February 16, 2017

SIGNAL RESEARCH AND MULTIFACTOR MODELS

QES Handbook of Active Investing, Part II

- **From Data to Knowledge.** This paper is Part II of our initial launch and the *QES Handbook of Active Investing*. In Part I – *The Big and the Small Sides of Big Data*, we discuss the vast contents of alternative data and data science. In this research, we elaborate on how to translate data into stock-selection signals. In the next few weeks, we will release Part III (*Style Rotation, Machine Learning, and the Next Frontier of Systematic Investing*) and Part IV (*Risk, Portfolio Construction, Trade Execution, and Performance Attribution*).
- **Signal Research.** In the first section of this paper, we review the methodologies behind single factor backtesting. Different approaches sometimes lead to conflicting views. We discuss the pros and cons of each method. In particular, we highlight one of the most damaging biases in investment research – the sin of story telling – once you find a pattern, you are obligated to come up with a story to explain it.
- **Review of Common Stock-Selection Factors**. We examine common stock-selection factors in seven categories (value, growth, momentum/reversal, quality, alternative, and Big Data). We provide a few interesting examples of how to improve our traditional factors and then move on to factors based on unconventional data such as news sentiment and securities lending.
- **Factor Selection, Signal Weighting, and Multifactor Models.** We inspect the details of factor selection (aka feature selection in machine learning language) via two broad methods – wrapper and filter. We use real-life simulations to assess the sensitivities of various factor selection algorithms to multicollinearity and noisy signals. Furthermore, we compare and contrast various factor weighting schemes (e.g., equally weighted, global minimum variance, regression based, Grinold & Kahn optimization based, and risk parity). We show how the alternative beta (also called risk premia or factor investing) approach matches up multifactor models. Lastly, we demonstrate how to incorporate downside risk, tail dependence and strategy crowding in the factor weighting decision process.

Source: Wolfe Research Luo's QES

Yin Luo, CFA, CPA
YLuo@wolferesearch.com

Javed Jussa
JJussa@wolferesearch.com

Sheng Wang
SWang@wolferesearch.com

QES Desk Phone: 1.646.582.9230
Luo.QES@wolferesearch.com

Table of Contents

A LETTER TO OUR READERS

Welcome to the second part of the QES Handbook series

Last week, we launched our QES research with the first paper on the Big Data evolution – we introduced the paradigm shift in the active management industry, the various conventional and non-traditional data contents, and data modeling (see Luo, et al [2017]). This paper extends what we wrote last week to the next level – how to translate data into knowledge.

Quantitative equity investment strategies often use factor based models. A factor can be broadly defined as any variable that is believed to be valuable in ranking stocks for investment and in predicting future returns or risks. Value and momentum are the classic examples. Many fundamental investors also implicitly or explicitly use factors in their investment process. For example, screening is widely used to identify potential investment opportunities.

In this paper, we start from the methodologies of conducting single factor backtesting. There are many ways to determine whether a specific investment strategy has produced excess return in the past. If it does, the hope is that it might also deliver alpha in the future. The problem is that these techniques often generate conflicting results. In particular, we highlight one of the most damaging biases in investment research – the sin of story telling – once you find a pattern, you are obligated to come up a story to explain it. In practice, portfolio managers and asset owners almost always demand an explanation before they invest in a strategy, as often the most important part of their due diligence process. The truth is whether the pattern will persist or not in the future has absolute nothing to do with how well we tell the story.

In the second part of this paper, we review some of the common factors used in stock selection. We break them down into seven style buckets – value, growth, price momentum/reversal, sentiment, quality, alternative, and Big Data. There are tremendous upsides even with traditional signals. The asset growth anomaly is such a classic example. Despite that it is being well covered in academic research, we can yet almost double the performance by making a small adjustment. We also demonstrate a few factors using unconventional data sources, e.g., news sentiment and securities lending.

The last part of this paper should be of the most interest to the majority of our readers. Now we have hundreds of these factors. Few managers use single factor models. However, at the same time, few investors include hundreds of factors in their models either. There are multiple hurdles from signal research to multifactor models. Factor selection (in machine learning literature, it is called feature selection) is often the first step. Discretionary factor selection is still the dominant approach in the industry. There is little in academic research on this topic. We introduce two broad machine learning based factor selection algorithms – wrapper and filter. Using real-life examples, we show the sensitivities of each feature selection technique on multicollinearity and noise data.

Furthermore, once we decide the factors to include in the model, how to weight them is still extremely important. We compare and contrast various common factor selection and factor weighting algorithms. More importantly, we use a series of practical examples and simulations to demonstrate the pros and cons of each approach – all on a global basis. In addition, we introduce one of our proprietary factor weighting techniques that incorporates extreme downside risk, tail dependence, and strategy crowding.

The next paper will discuss a series of more advanced topics. First, we will demonstrate the needs and the potential upside from factor timing, style rotation, and macro overlay. Then, we introduce machine learning into investment management. Almost all machine learning techniques are designed for cross-sectional data; therefore, we can't directly port these algorithms in finance. More importantly, we will elaborate our first global stock-section model – the LEAP model that incorporates all the data and techniques we have shown in this series of papers. The last paper in this series will examine practical portfolio implementation issues, e.g., risk models, transaction costs, constraints, performance attribution and active hedging. In a separate research series, we will discuss our framework for global macro, economics, and portfolio strategy research.

We hope you enjoy this paper. Any feedback and suggestion are more than welcome!

Regards,

Yin, Javed, Sheng, Kartik, and Luo's QES team

SIGNAL RESEARCH

Quantitative equity investment strategies often use factor based models. A factor can be broadly defined as any variable that is believed to be valuable in ranking stocks for investment and in predicting future returns or risks. Value and momentum are the classic examples. A factor strategy aims to identify significant factors that drive stock prices, and aim to construct a portfolio to have a positive bias towards such factors.

Many fundamental investors also implicitly or explicitly use factors in their investment process. Screening is widely used to identify potential investment opportunities. It is especially used by small-cap and global equity managers, because the investment universe is too large to conduct detailed fundamental analysis on all companies. In our previous research using both fund return and fund holding data, we find almost all fundamental portfolios have significant exposures to many of the common quantitative factors. In the end, it is not easy to find a manager who consistently bets against value and momentum.

In this section, we provide a broad overview on single factor backtesting methodologies, which sometimes deliver confliction results. We discuss the pros and cons of each approach and then give practical guidance.

THE SIN OF STORY TELLING

One of the very first questions about any factor is the direction. Are stocks with higher exposures to a factor more likely to lead to superior or inferior returns? For example, do we prefer companies with higher asset growth or bet against them? You may think it is obvious. Well, for some factors, it might be, but for many others, the empirical result may surprise you. Moreover, depending on the time horizon, the investment universe, and the backtesting methodology, the conclusion might be vastly different.

Let us take a simple example and use one of the foundations of modern finance – the book-to-market factor. This is the inverse of the price-to-book multiple[1] (price divided by book value per share) and one of the three factors in the Fama and French (see Fama and French [1993, 1996]) model[2]. A hedged portfolio that buys companies with the highest book-to-market ratios (i.e., value stocks) and simultaneously shorting the stocks with the lowest book-to-market (i.e., growth stocks), in theory, should produce positive returns in the long run. As shown in Figure 1 (A), the empirical result in the US does confirm our intuition. However, when we try to extend the study to the Australia and New Zealand market, we are surprised to see the opposite. Similarly, as documented in Jegadeesh [1990], the short-term mean reversal (defined as the total return of last month) depicts that investors should go against stocks with the biggest rallies in recent month, because these stocks are more likely to mean revert to lower levels. Figure 1 (B) shows a clear mean-reversal pattern in Japan. However, in AxJ, we see exactly the opposite.

Even in the same region, a factor may completely flip the sign during different periods. Figure 1 (C) shows the performance of the consensus five-year expected earnings growth factor in Europe in the

[1] Fundamental analysts tend to use the price-to-book or price-to-earnings multiples, while quantitative managers mostly use the inverse – book-to-market and earnings yield. As shown in Luo, et al [2017], there are a number of benefits of using the yield rather than multiple factors. Book values per share, for example, are mostly positive, but some firms do have negative book value at times. Therefore, using the yield factors generally lead to a larger breadth. The performance of yield factors also tends to be stronger.

[2] Fama and French call it HML (high minus low). We will explain the exact backtesting methodology shortly.

1980s-2000 era. High growth companies earned considerable premium in the technology bubble period. However, in the subsequent years from 2001 to present, the signal turned to the opposite direction, i.e., growth style plunged and that decline lasts for more than 15 years. We can easily come up with stories to argue either way. For example, we can attribute the growth phase in the 1990s to the technology bubble – when investors were irrational and chase high flying dot com stocks, these stocks kept going higher. Then, for the next 15 years, we can cite behavioral argument – investors like high growth and speculative stocks, which bids up the price and leads to lower returns in the future. This is what we call the "sin of story telling".

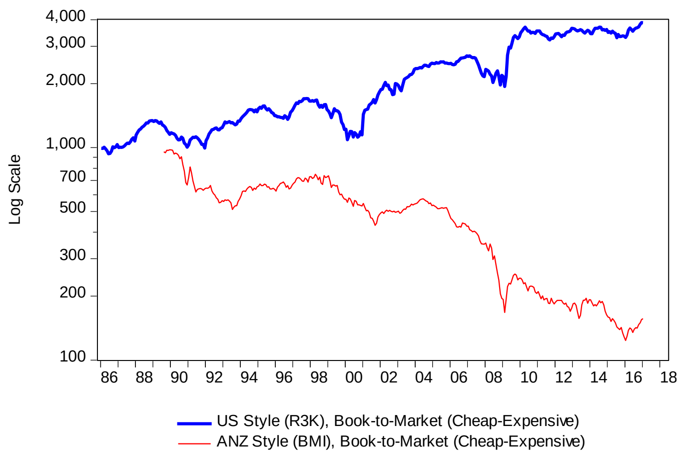
Admit it or not – we all love stories. Think about how many times that you went to a presentation and still remembered the numbers you saw in the presentation, even just a few days later. And then think about many long you still remember a funny story you heard during a lunch meeting. Human brains are designed to handle stories much better than numbers.

In the investment industry, one of the most important considerations of whether we should invest in a strategy (or a factor) rests on the intuition test. Similarly, whether you can get your paper published on a top tier journal crucially depends on your ability to explain your empirical results. However, we can't stop asking – does it really matter? Once we find a strong pattern, coming up with a story to explain it is easy. Whether the same pattern will repeat itself in the future has nothing to do with how well we tell our story, or does it?

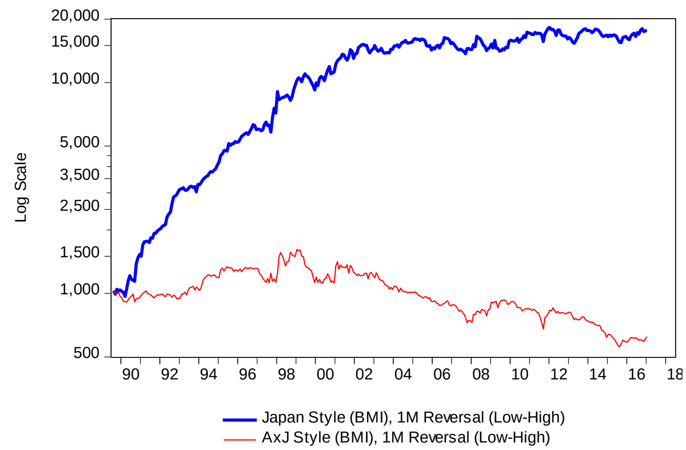
For our research, we have decided not to flip the direction of factors back and forth. Unless separately disclosed, most charts in this research assume that we overweight stocks with higher exposures (or scores) to a factor and short firms with lower scores. Therefore, when you see a consistent downward trend, or persistent negative returns (or ICs), it means lower values in that factor tend to be associated with higher subsequent returns.

Figure 1 The direction of a Factor may not always be what you think

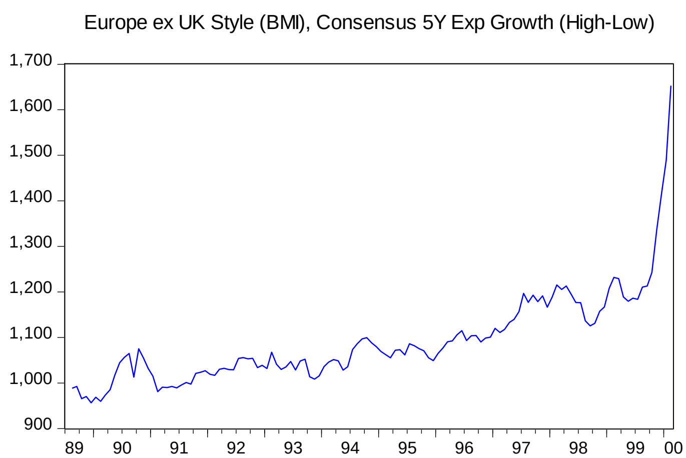
A) Book-to-Market in the US and ANZ



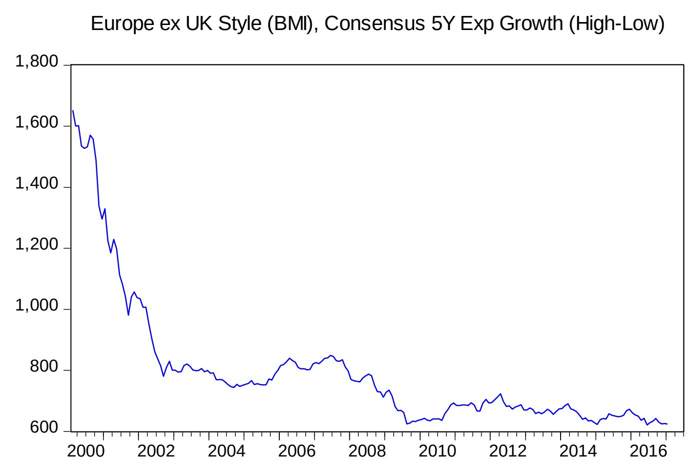
B) Short-Term Reversal in Japan and AxJ



C) Consensus 5Y EPS Growth in Europe, 1980s-2000



D) Consensus 5Y EPS Growth in Europe, 2001-Present



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

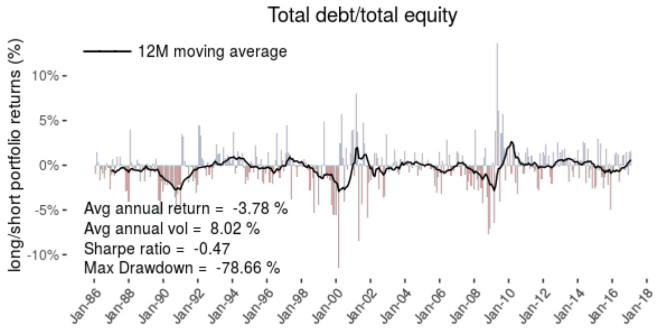
Alpha versus Risk Factors

Some analysts further differentiate alpha factors from risk factors. Traditionally, if a factor can explain cross sectional stock returns, it is potentially an alpha factor. However, if the factor return volatility is significant, it is considered as a risk factor. Figure 2 (A) shows an example of a risk factor – the debt-equity ratio. Although the average excess return is at a modest 3.8% per annum, the volatility of the factor is consistently high at 8.0%.

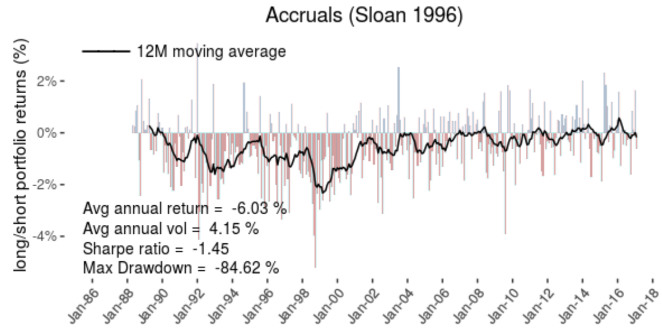
The requirement for an alpha factor is stricter – the factor performance also needs to be consistent. For example, the accounting quality factor based on Sloan's accruals concept (see Sloan [1996]) has more consistent return, with a Sharpe ratio of 1.45x – nearly three times higher than the debt/equity ratio factor (see Figure 2 B). In another words, while risk factors only need the unconditional variance to be significant, alpha factors also need to have a significant unconditional mean.

Figure 2 Alpha versus Risk Factors, US

A) Potential Risk Factor – Debt/Equity Ratio



B) Potential Alpha Factor – Sloan's Accruals



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

In recent years, the short term performance of most traditional alpha factors has shown substantial decay, likely due to arbitrage or changes of market regimes. Hence, taking a long-term and relatively static exposure to risk factors has become a popular investment theme. This is often called risk premium, smart beta or alternative beta investing.

While factor returns become more time varying, it reduces the payoff from static allocations. On the flip side, it does introduce opportunities for factor timing. We will discuss factor timing or style rotation in more details in a forthcoming paper.

SINGLE FACTOR BACKTESTING

Once a potential factor is identified, a portfolio manager needs to perform a series of tests to assess the factor's predictive power and efficacy. Some investors argue that the most important test is the "smell" test – does the factor make intuitive sense? A factor can often pass statistical backtesting, but if it does not make common sense and justification for the factor's efficacy is lacking, the manager may be data mining. The economic intuition test, however, lacks theoretical rigor and defeats the whole argument of being systematic. It is impossible to differentiate what accounts for being intuitive and what does not.

Investors therefore always need to remind themselves that what appears to be impressive performance from backtesting does not necessarily imply that the factor will continue to add value in the future.

Long/short Hedged Portfolio Approach

The most traditional and widely used method for implementing factor based portfolios is the hedged portfolio approach, pioneered and formulated by Fama and French [1993, 1996]. In this approach, having chosen the factor to be scrutinized and having ranked the investable stock universe by that factor, analysts divide the universe into groups referred to as quantiles (typically in deciles, quintiles or terciles) to form quantile portfolios. Stocks are either equally weighted or capitalization weighted within each quantile. A long/short hedged portfolio is typically formed by going long the best quantile and shorting the worst quantile. The performance of the hedged long/short portfolio is then tracked over time. This is the most intuitive approach, as the backtesting resembles a real-life investment strategy.

There are a few drawbacks to this approach. First, the information contained in the middle quantiles is wasted, as only the top and bottom quantiles are used in forming the hedged portfolio. Second, it is implicitly assumed that the relationship between the factor and future stock returns is linear (or at least monotonic) which may not be the case. Third, if many managers use similar factors, the resulting portfolios can be concentrated in specific stocks. Fourth, the hedged portfolio requires managers to short stocks. Shorting may not be possible in some markets or overly expensive in others. The hedged portfolio does not take into account of transaction costs, liquidity, and other institutional constraints; therefore, the performance is almost always overly optimistic. More importantly, the hedged portfolio is not a "pure" factor portfolio, as it has significant exposures to other risk factors.

Figure 3 (A) shows the performance of the current ratio (current assets/current liabilities) factor in Europe. The bars in the chart indicate the monthly portfolio returns (buying companies with highest financial liquidity ratios and shorting companies with the worse). The average annual return of the strategy is about -0.6% with a Sharpe ratio of -0.1x over the test period. The result is counter-intuitive and is inconsistent over time.

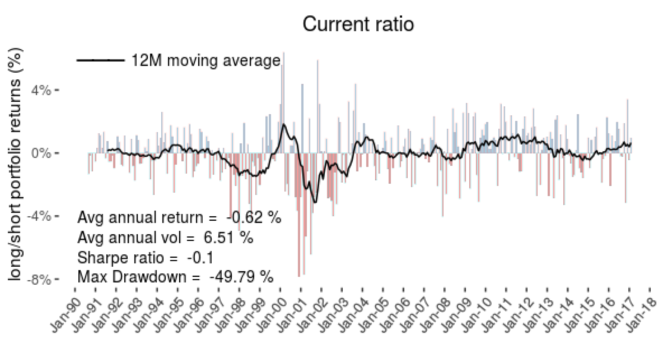
Figure 3 (B) shows the average monthly returns of the five quintile portfolios. Actually, other than companies with the highest liquidity, it forms a monotonic upward trend, suggesting financial liquidity is rewarded with higher subsequent returns. Too much financial leverage exposes a company to potential distress risk, but holding too much current assets may indicate either the company is not efficiently utilizing the "free" financing with current liabilities or holding too much unsold inventory. A simple long/short hedged portfolio approach would not reveal this important information.

For investors who desire a long-only factor portfolio, a commonly used approach is to construct a factor tilting portfolio, where a long-only portfolio with exposures to a given factor can be built with a target tracking error. The factor tilting portfolio tracks a benchmark index closely, but also provides exposures to the chosen factor. As such, it is similar to an enhanced indexing strategy.

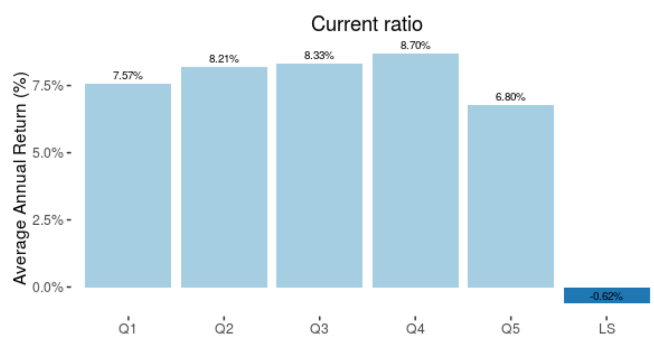
"Factor mimicking portfolio" or FMP, is an implementation method for a pure factor portfolio. FMP is a theoretical long/short portfolio that is dollar neutral, with a unit exposure to a chosen factor and no exposure to other factors. As FMPs are often too expensive to trade, managers typically construct the "pure" factor portfolio by following the FMP theory, but adding trading liquidity and short availability constraints in the portfolio construction.

Figure 3 Current Ratio, Europe

A) Long/Short Hedged Quintile Portfolio Return



B) Average Return, by Quintiles



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

Pearson and Spearman Rank IC

The Information Coefficient or IC is more commonly used by practitioners than the hedged portfolio approach as the standard measure of skill for factor performance in backtesting. Because most quantitative models are linear, IC captures the entire spectrum of stocks while long/short quantile portfolios only focus the top (and bottom) extremes. IC is generally considered to be a better measure with which to identify factors than the quantile-based hedged portfolio approach described earlier.

Pearson's IC is the simple correlation coefficient between the factor scores for the current period (e.g., earnings yield of all the stocks in our universe) and the next period's stock returns. As it is a correlation coefficient, its value always lies between -100% and 100%. The higher the IC, the higher the predictive power of the factor for subsequent returns is. In practice, any factor with an average monthly IC of 5%-6% is considered exceptional. The Pearson IC is sensitive to outliers as is illustrated below.

$$PearsonIC = @cor(f_t, r_{t+1})$$

A similar, but more robust measure is Spearman's rank IC, which is often preferred by practitioners. Spearman rank IC is essentially the Pearson correlation coefficient between the ranked factor scores and ranked forward returns.

$$SpearmanRankIC = @cor\big(@rank(f_t), @rank(r_{t+1})\big)$$

In the example shown in Figure 4, Pearson IC is negative at -0.8%, suggesting that the signal did not perform well, and was negatively correlated to the subsequent month's returns. Looking more carefully, however, it can be seen that the factor is generally in line with the subsequent stock returns, with the exception of stock I, where the factor predicts the highest return, while the stock turns out to be the worst performer. A single outlier can therefore turn what may actually be a good factor into a bad one as the Pearson IC is sensitive to outliers. In contrast, the Spearman rank IC is at 40.0%, suggesting that the factor has strong predictive power of subsequent returns. If three equally weighted portfolios (i.e., tercile portfolios) were constructed, the long basket (stocks G, H, and I) would have outperformed the short basket (stocks A, B, and C) by 56bps in this period. Therefore, in this case, Spearman rank IC is consistent with long/short portfolio, but Pearson IC is not.

Figure 4 Pearson Correlation Coefficient and Spearman Rank IC

Stock	Factor Score	Subsequent Return	Factor Score Rank	Return Rank
A	(1.45)	(3.00%)	9	8
B	(1.16)	(0.60%)	8	7
C	(0.60)	(0.50%)	7	6
D	(0.40)	(0.48%)	6	5
E	0.00	1.20%	5	4
F	0.40	3.00%	4	3
G	0.60	3.02%	3	2
H	1.16	3.05%	2	1
I	1.45	(8.50%)	1	9
Mean	0.00	(0.31%)		
Standard Deivation	1.00	3.71%		
Pearson IC		(0.80%)		
Spearman Rank IC				40.00%
Long/short Tercile Portfolio Return				0.56%

Sources: Wolfe Research Luo's QES

Univariate Regression

Another popular way to assess factor performance is via univariate (or multivariate[3]) regression:

$$r_{i,t+1} = \beta_0 + \beta_1 f_{i,t} + \varepsilon_{i,t}$$

Where,

$r_{i,t+1}$ is the return of stock i, in the subsequent period $t+1$,

$f_{i,t}$ is the score of the factor for stock i at time t

The inference centers on whether β_1 is statistically significant.

In an OLS (Ordinary Least Square) regression, β_1 can be computed as:

$$\beta_1 = \frac{COV\left(r_{i,t+1}, f_{i,t}\right)}{VAR\left(f_{i,t}\right)} = \frac{Corr\left(r_{i,t+1}, f_{i,t}\right)Disp\left(r_{i,t+1}\right)}{Disp\left(f_{i,t}\right)}$$

Where,

$COV\left(r_{i,t+1}, f_{i,t}\right)$ is the covariance between factor and subsequent return,

$VAR\left(f_{i,t}\right)$ is the variance of factor scores,

[3] Multivariate regression will be discussed extensively in the multifactor model section.

$Corr(r_{i,t+1}, f_{i,t})$ is the correlation between factor and the subsequent return, i.e., IC,

$Disp(r_{i,t+1})$ is the cross-sectional dispersion of returns, and

$Disp(f_{i,t})$ is the cross-sectional dispersion of factor scores

When the factor is standardized, $Disp(f_{i,t}) = 1$. If we further standardize our forward returns, $Disp(r_{i,t+1}) = 1$, then the regression coefficient is equivalent to IC.

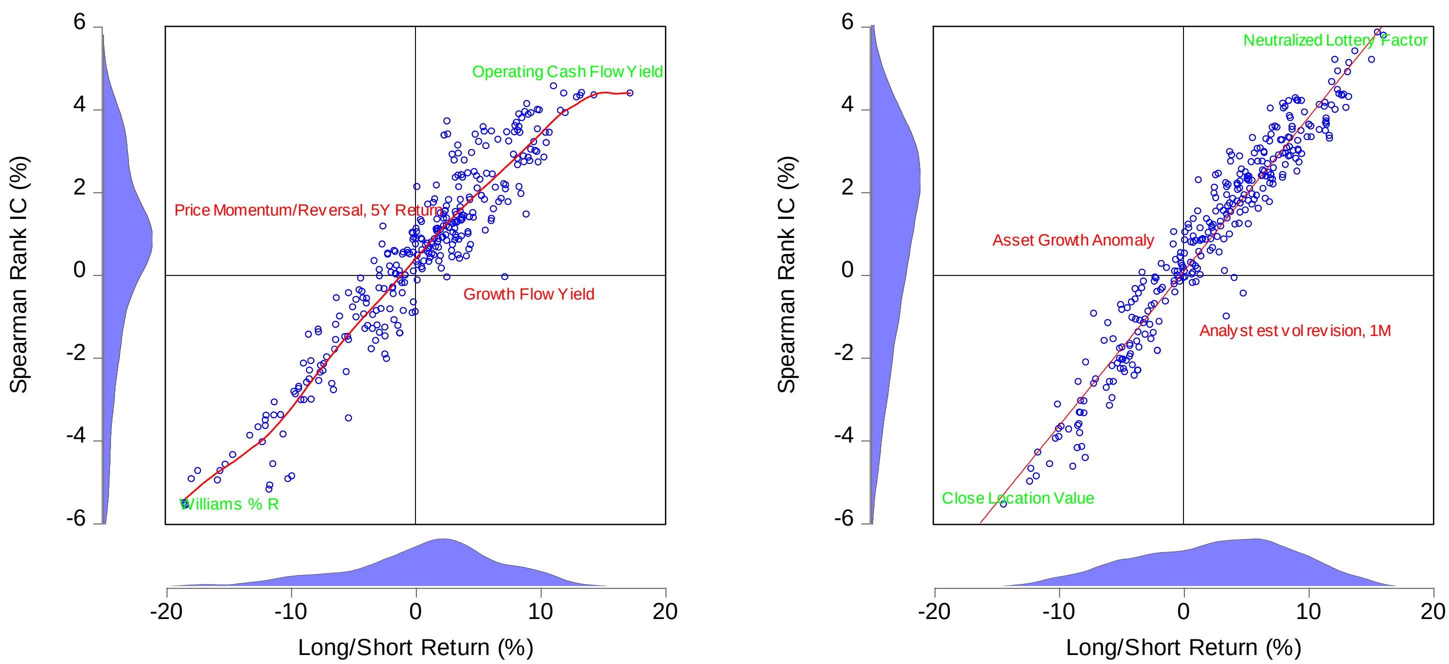
Do Different Backtesting Methodologies Tell the Same Story?

The big question is whether it makes any difference by choosing these different backtesting techniques. Using the ~400 factors in our factor library, we conduct an exercise. We backtest all of these factors in both US and Europe, using both long/short quintile portfolio and Spearman rank IC. As shown in Figure 5, in both US and Europe, these two approaches are generally consistent, but we can see noticeable differences and especially the divergence at the upper tail.

The most troubling part is the factors in the upper-left and bottom-right quadrants, representing those factors with opposite directions based on the two approaches. For example, in the US, the factor based on the past five-year return behaves like a price momentum (i.e., higher factor scores are correlated to higher subsequent month returns), using Spearman rank IC, but it is labeled as long-term reversal (i.e., the top quintile portfolio of stocks with higher scores underperform the bottom quintile), using the hedged portfolio approach. In the end, about 9% of factors in the US (see Figure 5 A) and 6% in Europe (see Figure 5 B) show different signs.

Figure 5 Long/Short Hedge Portfolio versus Spearman Rank IC

A) US **B) Europe**



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

Fundamental Law of Active Management

One of the building blocks of active investing is the fundamental law of active management (see Grinold [1989]):

$$IR = IC \times \sqrt{Breadth}$$

Where,

IR or information ratio represents the risk-adjusted performance of a portfolio,

$Breadth$ is the number of securities that we can express our views and normally assumed to be the same as N, i.e., number of stocks in our universe.

The fundamental law of active management states that the active performance of our portfolio is a function of stock picking skill (i.e., IC) and breadth (i.e., number of stocks in our coverage universe). Therefore, to improve our performance, we can either find an analyst or a factor with better stock-picking skills, or we can apply our skills to a broader investment universe.

There are a number of assumptions behind the law, but the most unrealistic one is that our skill, or IC is consistent over time. As discussed in the previous sections, in recent years, not only the average performance of most factors has declined, but also the volatility has increased. There are a number of studies attempting to extend the basic form of the law (see, Qian, et al [2007], Kroll, Trichilo, and Braun [2005], and Ding [2010], among many others).

Factor Coverage

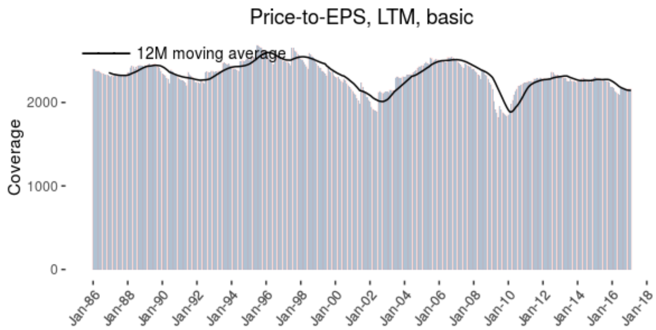
Based on the fundamental law of active management, everything else being equal, we prefer factors with broad coverage. One interesting example is the Price-to-Earnings multiple or PE. The PE ratio is probably the most commonly used valuation metric among investors – both professionals and retail investors regularly monitor the ratio before they make their investment decisions. However, there is a big flaw about the PE – you can't compute it if the denominator is zero or negative. Traditionally, most investors are long only, so it did not matter as much. For quantitative investors, we almost never use PE. Rather, we compute the inverse of PE, i.e., the earnings yield:

$$EarningsYield = \frac{EPS}{Price}$$

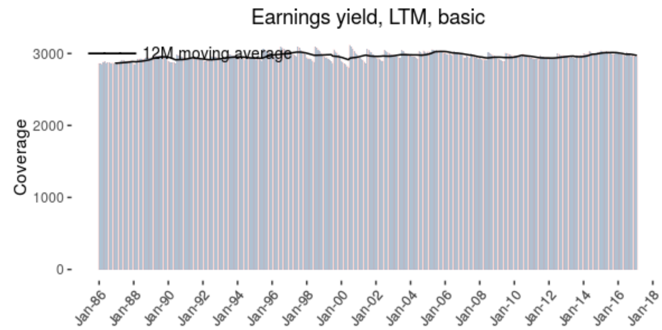
The earnings yield factor can be computed for any stocks with both EPS and price data (and price is almost never zero and can't be negative). As shown in Figure 6 (B), the coverage of earnings yield factor is about 27% higher than the PE ratio. Furthermore, the performance of the earnings yield factor is also 78% better, reflecting the strong information on the short side.

Figure 6 PE Ratio versus Earnings Yield

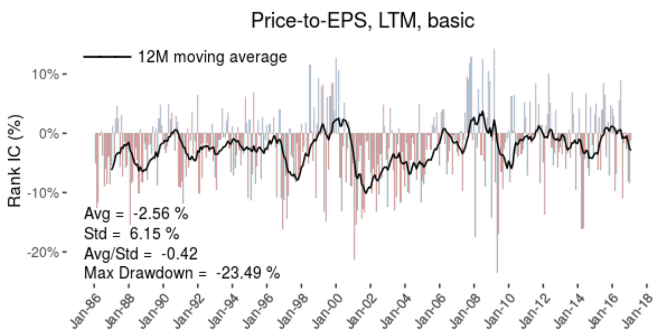
A) The Coverage of PE Ratio



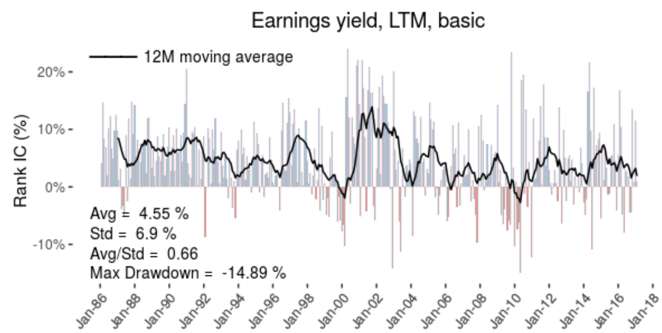
B) The Coverage of Earnings Yield



C) Rank IC, PE Ratio



D) Rank IC, Earnings Yield



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

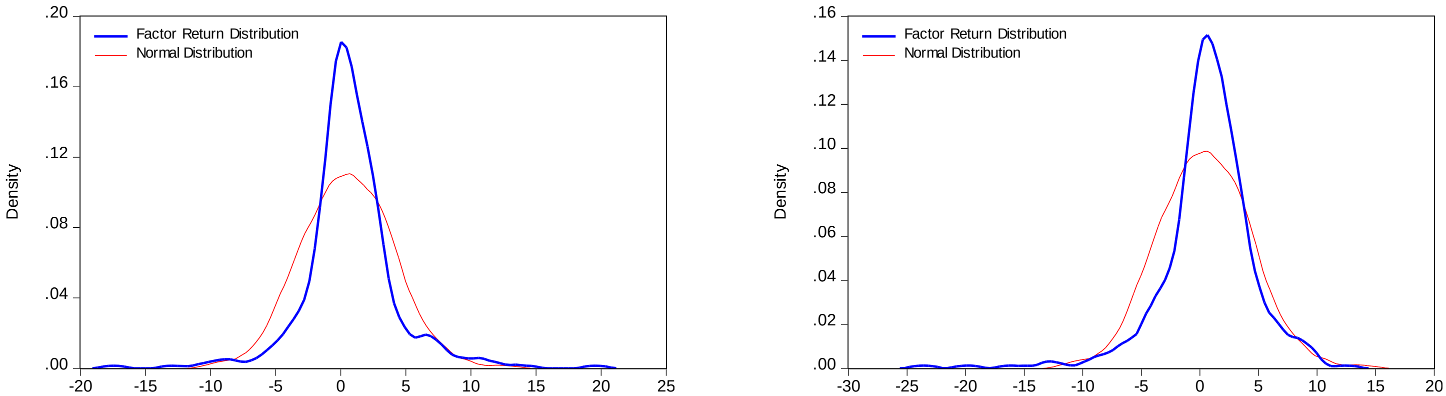
Factor Distribution

When we think about factor distribution, there are two issues that we need to address. One is factor score distribution. As discussed in Luo, et al [2017], if the distribution of scores is highly skewed with outliers, we need to transform the data. Otherwise, when we combine it with other factors, it may cause undesirable distortion to the resulting multifactor model. We have discussed how to deal with factor score distribution previously; therefore, we will zoom in the second issue of factor return distribution.

The traditional assumption of normal distribution on factor return is highly unrealistic. As shown in Figure 7 (A) and (B), the distribution of the classic value and momentum returns is clearly non-normal. More problematically, it tends to suffer from excess kurtosis and negative skewness. Excess kurtosis means that we are more likely to be surprised by extreme returns, while negative skewness hints that those surprises are more likely to be negative. We will deliberate the issue further in the multifactor model section later.

Figure 7 The Distribution of Factor Returns

A) Value (Earnings Yield) in the US **B) Price Momentum (12M-1M Total Return) in the US**



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

Factor Efficacy

It is a standard practice in academic research to show the statistical significance of a market anomaly over the entire history. To prove the robustness over time, it is also common to split the history into two or a few sub-periods. However, it is rare to see even a simple time series plot of the excess return over time.

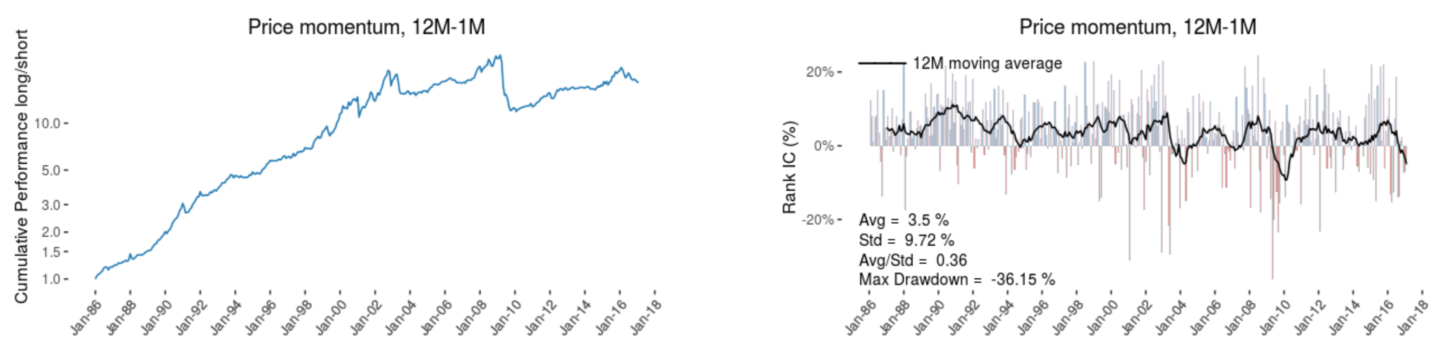
The classic example is price momentum – performance seems to have fallen in recent years with periodic significant drawdowns (see Figure 8). Therefore, in signal research, it is critical to check whether factor return is consistent over time. There are a range of econometric tools that can be used:

- Breakpoint regression to check for structural breaks;
- Markov regime switching model to examine potential regimes

We will discuss this topic further in a forthcoming paper, *Style Rotation, Machine Learning, and the Next Frontier in Systematic Investing*.

Figure 8 The Time Varying Nature of Factor Returns – Price Momentum

A) Cumulative Performance **B) Rank IC**



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

Factor Turnover and Decay

One of the most common questions that we hear from clients all the time is the turnover of a factor. The answer to that seemingly simple question, however, is rather complicated. The turnover of a portfolio depends on not only the frequency and magnitude of the change in factor scores over time, but also the portfolio construction process.

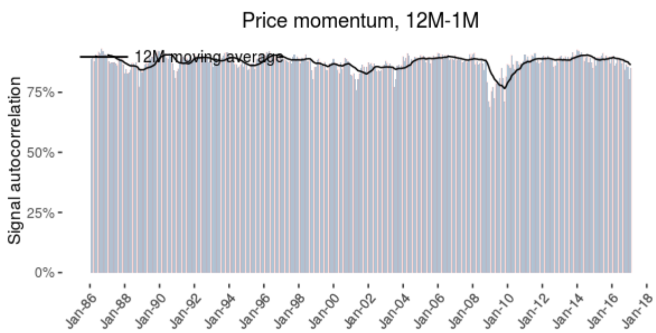
To isolate the effect of factor changes from portfolio construction, we use signal autocorrelation (serial correlation) to measure factor turnover. Signal autocorrelation is computed as the correlation between the vector of today's factor scores and the vector one period in the last:

$$SignalAutocorrelation_t = @correlation(f_t, f_{t-1})$$

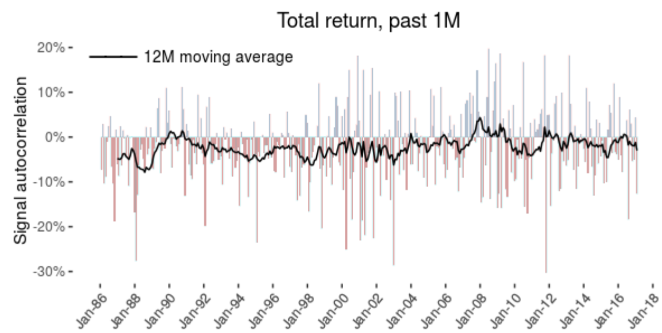
We then either plot the autocorrelation over time, or further compute the average. For example, Figure 9 (A) shows the signal autocorrelation for price momentum stays around 85%, indicating momentum is a low turnover strategy. On the other hand, mean-reversal strategy (based on past one month return) has a monthly serial correlation close to 0%, requiring almost a full rebalance each month (see Figure 9 B).

Figure 9 Signal Autocorrelation – Price Momentum versus Mean-Reversal

A) Price Momentum in the US



B) Mean-Reversal in the US

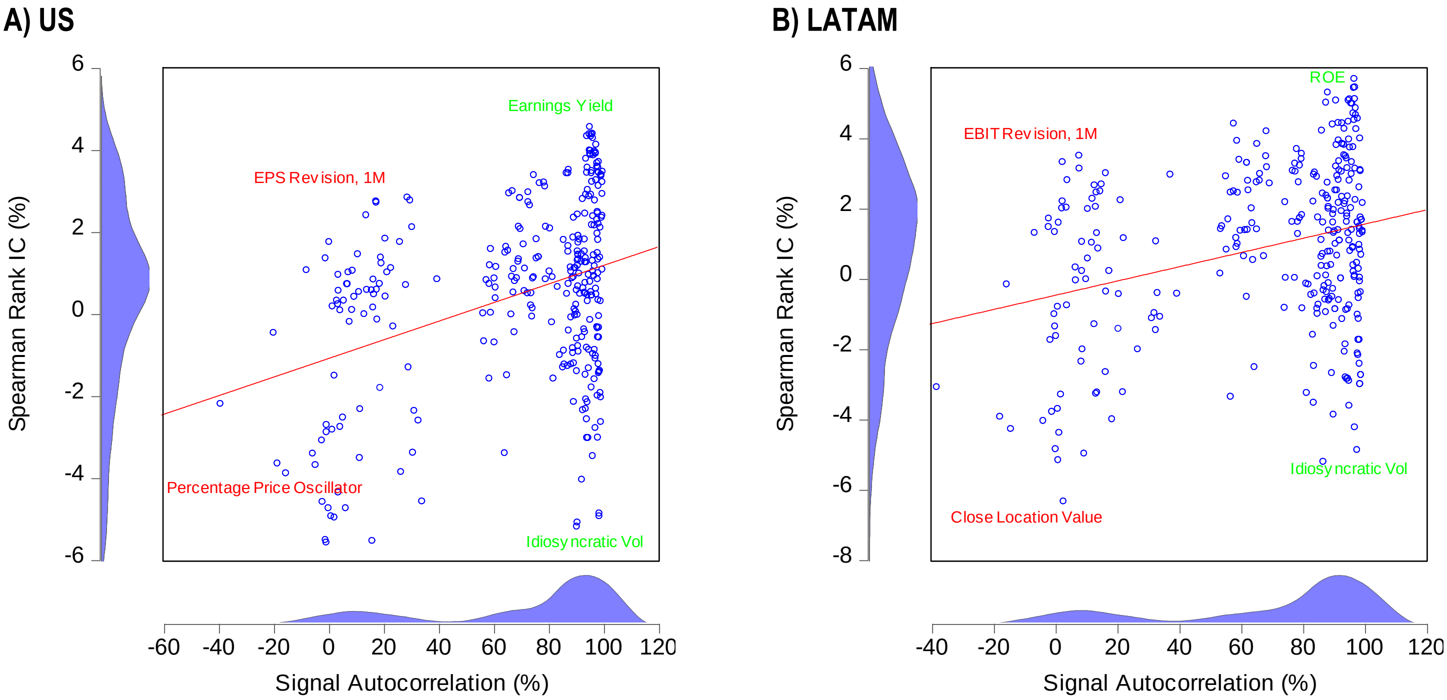


Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

Everything else being equal, we prefer factors with low turnover, i.e., high autocorrelation, because those factors lead to lower portfolio turnover in the end and therefore higher after-cost performance.

The month-to-month autocorrelation for most factors falls in the range of 80%-100%, in both the US and LATAM (see Figure 10). However, the distribution is clearly bi-modal, peaked at 80-100% and 5%-15% ranges. More importantly, we observe a weak positive relationship between signal serial correlation and performance, indicating superior factors are also likely to have modest turnover. This is particularly relevant for markets such as LATAM, emerging EMEA, and emerging Asia, where transaction costs tend to be much higher than in the US.

Figure 10 The Relationship between Factor Turnover and Factor Performance



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

A closely related, but different concept is information decay, which measures the decline in a factor's predictive power, as we extend the forecasting horizon. Essentially, we compute the Spearman's rank IC between today's factor scores and the next qth month return:

$$SpearmanRankIC_q = @cor\left(@rank(f_t), @rank(r_{t+q})\right)$$

We can then plot the IC decay chart to show how long the predicative power of our factor tends to last. For example, Figure 11 (A) shows the slow information decay for a value factor (earnings yield) – the predictive power remains positive even after a year. On the other hand, the mean-reversal factor[4] (based on past month return) is only predictive for the immediate subsequent month (see Figure 11 B).

[4] Please note that the relationship between mean-reversal signal and subsequent month's return is negative; therefore, in practice, we go long the stocks that had the largest rallies last month and short the ones with the lowest returns. We do not flip the sign of the factor in our backtest and that is why the IC is negative for the first month.

Figure 11 Signal Decay – Value versus Mean-Reversal

A) Value in the US

B) Mean-Reversal in the US

Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

COMMON STOCK-SELECTION FACTORS

Factors are the raw ingredients of quantitative investing and are often referred to as signals. Most quantitative managers spend the majority of their time studying factors. Traditionally, factors were primarily based on fundamental characteristics of underlying companies, however, many investors and asset owners have recently shifted their attention to unconventional and unstructured data sources in an effort to get an edge in creating strategies.

In this section, we review some of the common well-known stock-selection factors and show a few examples of why signals from unconventional Big Data can be useful to portfolio managers. In the coming months, we will shift our attention towards primarily non-traditional data sources and factors.

Where Do We Get the Ideas?

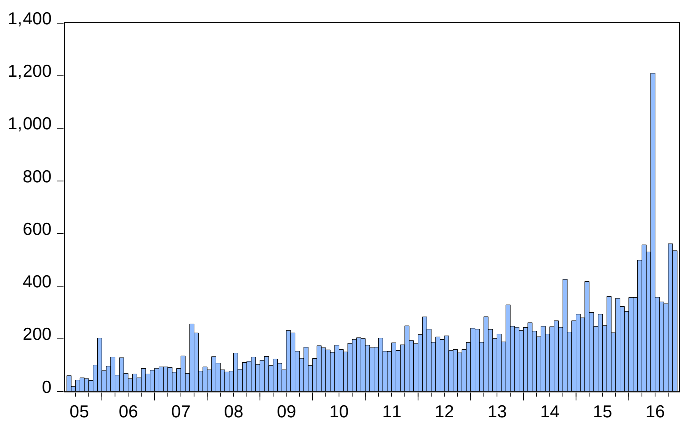
Factor strategies often rely on academic research in the search for new factor ideas. Given the lengthy review cycles for most academic journals, investors source more and more ideas from working paper repositories to access ideas well before the final articles are published. Research from the sell-side investment banks also provides interesting market color on factor performance and new factors.

There are thousands of scholars in academia sifting through tremendous amounts of information and hundreds of papers published in peer-reviewed academic journals around the world every day. More importantly, the idea of crowd sourcing seems to have also flourished in academia. Nowadays, most papers show up in working paper depositories (e.g., SSRN, RePEc, CEPR, NBER, EconPapers, etc.) long before they are published in the journals. The challenge is how to go through these gigantic piles of papers to find the information we need.

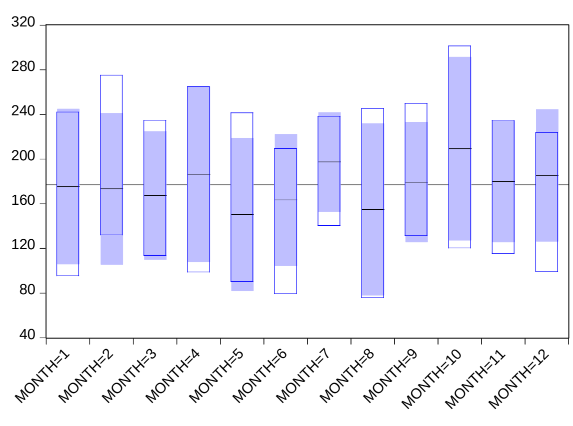
Figure 12 (A) shows the growth of the number of papers posted on SSRN related to investing – there are close to 600 investment related papers on SSRN alone. There are also strong seasonal patterns. As shown in Figure 12 (B), academic researchers seem to be particularly active in the summer and October. It is obviously quiet in the summer to conduct serious research. October coincides with job market papers for PhD students and the subsequent AEA/AFA conference in January.

Figure 12 # of Papers Posted on SSRN related to investing

A) The growth of # of papers on SSRN



B) Seasonality, # of papers posted by month



Sources: SSRN, Wolfe Research Luo's QES

We adopt a hybrid human-and-machine approach – a system we code name L-Scholar. We have developed a suite of web scraping algorithms to automatically pull together most of the published and working papers from various journals and working paper repositories. Then we store all these papers in a file-based database system, instead of a traditional relational database, for the ease and speed of textual data processing. We further develop a prototype natural language dictionary – similar to the famous Harvard psychology dictionary for general purpose Natural Language Processing (NLP). Our language dictionary is designed specifically for text processing of academic research papers relevant to investment management. We do not limit the papers only in the field of economics and finance, because a paper published in a computer science journal about machine learning can be equally or even more important.

The training of the NLP algorithms for research literature and the construction of our dictionary clearly required human intervention. Each of us on the team spends a considerable amount of time going through thousands of papers every month to manually pick the most interesting papers. Then we train our computer to proxy our selection process.

In the end, we publish a monthly piece called *Journal of QES* – a compilation of around 100 of the latest, most relevant and most interesting academic papers related to investing (see Luo, et al [2016] for a recent example). We hope that we are the best source of your signal research, especially on identifying new factors from both traditional and unstructured databases.

An Evolving Factor Library

We have over 400 relatively unique factors[5] in our global factor library. We further classify these factors into seven style categories:

- Value
- Growth
- Price momentum and reversal
- Analyst sentiment
- Quality
- Alternative
- Big Data

VALUE

Value is based on the Graham and Dodd's [1934] concept. Academic literature has a long history of documenting the value phenomenon. Basu (1977) found stocks with low PE or high earnings yield tend to provide higher returns. Fama and French (1993) formally outlined value investing by proposing book-to-market ratio as a way to measure value and growth.

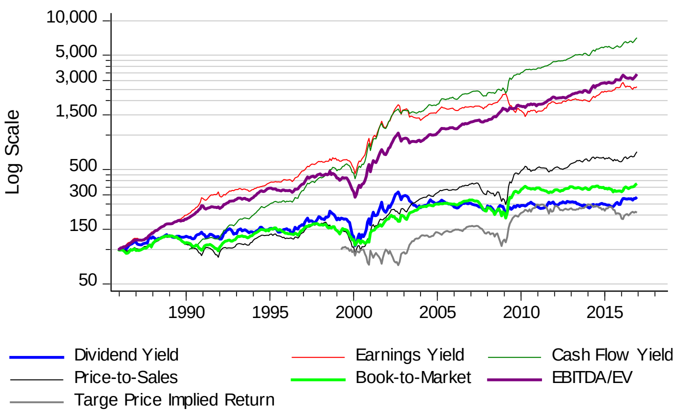
[5] From these 400 base factors, we can construct almost unlimited variations. For example, we only count one simple price momentum factor, based on the total return of the stock over the past 12 months, excluding the most recent month. However, you can compute the price momentum factor based on the past three, four, …, to 60 months. They are arguably highly correlated.

Although academics and practitioners agree that value stocks tend to deliver superior returns, they have considerable disagreements on the reasons. Fama and French [1993, 1996] suggest that the value premium exists simply to compensate for higher distress risk. Lakonishok, Shleifer, and Vishny [1994] cited behavioral arguments, suggesting the effect was a result of suboptimal behavior of the typical investor.

Value factors can also be based on other fundamental performance metrics of a company such as dividends, earnings, cash flow, EBIT, EBITDA and sales. Analysts often make two more variations on most value factors by adjusting for industry (and/or country) and historical differences. Most valuation ratios can also be computed using either historical (called trailing) or forward metrics. As shown in Figure 13 (A), the performance varies significantly among various value factors. The correlations among these factors are mostly modest. Target price implied return factor is even negative correlated to most of the other value factors.

Figure 13 Value Factor Performance

A) Different Faces of Value in the US



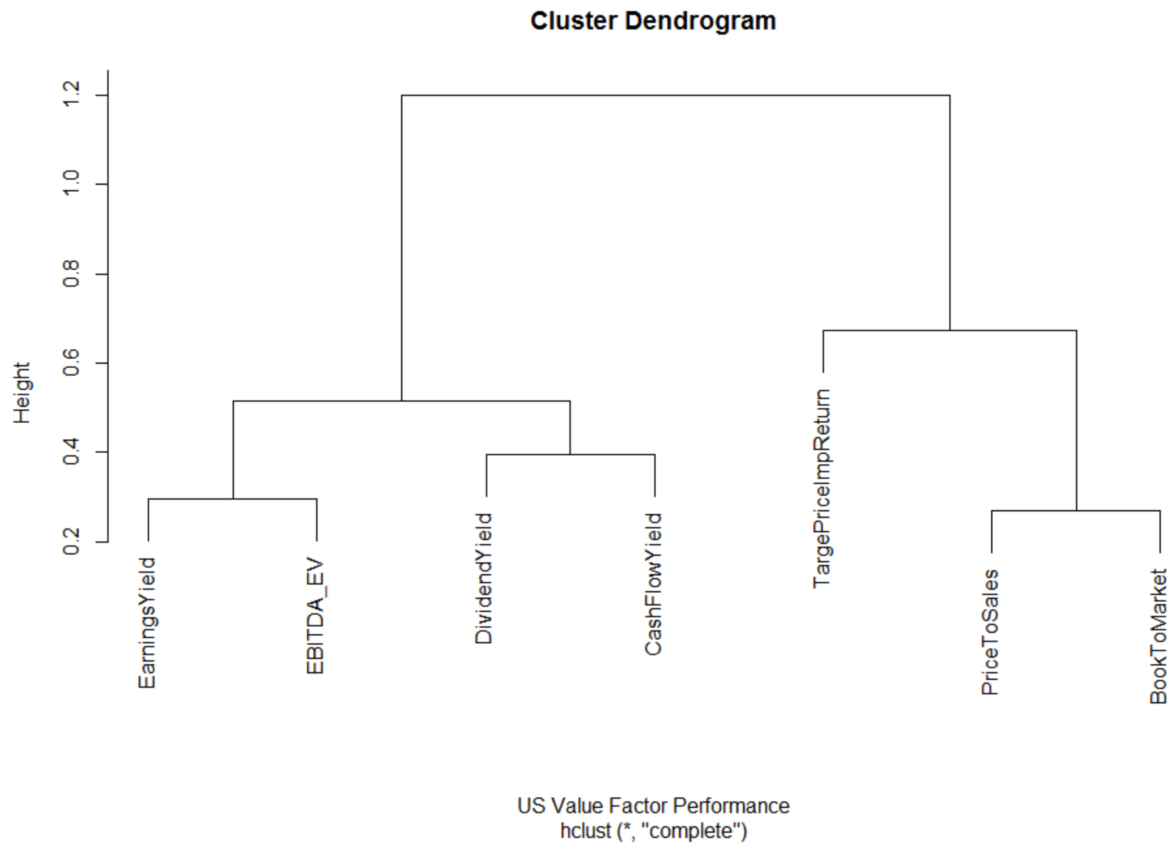
B) Correlation Matrix, Value Factors in the US

	Dividend Yield	Earnings Yield	CashFlow Yield	P/Sales	Book-to-Market	EBITDA/EV	TargetPrice Imp Return
DividendYield	100%						
EarningsYield	67%	100%					
CashFlowYield	81%	60%	100%				
PriceToSales	33%	10%	56%	100%			
BookToMarket	19%	27%	36%	90%	100%		
EBITDA/EV	79%	85%	81%	25%	4%	100%	
TargePriceImpReturn	42%	65%	21%	51%	70%	54%	100%

Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

As shown in Figure 14, value factors form two distinctive clusters, with flow-based variables on one side (e.g., earnings yield, EBITDA/EV, dividend yield, and cash flow yield) and stock-based variables on the other side (i.e., target price implied returns and book-to-market).

Figure 14 A Cluster Analysis of Value Factors

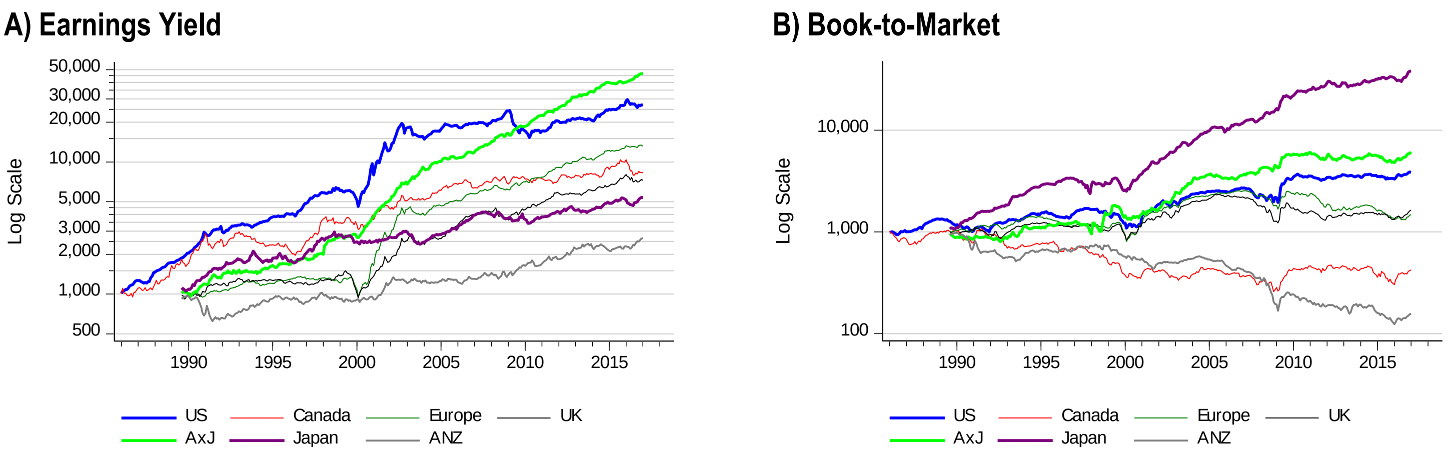


Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

Finding Value in the World

The performance of value factors around the world is vastly different. As shown in Figure 15 (A), the performance of earnings yield is exceptional in AxJ – in a region when typical investors chase for growth, value investing delivers superior results. On the other hand, book-to-market is disastrous in commodities-led markets such as ANZ and Canada, while it thrives in a consumption driven country like Japan (see Figure 15 B).

Figure 15 Value Factor Performance Globally



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

Valuation models, such as DDM, DCF, EVA, etc.

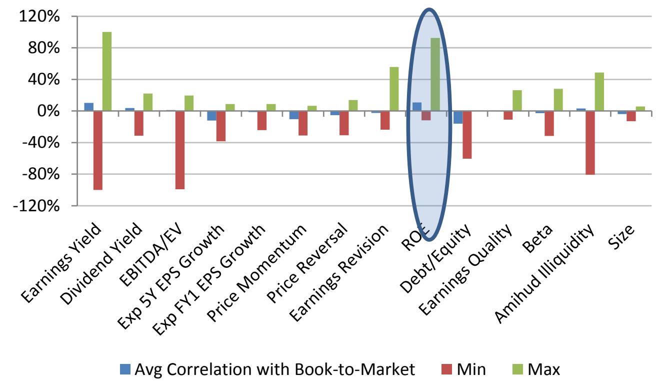
Fundamental analysis almost always starts from valuation and primarily focuses on valuation. There are many books about the many ways to conduct valuation. Damodaran [2012] offers a comprehensive treatment. The valuation theory generally provides a far richer treatment of valuing a company than the simple multiple/ratio approach introduced above. However, at the same time, valuation models also require a large number of assumptions, such as long-term growth rate, discount rate, etc. Most of these assumptions are difficult to estimate, suffer from large estimation errors, and are subjective in nature. More importantly, the valuation results tend to be fairly sensitive to the input assumptions. In an upcoming research, we will compare the performance of valuation models with value factors.

Fundamental Value

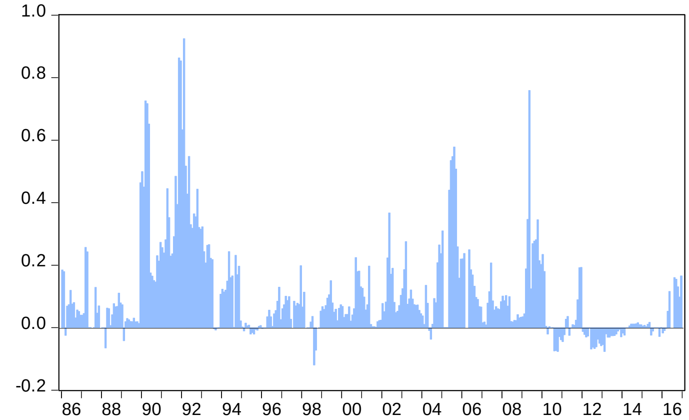
If we use book-to-market as an example, on average, it is positively correlated to earnings yield and ROE; and negatively correlated to earnings growth and price momentum (see Figure 16 A). However, the correlation can vary considerably over time. The maximum and minimum correlation range is rather wide. As shown in Figure 16 (B), the correlation between book-to-market and ROE has changed drastically over the past 30 years.

Figure 16 The Correlation between Book-to-Market and other Common Factors

A) Correlation with Book-to-Market: Average, Minimum, and Maximum,



B) Correlation between Book-to-Market and ROE over time



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

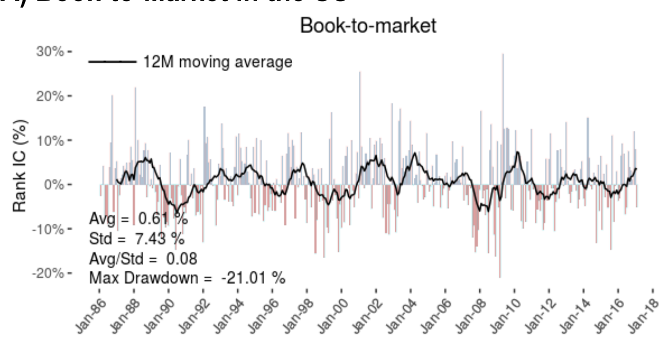
Therefore, it is natural to adjust valuation ratios on other fundamental factors, especially ROE. Figure 17 shows a simple example of adjusting book-to-market for ROE, by taking the residual ($\varepsilon_{i,t}$) from the following cross-sectional regression, at each point in time:

$$BookToMarket_{i,t} = \beta_{0,t} + \beta_{1,t} ROE_{i,t} + \varepsilon_{i,t}$$

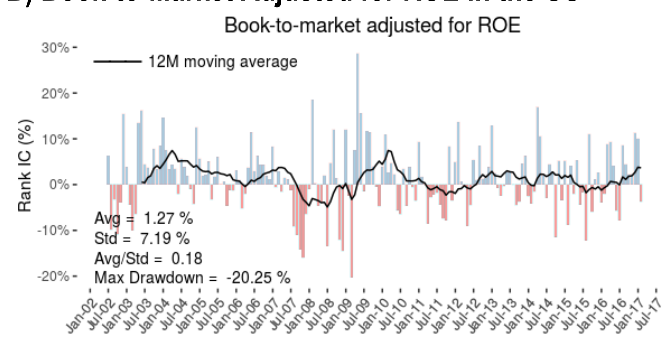
Adjusting for ROE more than double the performance of the book-to-market factor, albeit is still relative weak compared to other value factors.

Figure 17 Value Factor Performance

A) Book-to-Market in the US



B) Book-to-Market Adjusted for ROE in the US



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

GROWTH

Growth is another traditional investment style used in style investing. Growth factors try to measure a company's growth potential, and can be calculated using the company's historical growth rates, or forward projected growth rates. Growth factors can also be classified as short-term growth (last quarter's, last year's, next quarter's, or next year's growth) and long-term growth (last five year's or next five year's growth).

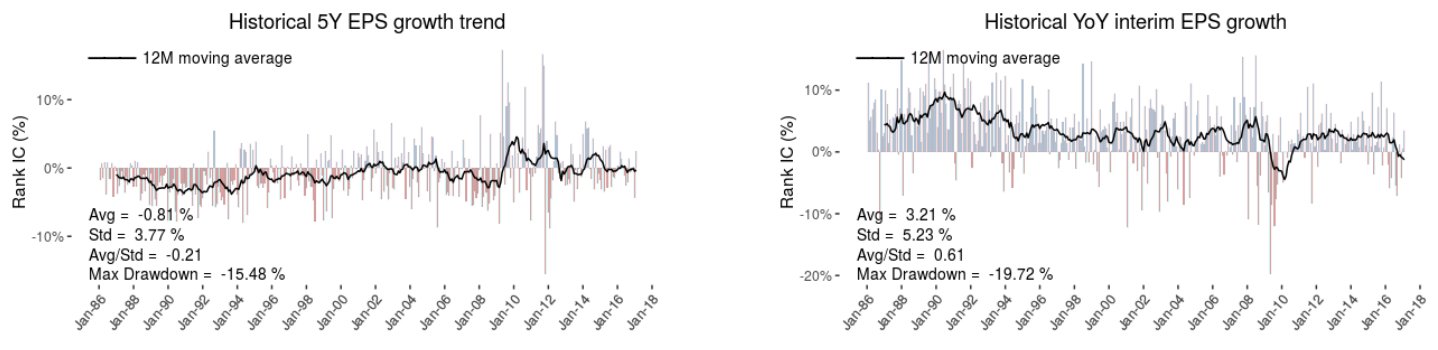
Historical Growth

Historical growth factors extrapolate past growth into the future. In a typical three-stage DCF (Discount Cash Flow) model, for example, it is common to assume the near term and mid-term growth rates to be in line with the recent growth profiles.

Factors that are based long-term historical growth show much weaker performance than short-term growth signals in the US (see Figure 18).

Figure 18 Historical Growth Factor Performance, US

A) Historical Five-Year EPS Growth **B) Historical YoY EPS Growth**



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

One argument about historical growth is that maybe acceleration is more than the simple growth rate. To measure acceleration, we perform the following regression, for each company, as of each point-in-time, using a rolling window:

$$EPS_{i,t} = \beta_{i,0} + \beta_{i,1}Trend + \beta_{i,2}Trend^2 + \varepsilon_{i,t}$$

Where,

$EPS_{i,t}$ is the earnings per share for company i, as of time t,

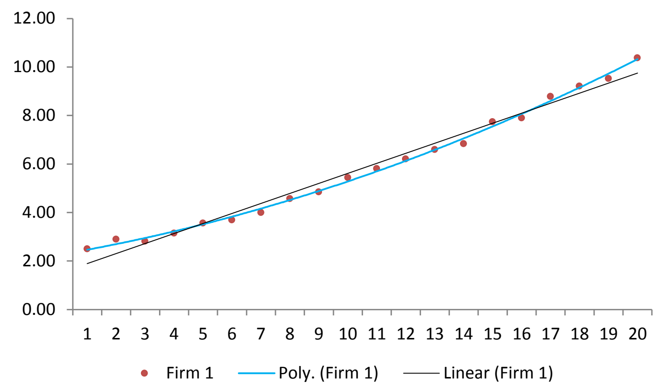
$Trend$ is a vector of time trend, e.g., $(1,2,\cdots,T)$

The estimated coefficient $\beta_{i,1}$ represents growth rate, while $\beta_{i,2}$ is the rate of acceleration.

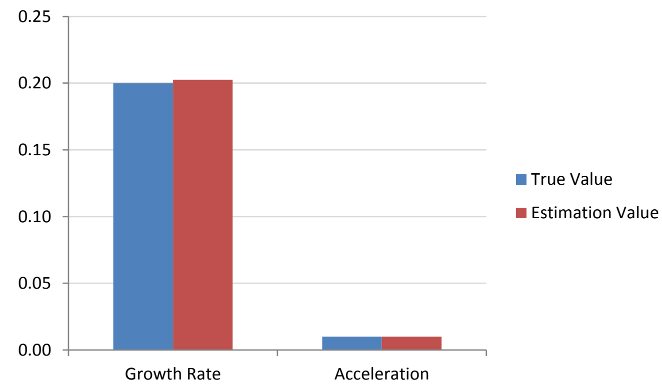
To show how the above procedure works, we design a simple simulation. We simulate two companies with 20 quarters of EPS data. Firm 1 has positive growth rate and positive acceleration (see Figure 19 A). We add some random noise to the EPS data. Figure 19 A shows the true and estimated β_1 and β_2 – the model fits the data very well. Similarly, Firm 2 had positive growth rate, but it is overshadowed by a negative rate of acceleration (see Figure 19 C and D). In this case, the estimated growth rate is quite different from the true value, but the rate of acceleration is fairly accurate.

Figure 19 A Simulated Example

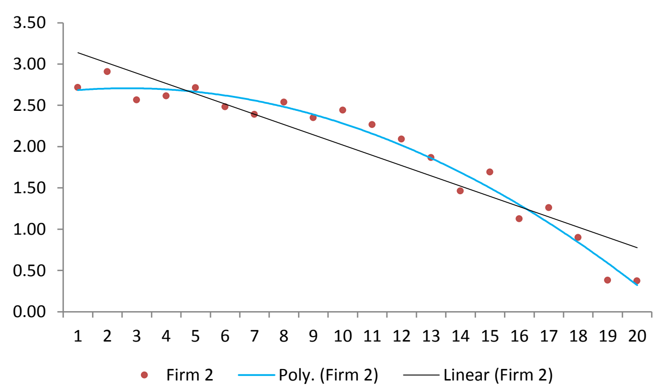
A) Positive Growth and Positive Acceleration



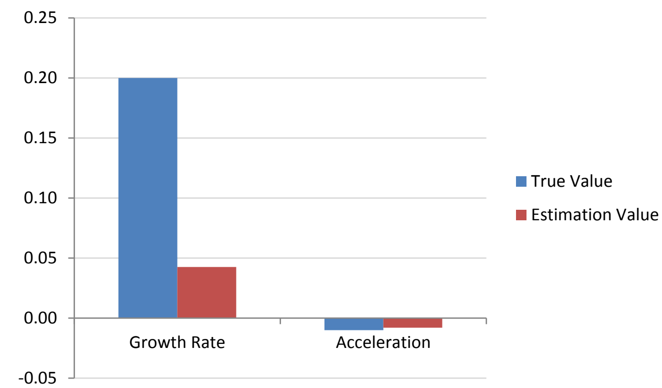
B) True and Estimated Parameters



C) Positive Growth but Negative Acceleration



D) True and Estimated Parameters

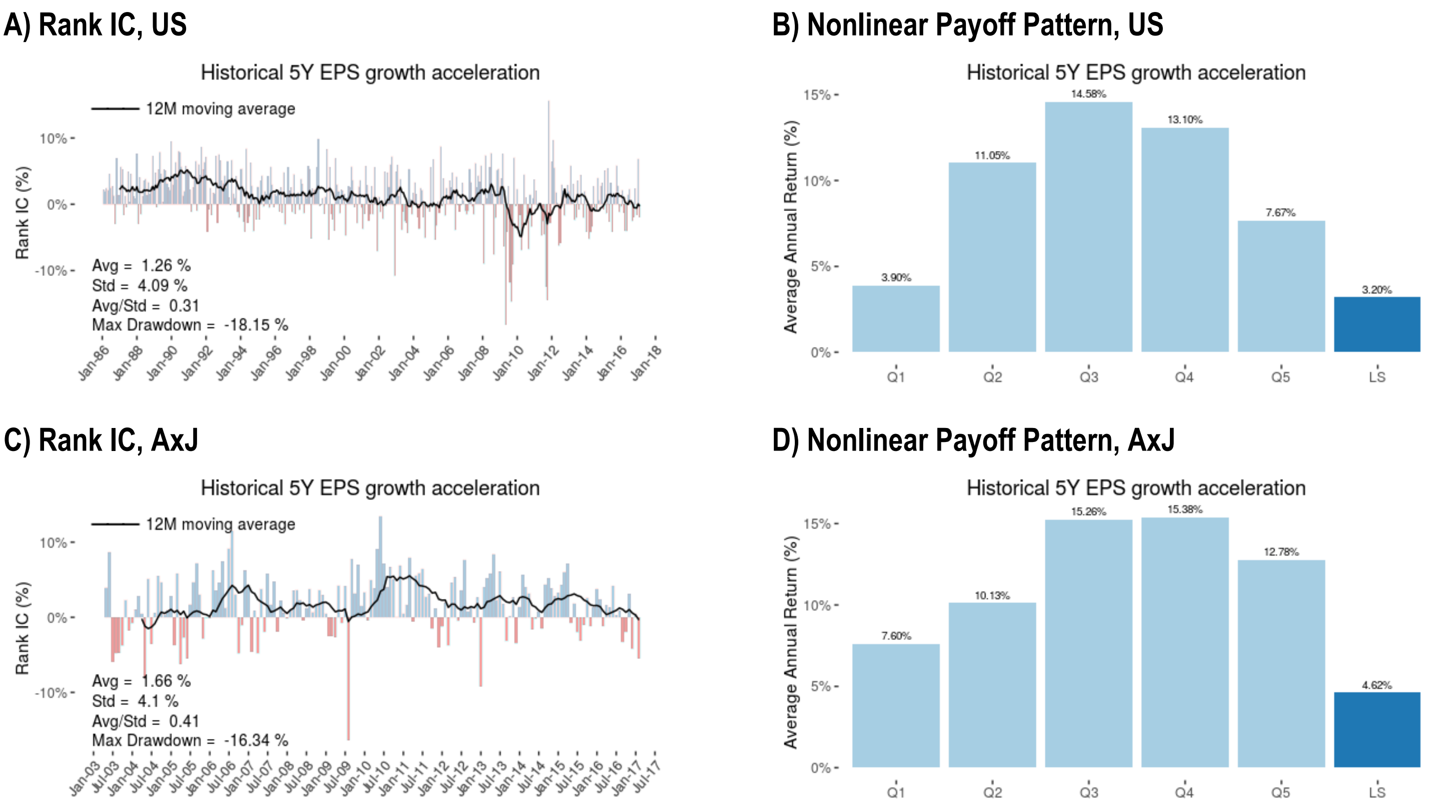


Sources: Wolfe Research Luo's QES

We estimate the above rate of acceleration for all companies globally. The performance of the factor is reasonable, but not spectacular (see Figure 20 A). However, a more careful examination of the payoff pattern reveals a very strong nonlinear shape – stocks with the highest and lowest earnings accelerations significantly underperform the middle quintiles (see Figure 20 B). This might reflect the fact that either high growth companies are too risky or the human behavior bias of bidding up growth companies beyond the level justified by their underlying fundamentals. Similarly, extremely low growth firms are likely to be the laggards. We observe similar patterns in AxJ (see Figure 20 C and D).

WOLFE RESEARCH

Figure 20 The Performance of Earnings Acceleration Factor



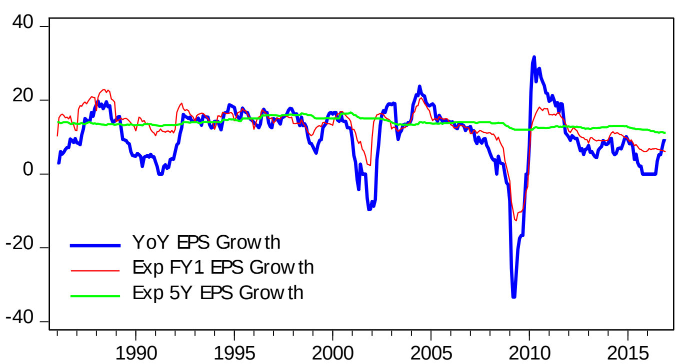
Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

Expected Growth

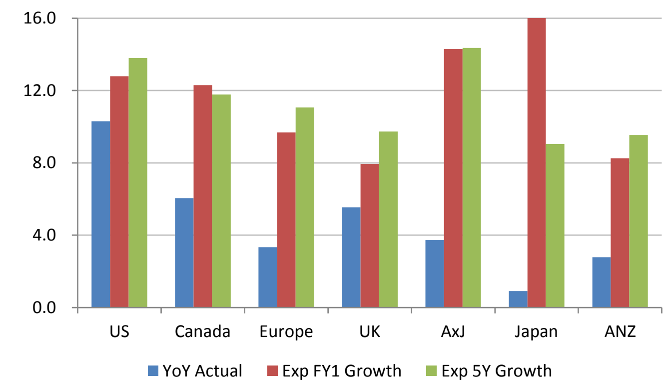
It is well known that sell-side analysts tend to be overly optimistic in their growth projections, especially on the long-term expectations. As shown in Figure 21 (A), analysts have been over-estimating both short-term and long-term earnings growth rates most of the time. Globally, analysts seem to be particularly complacent in Japan and Europe (see Figure 21 B). As a result, the performance of projected long-term earnings growth factor is dreadful in most regions (see Figure 21 C). Factors based on near-term growth expectations have fared better (see Figure 21 D).

Figure 21 Projected Growth Factor

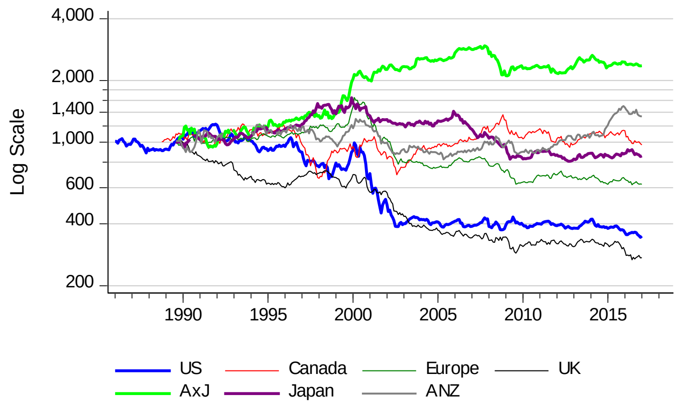
A) Projected Long-Term Growth vs Actual, US



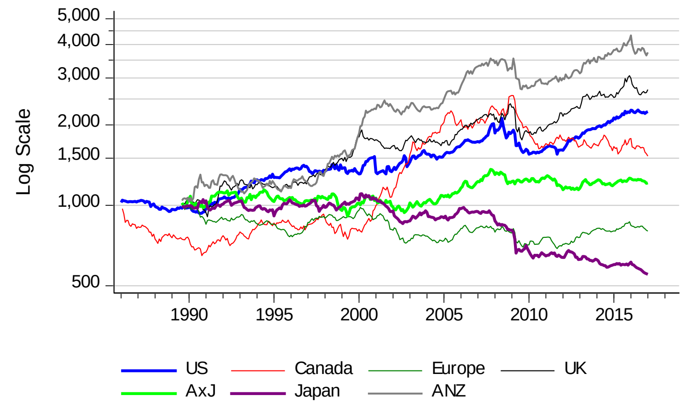
B) Projected Long-Term Growth vs Actual, Global



C) Exp Five-Year EPS Growth Factor Performance



D) Exp FY1/FY0 EPS Growth Factor Performance



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

PRICE MOMENTUM AND REVERSAL

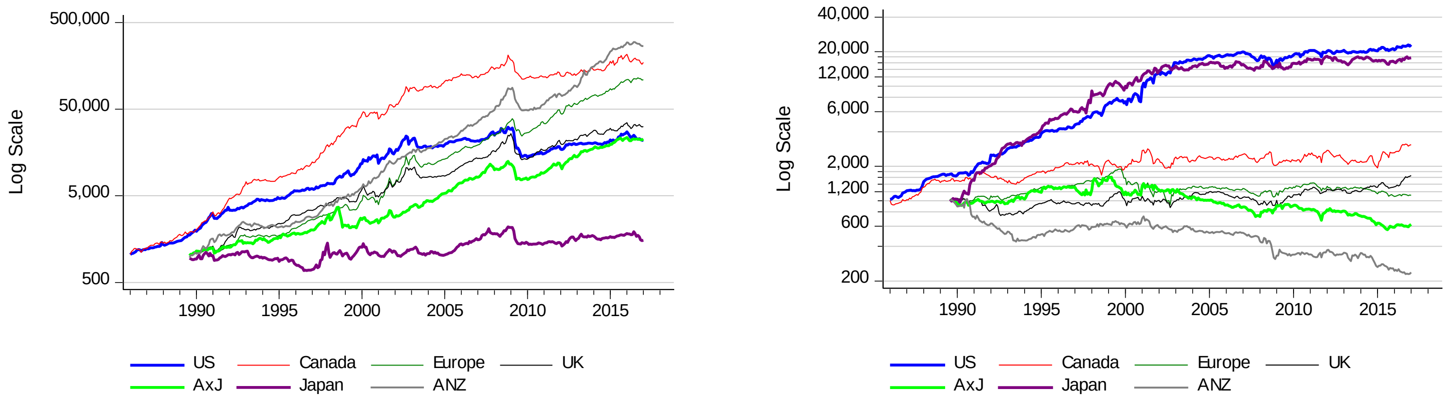
Researchers have found a strong price momentum effect in almost all asset classes in most countries. In fact, value and price momentum have long been the two cornerstones of quantitative investing.

Jegadeesh and Titman [1993] first documented that stocks that are the 'winners' over the last 12 months tend to outperform past 'losers' over the next two to 12 months in a study of the US market during the 1965 to 1989 period. The authors also found there is also a short-term reversal effect, where stocks which have a high price momentum in the last one month, tend to underperform in the next two to twelve months. In addition to the short-term reversal, mid-term momentum effect, some researchers also suggest that there is a long-term reversal pattern (see De Bondt and Thaler [1985]).

Empirically, mid-term price momentum effect can be observed in almost all regions, except Japan (see Figure 22 A). The March-May 2009 drawdown is also evident in all regions. On the other hand, the short-term mean reversal (based on last month's total return) effect is obvious in the US and Japan, but much more muted in other regions (see Figure 22 B).

Figure 22 Momentum and Reversal Around the World

A) Mid-Term Price Momentum (12M-1M) **B) Short-Term Mean-Reversal (1M)**

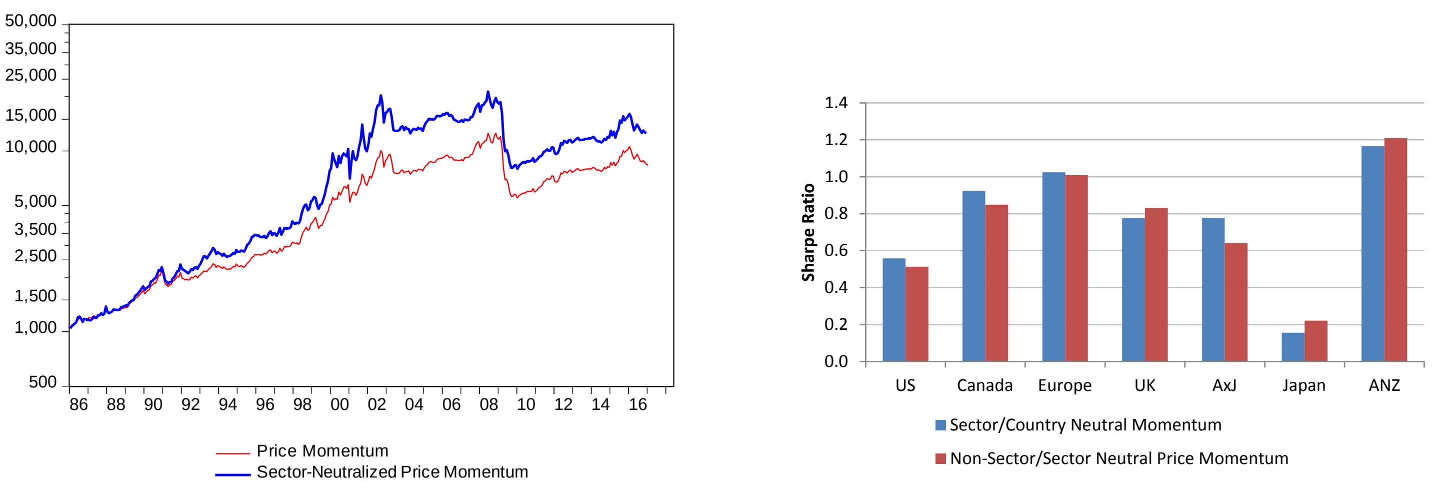


Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

This price momentum anomaly is commonly attributed to behavioral biases. It is interesting to note that since the academic publication of these findings, the performance of the price momentum factor has become much more volatile. Price momentum is also subject to extreme tail risk. As evidenced by the March-May 2009 time-period, the simple price momentum strategy (as measured by the long/short decile portfolio) lost -45% in three months. By neutralizing sector exposures, we can modestly reduce the downside risk and improve risk-adjusted performance in the US, Canada, and AxJ (see Figure 23).

Figure 23 Price Momentum and Sector Neutralized Momentum

A) Performance in the US **B) Performance Globally**



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

Neutralizing Price Momentum for Risk

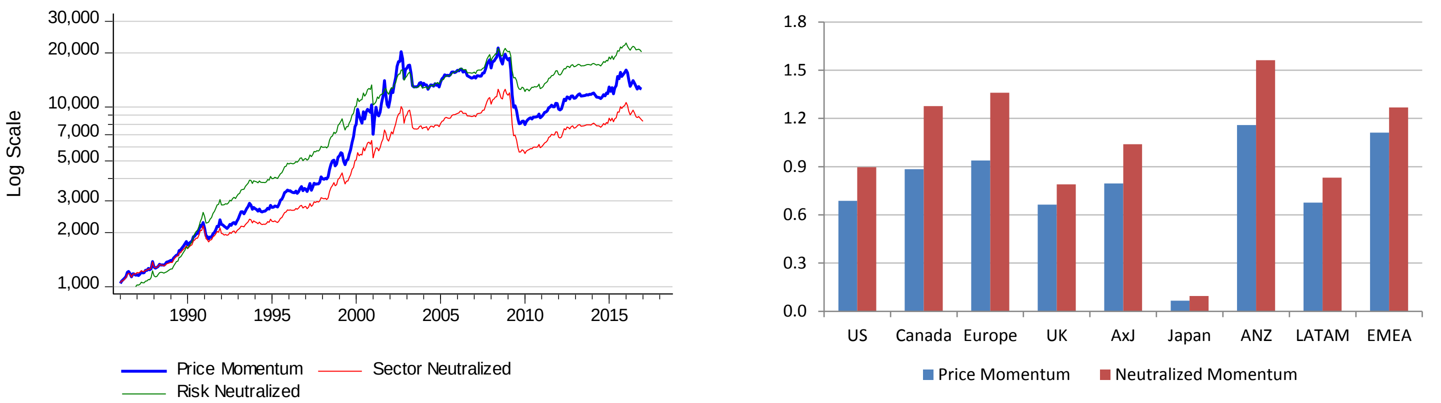
Ilmanen [2011] and Asness, et al [2012] point out that price momentum, as defined by past returns, should be adjusted for volatility. Otherwise, the most volatile stocks are much more likely to be in the best and worst momentum quantiles than the least risky assets. Therefore, in addition to sector neutral adjustment, we should also adjust for risk. One way to neutralize the impact of risk is via the following cross-sectional regression:

$$PriceMomentum_i = \gamma_0 + \gamma_1 BETA_i + \varepsilon_i$$

The regression residual (ε_i) becomes our new price momentum factor. As shown in Figure 24 (A), neutralizing risk boosts performance and cuts down risk at the same time in the US. Beyond the US market, we also observe the across the board boost in performance in the other eight regions (see Figure 24 B).

Figure 24 Neutralized Price Momentum Factor

A) Neutralized Momentum in the US **B) Neutralized Momentum (Sharpe Ratio) Globally**



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

The Lottery Factor

In academia, there has long been a great deal of evidence suggesting that investors like speculative opportunities – minimal amount of initial investment, but large potential payoffs. We define the lottery factor as the maximum daily return in the past one month (see Bali, Cakici, and Whitelaw [2011]):

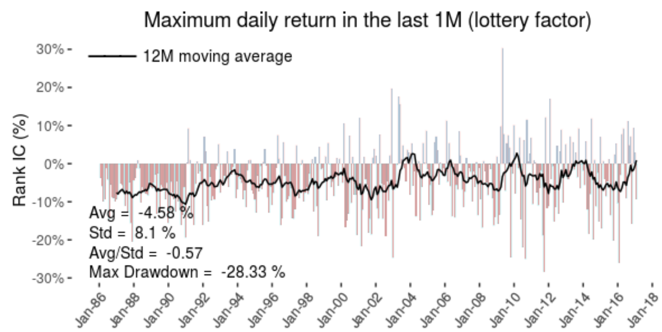
$$LotteryFactor_{i,t} = @Maximum\left(r_{i,t}, r_{i,t-1}, \cdots, r_{i,t=21}\right)$$

It is conceptually related to, but very different from the short-term reversal factor. Stocks with extreme returns in one day tend to attract a lot of attention from the media. Investors watching the news may expect these stocks to produce similar payoff in the near future. The temporary buying pressure is likely to push stock price beyond equilibrium, which leads to lower subsequent returns.

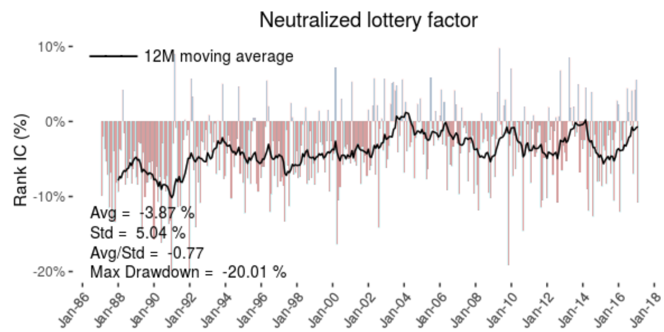
As Show in Figure 25 (A), the lottery factor generally has superior performance to traditional mean-reversal signals. Similar to price momentum, the lottery factor also has significant time-varying correlation to risk. Therefore, we also neutralize the lottery factor against beta. As shown in Figure 25 (B), the neutralized lottery factor's risk adjusted performance increases by over 30%. Furthermore, it also tends to have a much higher autocorrelation; therefore, lower turnover than reversal factors (see Figure 25 C and D).

Figure 25 The Lottery Factor

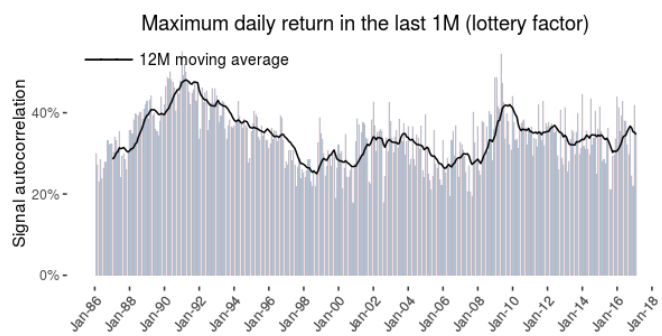
A) Lottery Factor, Rank IC



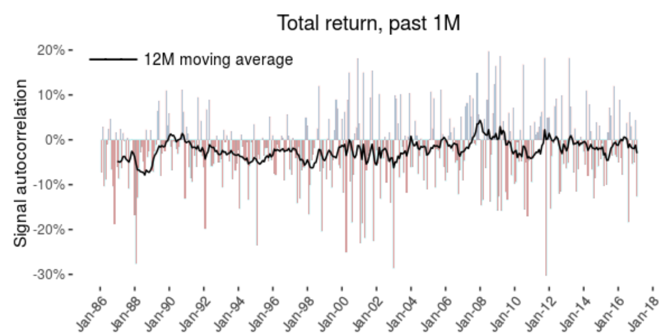
B) Neutralized Lottery Factor, Rank IC



C) Signal Autocorrelation – Lottery Factor



D) Signal Autocorrelation – 1M Reversal



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

ANALYST SENTIMENT

Factors based on consensus sell-side analyst estimates have long been extensively used by both quantitative and fundamental investors. There are a few vendors providing sell-side analyst estimates:

- Thomson Reuters IBES
- Factset
- S&P Capital IQ
- Bloomberg

There are also vendors that specialize in a given country:

- Morningstar (formerly CPMS) for Canada
- WIND for China
- Toyo Keizai for Japan

A really interesting alternative data vendor called Estimize uses estimates from buy-side, sell-side, and other participants. The crowdsourcing nature of Estimize's estimate makes it unique and less correlated. In our previous research, we find Estimize's estimates can be more accurate for large cap companies, especially when it is close to earnings reporting dates, than traditional sell-side consensus.

Metrics based on Consensus Estimates

Analyst estimate databases typically include a wide range of metrics:

- Metrics with Periodicity: EPS, CFPS (Cash Flow Per Share), DPS (Dividend Per Share), BPS (Book Value Per Share), Revenue/Sales, FFO (Funds From Operations), Capital Expenditure, Free Cash Flow, Gross Margin, EBITDA, Effective Tax Rate, NAV, ROA, ROE, etc.
- Metrics without Periodicity: Long-Term Growth Rate, Target Price, Recommendation, Volatility

For metrics with periodicity, we can compute over different horizons, e.g., FQ1 (the next fiscal quarter), NTM (next 12 month), FY1 (the next fiscal year), etc.

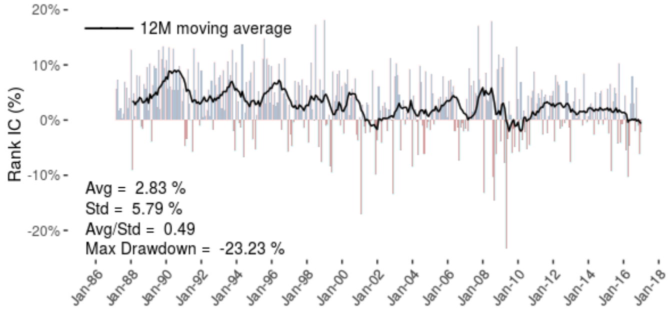
As shown in Figure 26 (A), the performance of the analyst revision factor in the US has declined substantially in recent years. We also detect similar performance downgrade in Canada and Japan (see Figure 26 B).

One interesting factor to note is the consensus recommendation itself. Sell-side analysts have different rating system, but vendors reconcile it into a one to five score, where one corresponds to strong buy, while five means strong sell. As shown in Figure 26 (C), the distribution is definitely skewed to buys. In particular, in the late 1990s technology bubble, recommendations were primarily buys and strong buys. The performance of sell-side analyst's rating signal is rather weak and it turns to the opposite direction since 2014.

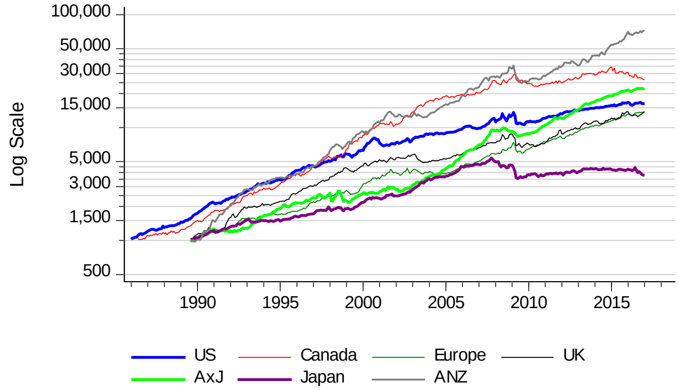
With the record low interest rates in most developed countries, investors have been chasing for yield. As a result, the price of high yield assets (e.g., high yield stocks and bonds) has increased significantly in recent years. With the fear of more rate hikes from the US Federal Reserve, investors look for alternatives. One area that generates strong interest is the dividend growth strategy. In our past research, we find dividend growth strategies stack up well with other factors. One way to gauge a company's ability to raise dividend is via the dividend revision factor. As a result, the dividend revision factor has shown exceptional performance in the US and Europe (see Figure 26 E and F).

Figure 26 Analyst Sentiment Factor Performance

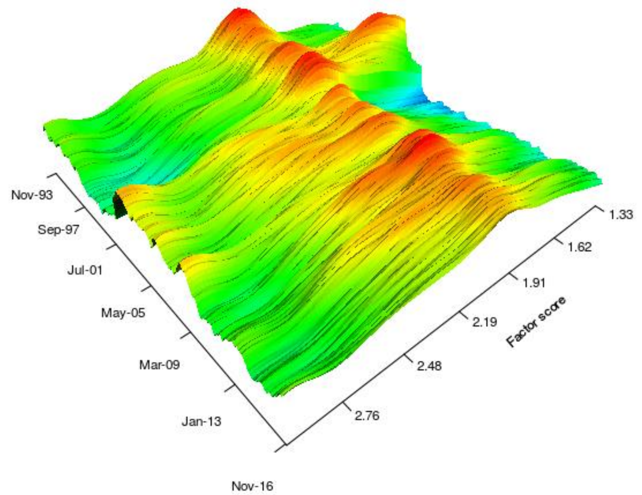
A) Earnings Revision, 3M, US



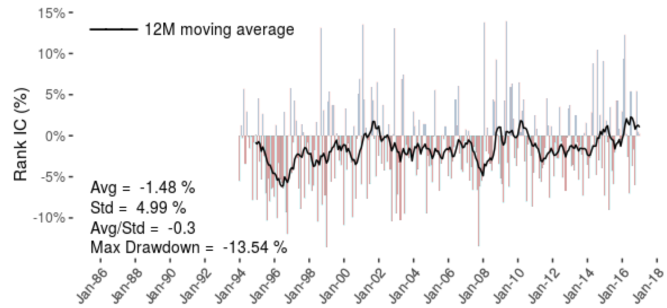
B) Earnings Revisions, 3M, Global



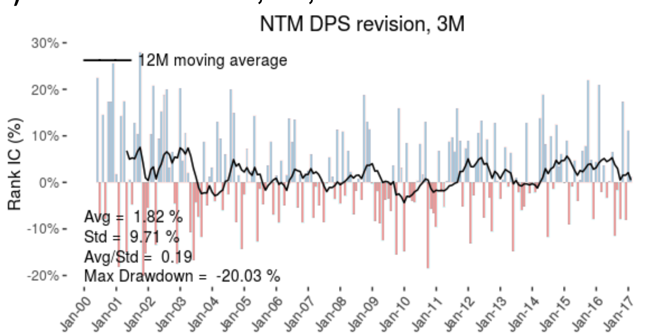
C) Consensus Recommendation Distribution



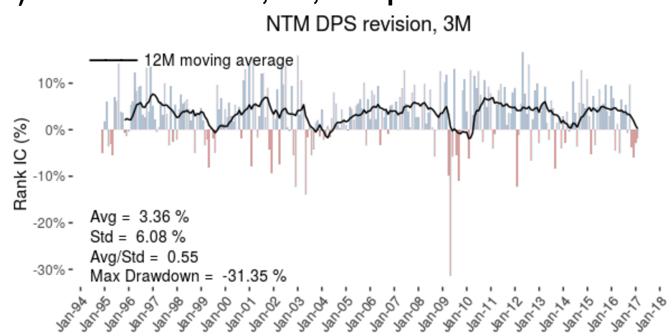
D) Consensus Recommendation, Rank IC, US



E) Dividend Revision, 3M, US



F) Dividend Revision, 3M, Europe

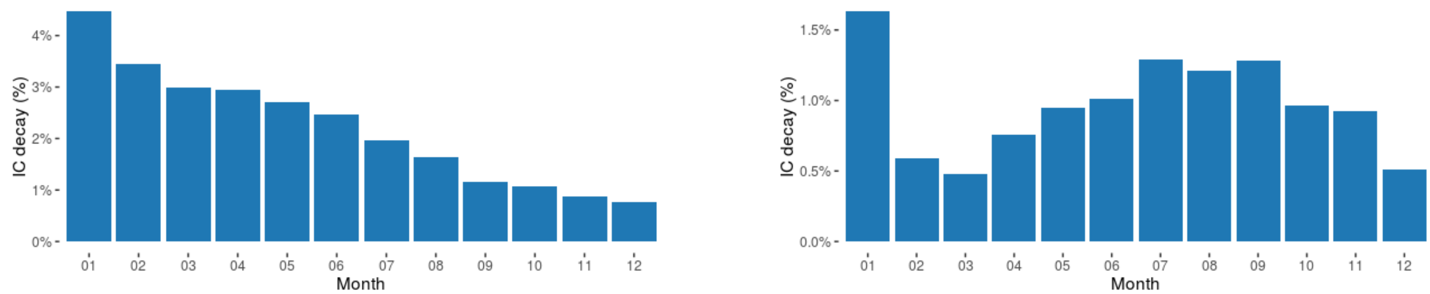


Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

To understand the speed of arbitrage, we backtest the three-month earnings revision factor in two periods: 1) 1985 to the end of 1999; and 2) 2000 to present. As shown in Figure 27 (A), in the 1980s/1990s, earnings revision factor had a smooth decay function. The IC was about 4.5% in the first month and then fell to 3.3% and 2.9% in the next two months. Even if we had to wait for a month to implement our strategy, we still had decent alpha. Since 2000, however, not only the average performance of the factor has come down to 1.6% IC for the first month, but also the speed of decay has accelerated (see Figure 27 B). There is essentially no predictive power after a month.

Figure 27 The Significant Decay in EPS Revision

A) Factor Decay, 1980s and 1990s **B) Factor Decay, 2000s-Present**



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

Detailed Estimates and more Sophisticated Consensus

It is traditionally done using consensus, but increasingly, detailed analyst-level data is being used. Many buy-side firms have built direct data feed from large sell-side banks to source data beyond traditional metrics on a more timely basis.

Investors can also use detailed estimates to construct more accurate consensus, for example, by overweighting:

- Analysts who have been more accurate
- Estimates that are more timely
- Estimates that are farther away from the consensus

This is an area that we are actively working on.

QUALITY

Apart from accounting ratios and share price data being used as fundamental style factors, analysts continue to create more complex factors based on the variety of accounting information available for companies. One of the best known examples of how in-depth accounting knowledge can impact investment performance is Richard Sloan's seminal paper on earnings quality (see Sloan [1996]), with the proposition of the accruals factor. It suggests that stock prices fail to reflect fully the information contained in the accrual and cash flow components of current earnings.

Many Faces of Quality

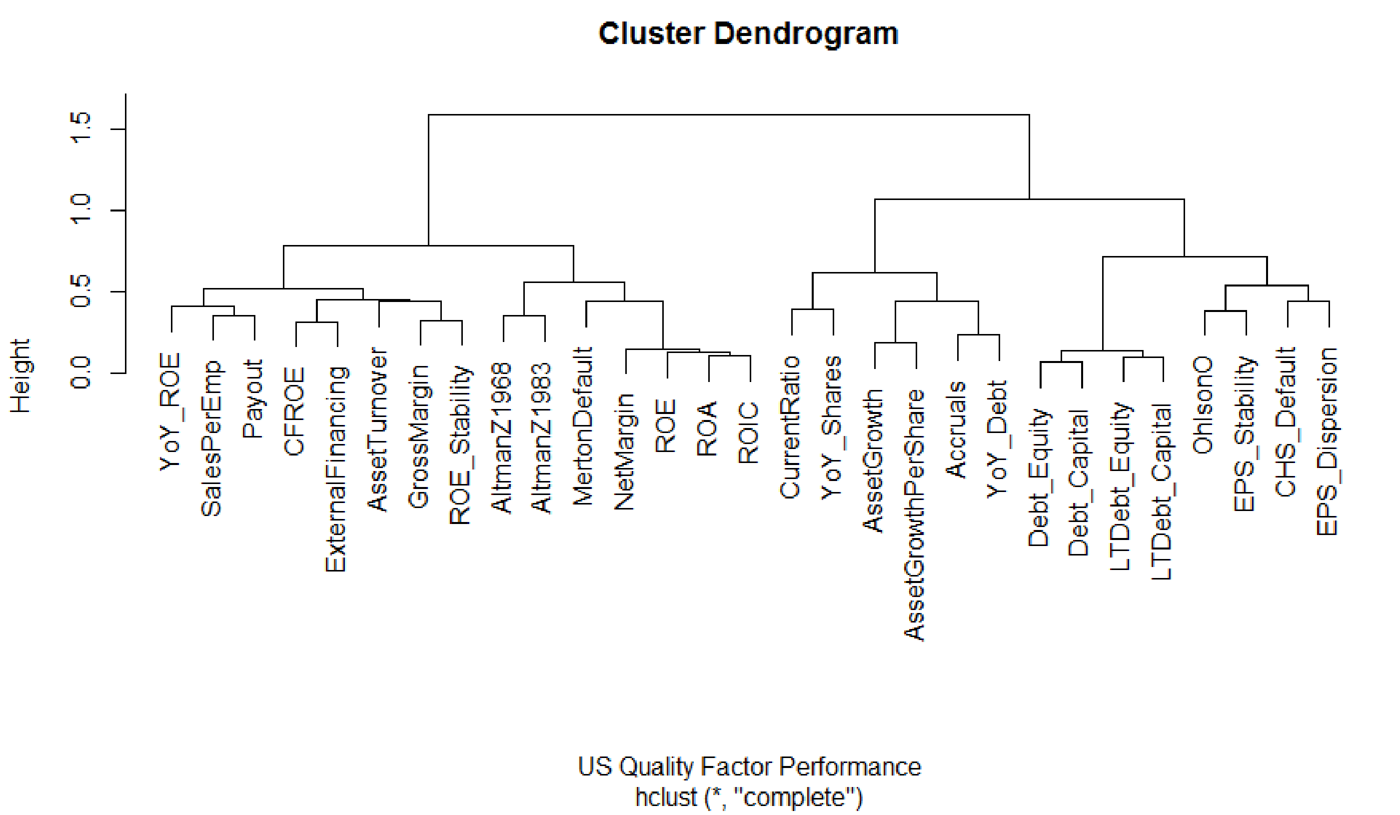
In addition to the accruals anomaly, there are many other potential factors based on a company's fundamental data, such as profitability, balance sheet and solvency risk, earnings quality, stability, sustainability of dividend payout, capital utilization, and management efficiency measures:

- Profitability: ROE, ROA, Asset Turnover, Profit Margin, Net Margin

- Financial Leverage and Liquidity: Current Ratio, Debt/Equity Ratio, Altman's z-score, Merton's distance to default
- Accounting Quality: Sloan's Accruals, Percent Accruals, Piotroski's F-Score, Mohanram's G-Score
- Stability: Historical EPS Stability, FY1 EPS Dispersion
- Capital Utilization: Asset Growth Anomaly, YoY Change in Shares Outstanding, Net External Financing/Net Operating Assets

We pick 29 representative quality factors from each of the above five categories and conduct a cluster analysis. As shown in Figure 28, the empirical classification is somewhat consistent with our manual categories. For example, ROE, Asset Turnover, and most profitability factors fall into the left branch. On the other hand, leverage ratios and default probability factors are mostly on the right side. Altman's z-score, however, appears to be closely related to the Merton's distance to default, but quite different from the Ohlson's O-score.

Figure 28 A Cluster Analysis of Quality Factors



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

An In-depth View of the Asset Growth Anomaly

One interesting quality factor is based on the asset growth anomaly (see Cooper, et al [2008] and Li and Sullivan [2015]). The asset growth anomaly argues that companies experiencing abnormally high growth in assets and capital spending tend to produce negative excess returns.

To empirically test the asset growth anomaly, we define our factor as:

$$AssetGrowth_{i,t} = \frac{TotalAssets_{i,t}}{TotalAssets_{t-q}} - 1$$

Where, q corresponds to a company's reporting frequency. For example, for firms reporting on a quarterly basis, $q = 4$. Similarly, if a company reports its financials twice a year, $q = 2$. Therefore, we essentially compute the year-over-year percentage change in a firm's total assets on the balance sheet.

Our first definition of asset growth is similar to Cooper, et al [2008], but defers in a few aspects:

- We use total assets from interim financial statements, rather than annual balance sheet. As shown in Luo, et al [2017], we find factors based on interim financial statements provide far more timely measures of firm performance and have much stronger predictive power of future stock returns. Using interim financial statement data, however, introduces a great deal of complexity in data management, as companies in different countries have different reporting frequencies. Furthermore, the same company may also change its reporting frequency and even fiscal year end from time to time.
- Instead of computing the signal once a year and limiting the universe to fiscal year end of December, we calculate our factor daily.
- Both Cooper, et al [2008] and Li and Sullivan [2015] use non-point-in-time database, which suffers from re-statement bias and arbitrary reporting lag assumptions. We use point-in-time global financial statement data. As shown in Luo, et al [2017], the potential look-ahead bias from restated financial statement data can be significant.
- Cooper, et al [2008] focus on the US market, while Li and Sullivan [2015] extend the study to global developed markets. We further expand the universe to the emerging markets and find very different results.
- Prior research typically computes total asset growth from two years ago, while we define our factor as year-over-year growth. The key reason is that we do not require a company to have at least two years of data, which increases the size of our universe slightly.

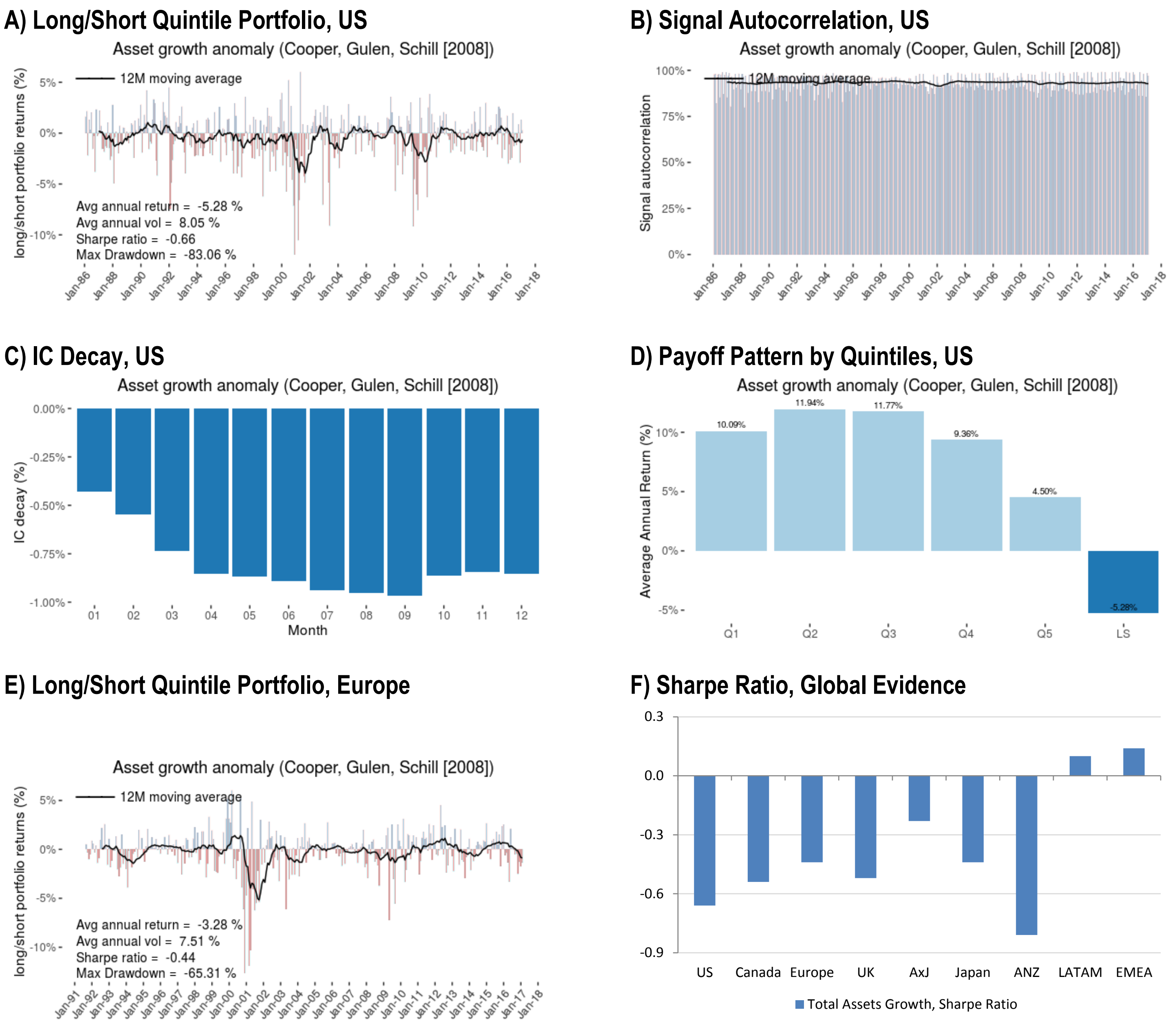
The mispricing argument of why asset growth anomaly exists suggests that investors tend to over-extrapolate past growth into the future. As you probably remember, in almost all university and CFA valuation courses, one of the key inputs for DCF (Discounted Cash Flows) type of models is the long-term growth rate. Almost all estimates of long-term growth start from the revenue growth in the past several years.

Another alternative explanation rests on systematic risk (see Fama and French [2014]), which argues that the mix of growth options and assets in place may introduce time varying risk for high growth companies.

As shown in Figure 29 (A), the highest asset growth quintile portfolio underperforms the lowest one by -5.3% per year. The asset growth factor has a Sharpe ratio of 0.66x. The turnover of the signal is extremely low (see Figure 29 B). The predictive power of the factor lasts for nine months (see Figure 29 C). Lastly, Figure 29 (D) suggests that performance of the signal is mostly from the underperformance of high asset growth firms, rather than the reward from low growth companies.

We also observe the asset growth anomaly in Europe (see Figure 29 E). Indeed, the asset anomaly is present in all developed markets in the US, Canada, Europe, UK, Japan, and ANZ. It is particularly strong in the Australian market. On the other hand, the phenomenon is much weaker in AxJ, which comprises both developed countries such as Hong Kong and Singapore and emerging countries such as China and Malaysia. Lastly, the asset growth factor turns into the opposite direction in LATAM and emerging EMEA, where growth is more likely to be encouraging news.

Figure 29 The Performance of the Asset Growth Factor



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

Now, let's dig a little deeper into the balance sheet. There is nothing magic about the total assets on the balance sheet. We want to see which line item causes the anomaly. We decompose the balance sheet into a few major sub-components, by both the assets side and liabilities & shareholders' equity side.

As shown in Figure 30 (A), we see substantial differences across different balance sheet items. The growth in Property, Plant and Equipment (PP&E) causes the largest damage to share price, with a Sharpe ratio of 1.13x, almost twice higher than the growth in total assets. The PP&E growth factor also exhibits consistent time series properties (see Figure 30 B). Globally, the PP&E growth factor also beats the plain vanilla asset growth factor in almost every single region (see Figure 30 C).

Another dimension that we have tested is to compute growth rate in per share assets:

$$PerShareAssetGrowth_{i,t} = \frac{TotalAssetsPerShare_{i,t}}{TotalAssetsPerShare_{t-q}} - 1$$

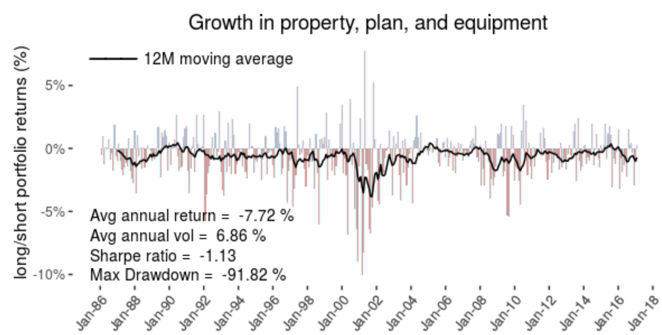
The performance of the per share asset growth factor, however, is considerably weaker (see Figure 30 D), across all balance sheet items. The per share growth more closely resembles organic growth – if a company is really able to grow its asset base without issuing new shares, excessive borrowing or external acquisitions, maybe it should not be penalized in the end.

Figure 30 Refining the Asset Growth Factor

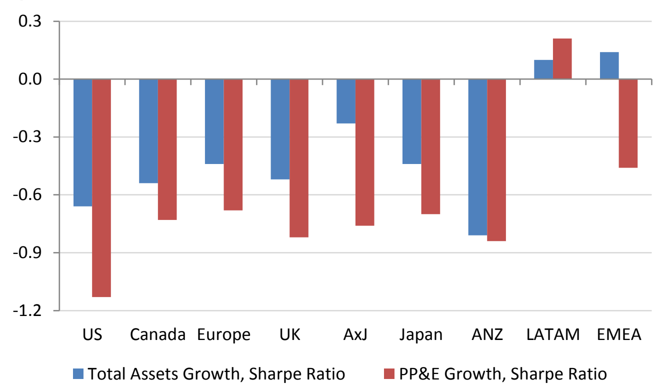
A) Decomposing the Balance Sheet Growth, US

	Sharpe Ratio	Risk-Adjusted IC
Assets		
Current Assets	(0.37)	0.02
Property, Plant, and Equipment	(1.13)	(0.27)
Other assets	(0.90)	(0.17)
Total Assets	(0.66)	(0.08)
Liabilities and Shareholders' Equity		
Current Liabilities	(0.70)	(0.10)
Total Liabilities	(0.97)	(0.21)
Shareholders' Equity	(0.26)	0.09
Total Liabilities and Shareholders' Equity		

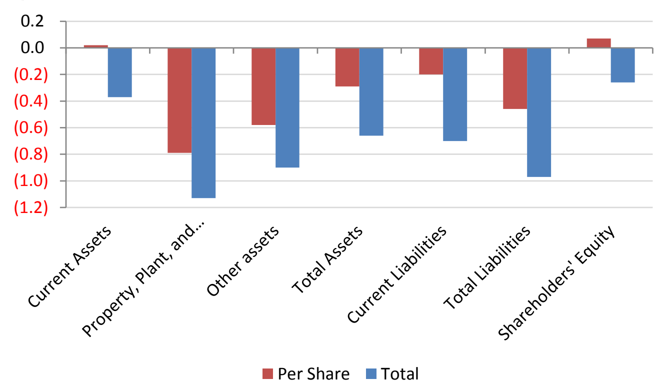
B) Long/Short Quintile Portfolio, US



C) PP&E Growth Factor, Global Evidence



D) Per Share Asset Growth Factors, US



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

ALTERNATIVE

We broadly label the last category of traditional factors, including any factors that do not fall into the previous five buckets.

Technical

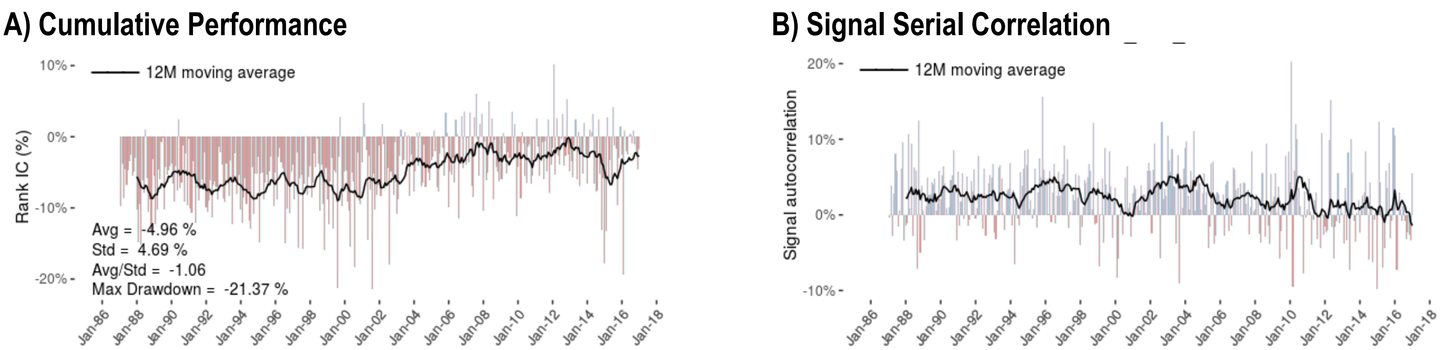
Technical factors are primarily computed on price and trading volume data. The weak form of EMH (Efficient Market Hypothesis) suggests that there should not be any alpha in technical indicators. However, they have been used in the investment industry for a long time, by both technical analysts and StatArb managers. Most of them are reversal in nature; therefore, suffer from extremely high turnover.

Figure 31 shows an example – the Close Location Value (CLV) indicator, which is used to determine where the price of the stock closes relative to the day's high and low:

$$CLV = \frac{(Close - Low) - (High - Close)}{(High - Low)}$$

The CLV ranges between +1 and -1 and behaves like a mean-reversal indicator. The performance of the signal is extremely strong pre-cost. However, the turnover of the factor is also steep (see Figure 31 B).

Figure 31 An Example – Close Location Value



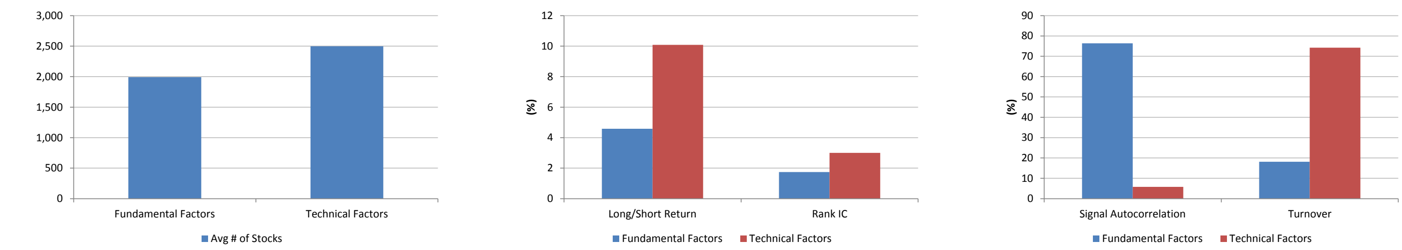
Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

We have a large number of technical indicators in our factor library. On average technical factors have much better coverage[6] (see Figure 32 A), stronger pre-cost performance (see Figure 32 B), but suffer from higher extraordinarily turnover (see Figure 32 C) than signals based on fundamental data. In Part IV of this research series, we will show the implication of transaction costs on model performance and discuss how to incorporate costs in our portfolio construction.

[6] Technical factors are primarily based on market data, e.g., price and volume data, which is available to almost all companies.

Figure 32 Technical versus Fundamental Factors

A) Coverage **B) Performance (pre-cost)** **C) Turnover**



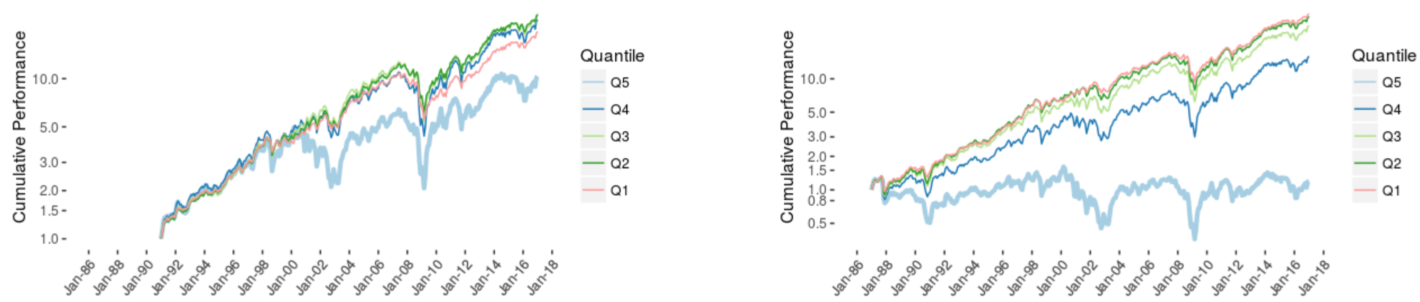
Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

Risk

There are many measures of a firm's risk, e.g., Beta, Volatility, Idiosyncratic Volatility, and Skewness. Classic finance suggests that diversifiable risk (e.g., idiosyncratic volatility) should not be compensated with higher returns, while non-diversifiable risk (e.g., beta) should be rewarded. However, empirical research often finds conflicting results. As shown Figure 33 (A) and (B), stocks with higher risk (defined by either beta or idiosyncratic volatility) are more likely to earn lower rather than higher returns. The payoff patterns seem to suggest that the low risk anomaly is more about the underperformance of high risk stocks rather than the outperformance of low risk names.

Figure 33 Risk-Based Factors

A) Beta **B) Idiosyncratic Volatility**

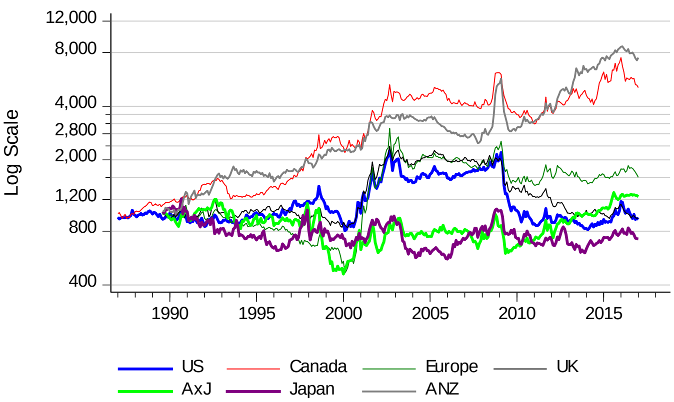


Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

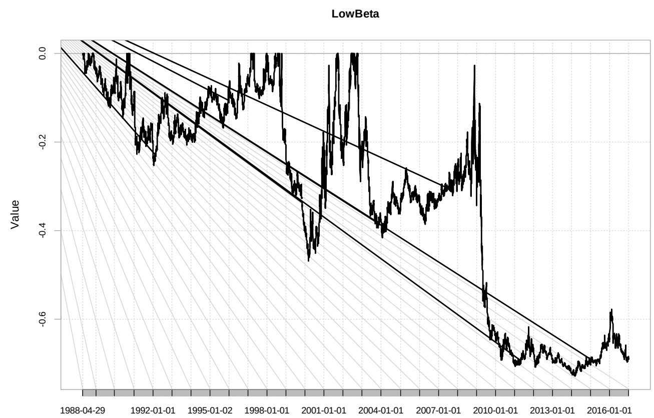
The so-called low risk anomaly has puzzled both academics and practitioners. Low risk oriented funds attracted great attention and asset inflow in the 2008 global financial crisis and subsequent years. However, our previous research (see Luo, et al [2017]) suggests that low risk strategies are particularly sensitive to risk-on/risk-off environment and vulnerable to crowded trades. As shown in Figure 34 (A) and (B), the performance of low beta strategy is quite volatile with substantial drawdown.

Figure 34 Low Risk Strategies are not Necessarily Low Risk

A) Low Beta Strategy Performance, Global



B) Low Beta Strategy Drawdown, US



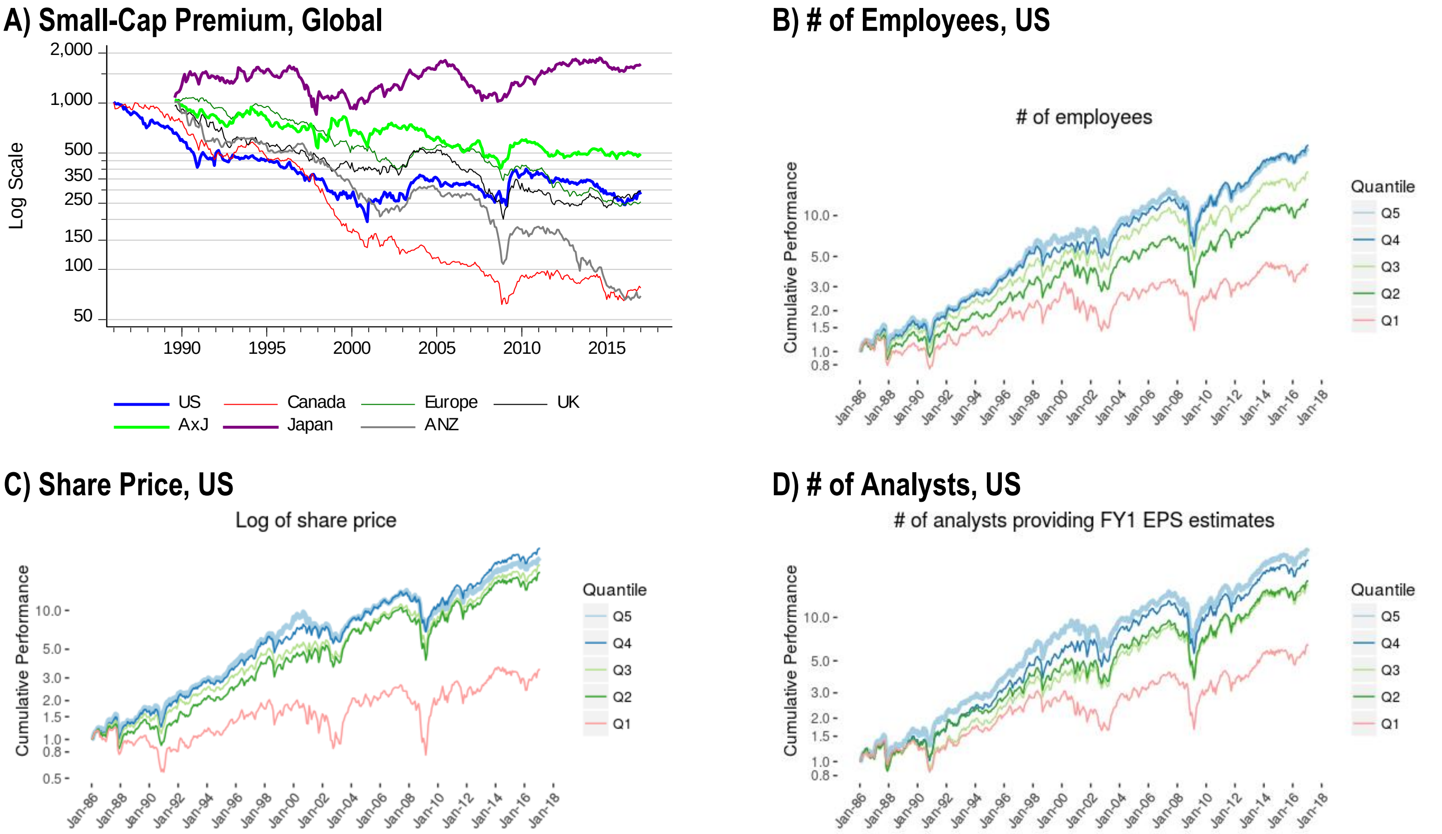
Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

Size

Fama and French [1993, 1996] and Carhart [1997] argue that the three key factors: value (book-to-market), size (market capitalization) and momentum can explain the majority of cross-sectional variation in stocks returns and dominate other market anomalies. The so-called small cap premium, however, is fairly inconsistent over time in the US (see Figure 35 A). Globally, we can only observe a visible small cap premium in Japan (see Figure 35 A).

In addition to market cap, there are also many other measures of size, e.g., revenue/sales, number of employees (see Figure 35 B), share price (see Figure 35 C), number of analysts (see Figure 35 D), length of history, etc. Similar to what find for market cap, almost all size factors suggest that small companies underperform the market in the long term.

Figure 35 Size Premium?



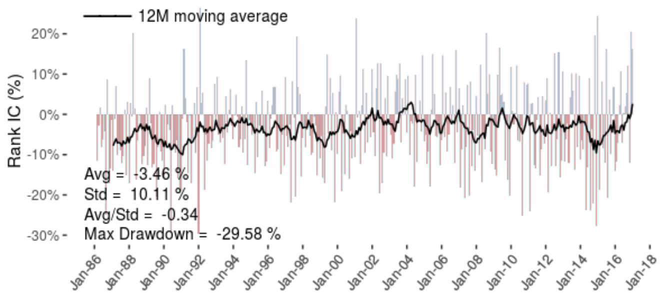
Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

Liquidity

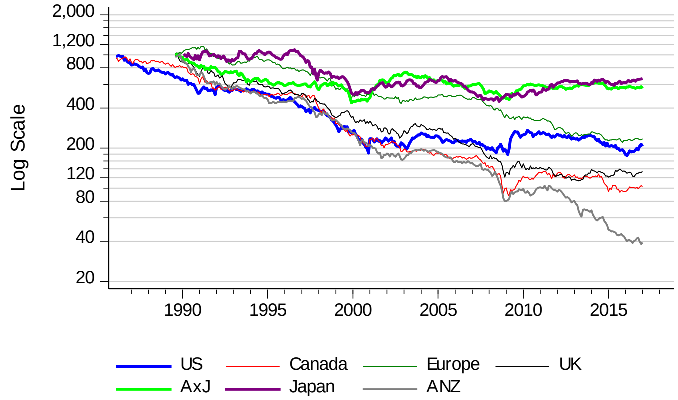
Similar to the argument of size premium, Amihud [2002] introduces a measure of illiquidity. He argues that illiquid stocks should earn higher returns to induce investors. Empirically, however, similar to what we find about size, the payoff to illiquidity is also time varying (see Figure 36 A) and differ greatly from country to country (see Figure 36 B).

Figure 36 Illiquidity Premium?

A) US



B) Global



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

Examples of Big Data Factors

The last but also the most exciting category is on Big Data factors, i.e., factors that are derived from alternative data sources. All the Big Data contents discussed in Luo, et al [2017] can and should be used to design new factors.

The opponents of using Big Data factors include:

- The coverage tends to be poor
- The history tends to be short
- The alpha itself tends to be modest at best
- There is likely data mining bias

However, Big Data factors also have a number of benefits:

- The alpha decay tends to be much weaker than traditional factors
- They tend to be uncorrelated to traditional factors
- They tend to be less crowded, as it takes tremendous time, efforts, and investment to integrate these factors; therefore, most investors do not use these factors in their investment process

This is an active area of research that we will devote a significant percentage of our time in the future. In this section, we will only show you two examples.

News Sentiment

A new and exciting area of research involves news sentiment. Rather than just relying on the output of sell-side analysts, the suite of Natural Language Processing (NLP) algorithms could be used to analyze the large volume of news stories to quantify the news sentiment on stocks. Our previous research finds that news sentiment based factors have decent predictive powerful and are uncorrelated to traditional analyst sentiment signals.

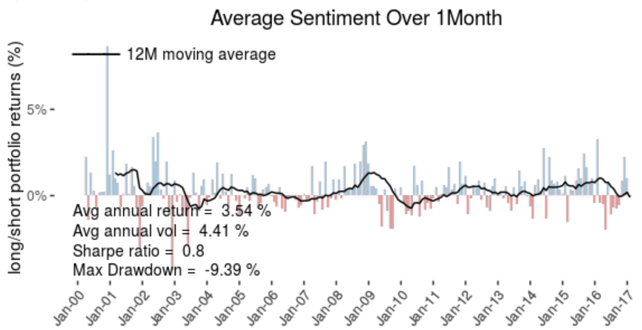
In this section, we show a simple example using Ravenpack's news sentiment engine. Ravenpack tracks the news of tens of thousands of companies, entities, currencies, commodities, and people worldwide. Then, the firm uses a suite of machine learning and NLP algorithms to quantify the sentiment in the news.

The factor of interest to us is a simple average of news sentiment from all stories about a company in the past month. Even on a monthly horizon, the portfolio generates a Sharpe ratio of 0.8x (see Figure 37 A) in the US. Furthermore, most of the access return comes from the long side (see Figure 37 B); therefore, it can be equally useful for long-only managers. Interestingly, the news sentiment indicator is uncorrelated to most traditional factors (see Figure 37 C). In particular, the correlation with analyst sentiment is less than 11%. Since the current NLP focuses on English language, it is also useful in other English language speaking countries, e.g., Canada (see Figure 37 D).

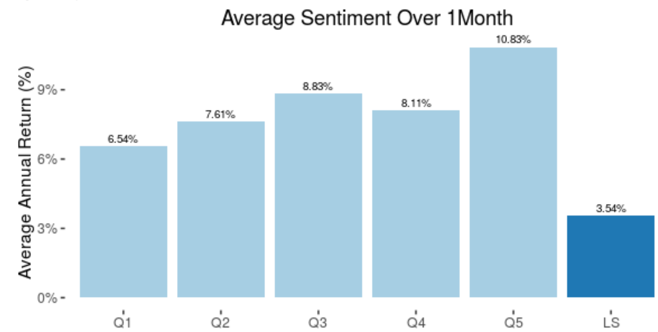
This barely scratches the surface of what we can do with textual information. We plan to cover this topic with more details in the future.

Figure 37 Ravenpack News Sentiment – Average Sentiment, 1M

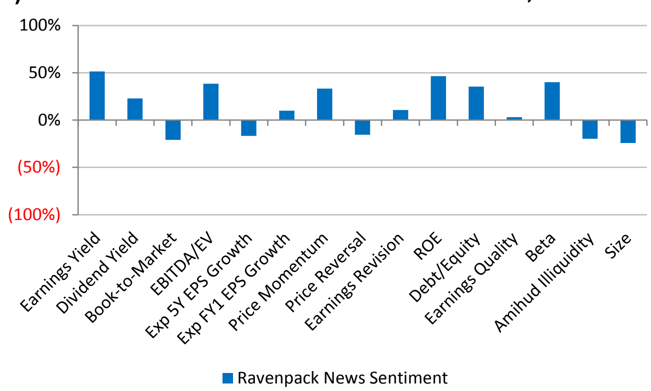
A) Long/Short Portfolio Return, US



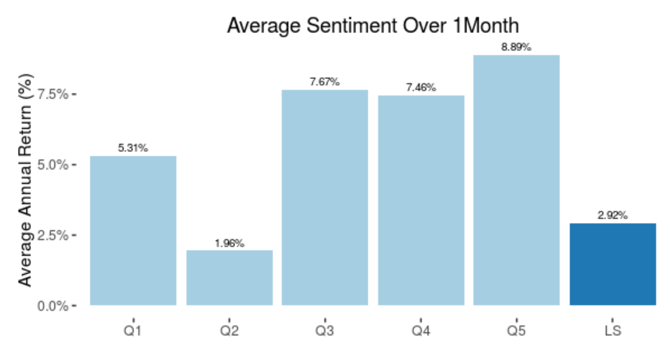
B) Payoff Pattern in the US



C) Correlation with other Common Factors, US



D) Payoff Pattern in Canada



Sources: Ravenpack, Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

Securities Finance

Markit Securities Finance (formerly DataExplorers) provides a unique source of securities lending data. It collects information from a wide range of participants in the stock loan trading market, including beneficial owners, buy side investors, and intermediaries globally. It is important to note that the securities lending market is OTC; therefore, data on this market is scarce, incomplete, and delayed. Without crowdsourcing from various participants, it would otherwise be impossible to have a timely data on this market.

Traditionally, investors collect short interest data from exchanges (e.g., NASDAQ, NYSE, TSX in Canada) or regulatory filings (e.g., FINRA in the US and FCA in the UK). However, there are a number of issues:

- It is often reported with delays. For example, Compustat collects and consolidate short interest data for US and Canadian stocks, twice a month at mid-month and month end. The month end short interest reflects the short positions as of mid-month.
- It is not available in many markets.
- Is only reflects the demand side, i.e., how many shares that short sellers want to short, but not the supply side, i.e., how many shares that asset owners are willing to lend.

Market Securities Finance database nicely fills in the gap. It has a number of interesting features to make it an invaluable source of information:

- It is updated daily, with a T+2 reporting lag[7].
- It covers 30,000 equity instruments[8] globally.
- It has a large number of data fields, covering demand (short interest), supply (inventories available), cost of borrow, etc.

One simple yet interesting factor is to combine the demand and supply sides to measure the true short interest of a stock – the utilization factor:

$$Utilization_{i,t} = \frac{ShortInterest_{i,t}}{AvailableInventory_{i,t}}$$

As shown in Figure 38 (A) and (B), the factor covers the vast majority of the market in both US and Europe. A monthly rebalanced long/short portfolio based on the short interest (utilization) factor has delivered an average return of 12.6% and a Sharpe ratio of 1.6x, pre-cost[9] (see Figure 38 C). Lastly, the correlation of the signal with other common factors seems to be reasonable, indicating the information derived from the securities lending market is unique.

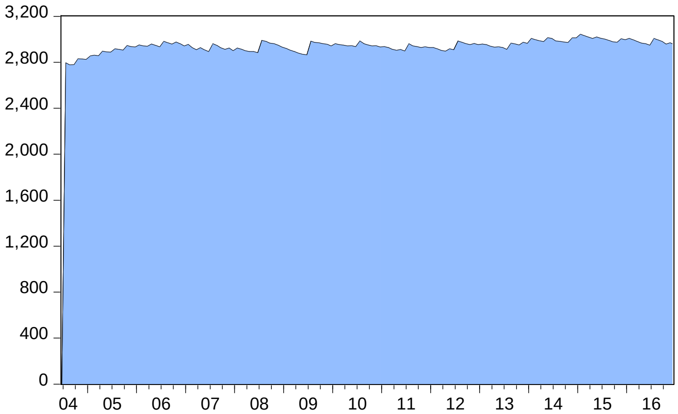
[7] Data is also reported intra-day, with roughly 70% data is available on a T+1 basis.

[8] Data is collected and reported on an issue level, including common shares, depository receipts, and ETFs. Markit also has short interest data on government and corporate bonds.

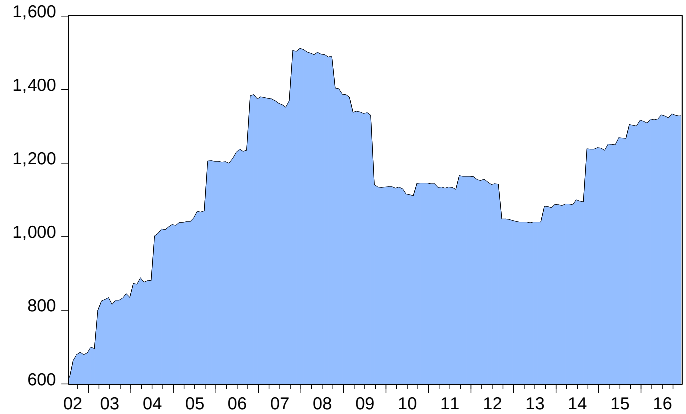
[9] The performance is pre-transaction costs and assumes that we can short all stocks without stock lending fees. As will be discussed in a forthcoming research, stocks that are heavily shorted tend to be difficult to borrow and expensive to short. Therefore, the after-cost performance is likely to be different. We will elaborate how to best use the signal in the future.

Figure 38 Markit Securities Finance – Short Interest (Utilization) Factor

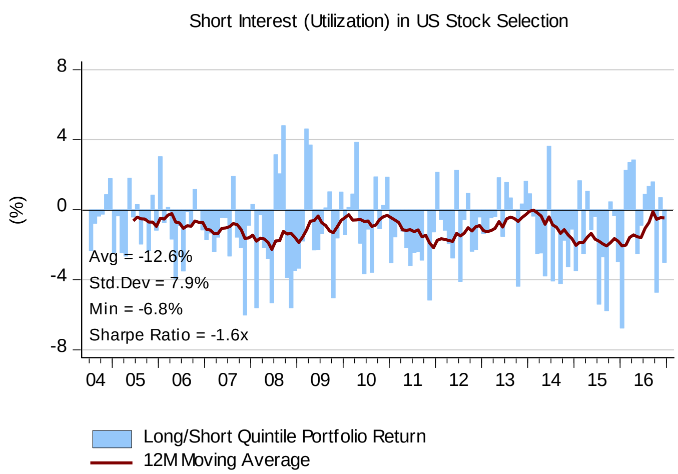
A) Cover in the US



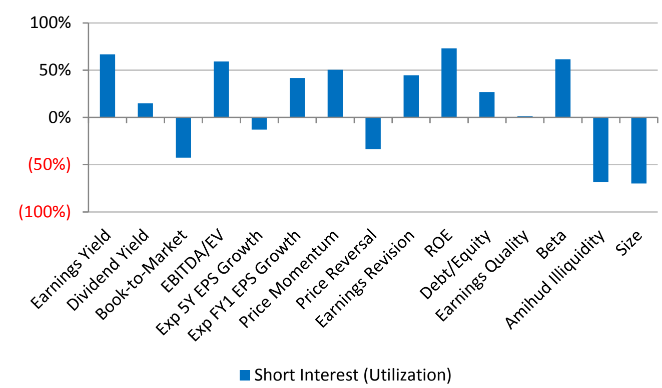
B) Coverage in Europe



C) Performance in the US



D) Performance Correlation with other Factors, US



Sources: Markit, Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

MULTIFACTOR MODELING

Few active managers use single factor models[10], and similarly, almost nobody includes hundreds of factors in the same model either. In practice, most stock selection models share some common multifactor structure, with linear combination being the dominant framework. In this section, we start from factor selection – an extremely important topic but has limited coverage in academic research. Then we move on to the factor weighting decision. We review some common factor weighting schemes and propose our own. We also compare multifactor active investing with a recent phenomenon of alternative beta portfolios. Advanced topics such as machine learning, style rotation, and our proprietary LEAP (L-Economic-Alpha Processing) model will be introduced in Part III of the *QES Handbook of Active Investing* series.

FACTOR SELECTION

Factor selection (in machine learning literature, it is also called *feature selection*) process can be modeled together with or separate from factor weighting. Either way, they are closely related. In the end, if you have a zero weight on a factor, it is the same as not to include it in the model. Some factor weighting algorithms have built-in feature selection, while others may require pre-selecting candidate factors.

Factor selection is the first and probably the most important step in building a multifactor model. However, there are a few significant challenges:

- **Dimensionality.** We have too many factors to choose from. In our standard factor library, we have about 400 relatively unique factors. Many of these factors have a few underlying parameters and therefore can be further expanded into other variations. As discussed in Luo, et al [2017], there are also increasingly more unstructured data sets being made available, which lead to even more factors.
- **Multicollinearity.** Obviously not all factors are unique. Many of them are highly correlated. For example, price momentum based on the past three, six, nine, and 12 months of returns are highly correlated. Analyst sentiment based on EPS, cash flow per year, revenue, and recommendations also point to similar directions. Some modeling techniques are more robust to linearly correlated factors (e.g., random forest), but the most widely used linear regression and factor optimization (e.g., Grinold & Kahn) are particularly vulnerable to multicollinearity.
- **Noisy Data.** In finance, the signal to noise ratio is extremely low, compared to what you normally see in science and engineering. Accidentally adding noisy or redundant factors to the model may lower performance.
- **Limited Prior Research from Academia.** Factor selection receives little coverage in the classic finance literature.

In the investment community, factor selection is generally done either by a discretionary `process or a systematic framework.

[10] The alternative beta (also known as smart beta and risk premia) investing is typically designed on single factors, as a middle ground between active and market capitalization weighted passive indices. We will discuss alternative beta investing in a later section.

Discretionary Factor Selection Process

In both academia and the investment industry, factor selection is mostly a discretionary process. The proponents suggest that it enforces economic intuition and avoids data mining bias. It normally follows the following steps:

- Researching academic and industry research publications
- Meeting and discussing with potential data vendors
- Retaining domain experts from academia or consulting firms
- Internal single factor backtesting
- Internal debating on the merit and economic rationale of the signal
- Assessing whether the signal is correlated with existing factors
- Deciding whether to include in the factor model and if yes, at what weight

The biggest problem with the above discretionary process is that it is not reproducible, which violates the first principle of systematic investing. It is nearly impossible to pinpoint whether it is due to skill or luck to include the signal at the first place.

Furthermore, when we debate on whether to include a factor in the model, the first and most important criterion is almost always on the factor's past performance. However, don't we know that "past performance is not indicative of future results"? If we select the factors with the best performance in the past, by definition, they would add incremental "alpha" in a backtesting, because the factor selection itself is in-sample.

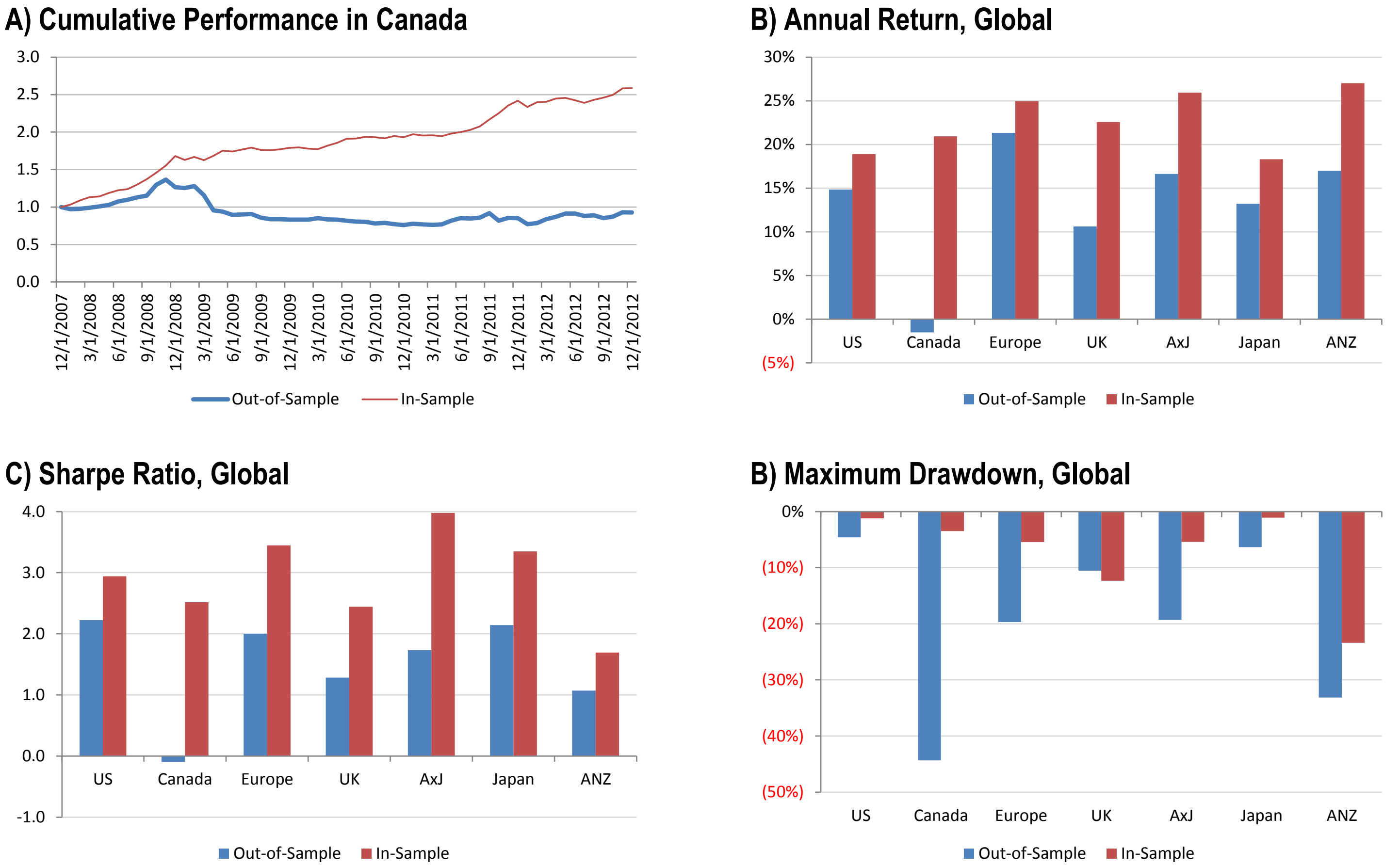
To demonstrate the impact of in-sample factor selection bias, we perform two backtests, for seven regions (US, Canada, Europe, UK, AxJ, Japan, and ANZ), respectively:

- **In-sample factor selection** – in this simulation, we select the best six factors – one from each of the six style categories (value, growth, momentum & reversal, sentiment, quality, and alternative), based on data from 2008-2012. We equally weight the six factors as our alpha model. Then, we backtest the performance of this model in 2008-2012. You can immediately spot the problem – at the end of 2007, we did not know what factors would generate the highest return in the next five years.
- **Out of sample factor selection** – in this backtesting, we select the best six factors – one from each of the six style buckets, based on data from 2002-2007. Then we equally weight the six factors and test the performance of the model for the subsequent five years in 2008-2012.

As shown in Figure 39 (A), the in-sample model had annual return of 20.9%, while the out-of-sample model was down by -1.5% in Canada. While it is an extreme case in Canada that the in-sample and out-of-sample models go on the opposite directions, we can clearly see that the in-sample model inflates performance significantly in every single region, with significantly higher returns (see Figure 39 B), greater Sharpe ratio (see Figure 39 C), and more modest drawdown (see Figure 39 D). Furthermore, the top six factors selected in-sample are almost never the same as the ones chosen out-of-sample.

While most of the other processes are systematic, it is odd to conduct factor selection in a discretionary way. We now shift our attention to systematic factor selection.

Figure 39 In-Sample versus Out-of-Sample Factor Selection



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

Systematic Factor Selection

In machine learning literature, factor selection is commonly known as feature selection. The factor selection process is rarely addressed in the finance literature. In our opinion, this is the area that is most suspicious to the model snooping bias. Using the machine learning terminology of John, Kohavi, and Pfleger [1994], dimension reduction can be achieved with two approaches:

- **Wrapper methods** evaluate multiple models using procedures that add and/or delete predictors to find the optimal combination that maximizes model performance. Essentially, wrapper methods are search algorithms that treat the predictors as the inputs and utilize model performance as the output. The traditional stepwise regression is a simple example, but there are far more complex and sophisticated models such as recursive feature elimination, genetic algorithms, and simulated annealing.
- **Filter methods** evaluate the relevance of the predictors outside of the predictive models and subsequently choose only the predictors that pass some criterion. For example, we could use a classification tree model to determine whether a predictor has any predictive power of stock returns. Then only those ones passed through the tree model will be used in a subsequent

linear regression. Saeys, Inza, and Larranaga [2007] provide a comprehensive review of filter methods.

Some machine learning algorithms have built-in factor selection (feature selection), such as tree and rule-based models, MARS, LASSO. However, the performance of most modeling techniques, in particular, linear regression tend to downgrade substantially in the cases of large number of highly correlated and non-informative predictors.

The classic forward selection algorithm starts from adding one factor at a time and normally uses p-value as the decision criterion, until no variable's p-value is below the decision threshold. Similarly, the backward selection algorithm begins with all factors and then removes one variable at a time. This procedure suffers from a number of issues:

- The search procedure is greedy, meaning that it does not reevaluate past solutions.
- The use of repeated hypothesis tests invalidates many of their statistical properties, since the same data are being evaluated multiple times (see Harrell [2001] for extensive discussion on stepwise regression). This is often referred by many investors as model overfitting.
- Maximizing statistical significance may not be the same as maximizing forecasting accuracy or strategy profitability.

Stepwise regression improves the forward regressions by reevaluating each term for removal after each factor is added, which makes the algorithm less greedy, but exacerbates the problem of repeated hypothesis testing. AIC (Akaike Information Criterion) or BIC (Bayesian Information Criterion) can be used in place of p-value, by adding a penalty to complex models.

EMPIRICAL ANALYSIS OF FACTOR SELECTION ALGORITHMS

In stock-selection modeling, the data is normally structured in three dimensions: company (i.e., stock), factor, and time, which complicates the modeling process considerably. In econometric jargon, it is called panel data, i.e., the same set of sample (i.e., company stocks) is observed repeatedly over time.

Panel Regression

The most conventional model is panel regression. Essentially, we stack all observations over multiple periods together, and then estimate the individual effects (at the company level) or time effects jointly. A theoretical model can be specified as:

$$r_{i,t} = \beta_0 + \sum_{k=1}^{K_t} \beta_{i,k,t} f_{i,k,t-1} + \delta_i + \gamma_t + \epsilon_{i,t}$$

Where,

$r_{i,t}$ is the return for stock i in period t;

β_0 is the overall intercept term,

$\beta_{i,k,t}$ is the estimated coefficient (i.e., orthogonal return) for stock i, factor k in period t;

$f_{i,k,t-1}$ is the normalized score of stock i, for factor k, in period $t-1$;

δ_i is a company specific effect for stock i,

γ_t is a time specific effect at time t,

$\epsilon_{i,t}$ is the regression residual (i.e., the random noise that can't be explained by our factors), for stock i in period t;

$i = 1, \ldots, N_t$, where N_t is the number of stocks in our universe in period t;

$k = 1, \ldots, K_t$, where K_t is the number of factors in period t; and

The above model can't be directly estimated, because there are more parameters than the number of data points. Therefore, a structure has to be imposed. There is a long and growing list of econometrics techniques in panel data modeling. For example, the $\beta_{i,k,t}$ coefficient is often assumed to be invariant to individual and time, i.e., we only estimate β_k, while the cross-section and/or period specific effects terms δ and γ may be handled using fixed or random effects models. The fixed effects models (see Baltagi [2005]) resemble our traditional OLS regression, but it is important to remove the means from the dependent and exogenous regressors. The random effects models are unfamiliar to most equity analysts. They assumes that the δ and γ coefficients are realizations of independent random variables with mean zero and finite variance. Random effects can be estimated via feasible GLS (Generalized Least Squared) techniques (details see Baltagi [2005]).

Pooled Time Series Cross Sectional Regression (PTSCS)

Because of the complexity involved in panel regression, equity analysts mostly stay away from the technique. In addition, in the equity world, individual effects (at the company level) are rarely of interest. Rather, the time effects are essential, since we need to account for the time-varying nature of factor returns.

A more commonly used technique by practitioners is the Pooled Time Series Cross Sectional Regression (PTSCS). The PTSCS essentially stacks all observation over time together, as if they all occur at the same time by ignoring the time dimension altogether:

$$r_{i,t} = \beta_0 + \sum_{k=1}^{K_t} \beta_k f_{i,k,t-1} + \epsilon_{i,t}$$

As you can see, the regression coefficients (β_k) are no longer vary by company or time. The above equation is often estimated via a simple OLS regression and our familiar statistical tests (e.g., T-statistics and p value) can be used to assess whether a factor is a statistically significant pricing factor.

The PTSCS approach assumes factor payoffs are static over time, which is highly unrealistic in our applications. We will now discuss a more robust technique – the Fama-MacBeth regression.

Fama-MacBeth Regression

In the Signal Research section, we form our long/short quantile portfolios at each month end and track the performance in the following month. The procedure is repeated every month, so that we have a time series return of our hedged portfolio. Alternatively, if we measure performance using the Spearman rank IC, we compute the correlation (rank correlation) between factor scores at a month end and the returns of all stocks at the subsequent month. Then, the same calculation is repeated over time to get a time series of IC's.

When we deal with multifactor models, we can naturally extend the above procedures to a multivariate setup. One of the most traditional, but also extremely powerful factor selection algorithms is the Fama-MacBeth Regression (see Fama and MacBeth [1973]. The Fama-MacBeth regression is essentially a cross-sectional multivariate regression, repeated over time:

$$r_{i,t} = \beta_0 + \sum_{k=1}^{K_t} \beta_{k,t} f_{i,k,t-1} + \epsilon_{i,t}$$

Where,

Once we repeat the above regression over a number of periods (T), we can collect the time series of regression coefficients ($\beta_{k,t}$). The statistical significance test is conducted on the time series of $\beta_{k,t}$. For example, a simple T-statistic can be computed as:

$$T - Statistic = \frac{\bar{\beta} - \mu}{\sigma/\sqrt{T}}$$

Where,

μ is typically set as 0 (i.e., whether factor return is statistically different from 0),

σ is the sample (time series) standard deviation of β, and

T is the number of time periods (normally we use a rolling window).

Since the time series of factor returns is likely to show serial correlation, we highly recommend a more robust version of statistical testing, e.g., the Newey-West (see Newey and West [1987]) procedure.

An Empirical Study

To gauge the ability and sensitivity of the Fama-MacBeth regression, we conduct an interesting simulation, with the following eight factors:

- Value: Trailing earnings yield – we prefer companies with high earnings yield
- Value: Book-to-market – we buy companies with high book-to-market, i.e., cheap stocks on valuation
- Growth: Consensus FY1/FY0 EPS Growth – we prefer companies with high earnings growth
- Price Momentum: 12M total return excluding the most recent month – we prefer companies with positive price momentum
- Analyst Sentiment: 3M EPS revision – we buy companies with positive earnings revisions
- Quality – Profitability: Return on equity – we like firms with high ROEs
- Quality – Leverage: Debt/Equity ratio – we prefer companies with low financial leverage
- Quality – Earnings Quality: Sloan’s accruals – we buy companies with low accruals

by running three models:

- **Baseline Model.** The baseline model starts from the above eight common stock-selection factors. We regress the forward one-month stock return on the eight factors, at each month end. Then we repeat the regression every month.
- **Multicollinearity Model.** To measure the sensitivity to highly correlated factors, we add six more price momentum factors to the baseline model – past 6-, 7-, until 11-month returns. The six price momentum factors are closely related with our 12M-1M price momentum signal.
- **Uninformed Model.** To measure the model's ability to distinguish uninformed random noise, we add six completely random simulated data to the model. The six new factors are sampled from a standard normal distribution, i.e., completely uninformed.

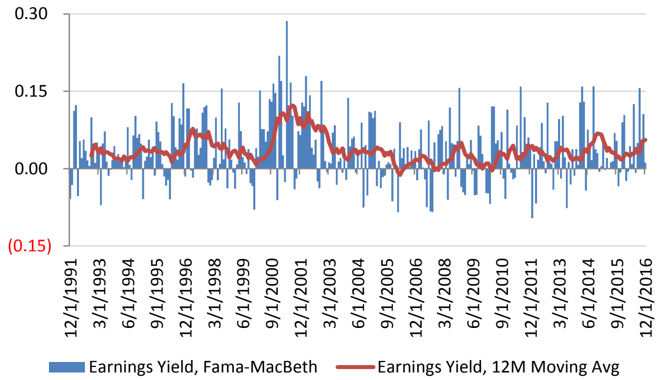
Multivariate Regression versus Single Factor Backtesting

Let's start from the baseline model. Figure 40 (A) plots the time-series coefficient of the (cross-sectional) regression on earnings yield. Clearly, the coefficient is consistently positive. Even after adjusting for other common factors, earnings yield generates consistent positive payoff. Figure 40 (B) shows the Spearman rank IC of earnings yield, based on univariate backtesting. Although Figure 40 (A) and (B) demonstrate broadly similar patterns, there are noticeable differences. Actually, the correlation between them is less than 20%. Lastly, Figure 40 (C) presents the performance based on a long/short quintile portfolio. It appears that univariate backtesting based on long/short hedged portfolio is more consistent with single factor backtesting with IC (the correlation is almost 90%), but they are both different from multifactor regressions.

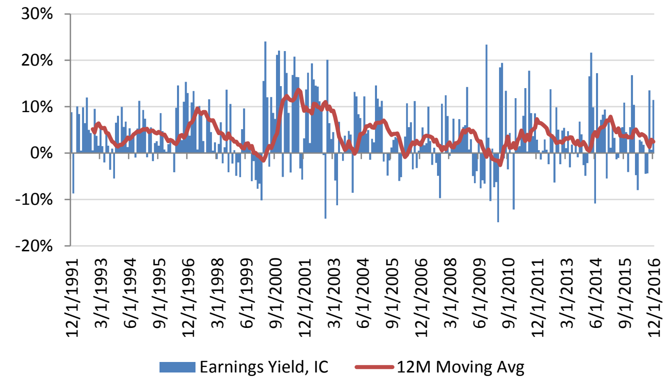
Now, let's shift our attention to debt/equity ratio. The debt/equity ratio behaves more like a risk factor than an alpha signal, with significant time variation. Furthermore, as shown in Figure 40 (D), (E), and (F), the two univariate backtestings are highly correlated (correlation is almost 85%), but neither one is correlated to the multifactor regression coefficient. Lastly, both Fama-MacBeth regression and IC studies find highly leverage firms deliver higher subsequent returns (statistically insignificant), while the long/short portfolio tells the opposite story (again, statistically insignificant).

Figure 40 Fama-MacBeth Regression versus Univariate Backtesting

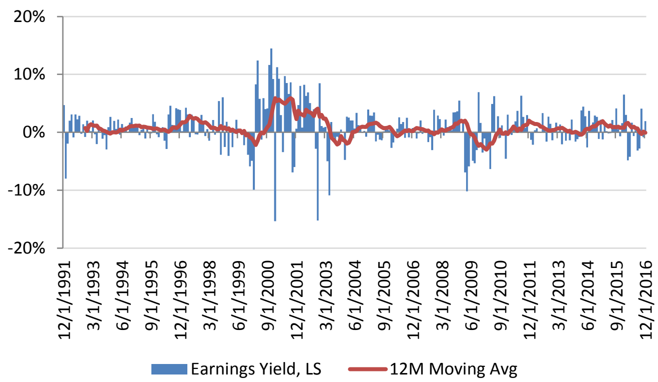
A) Earnings Yield, Fama-MacBeth Regression Coef



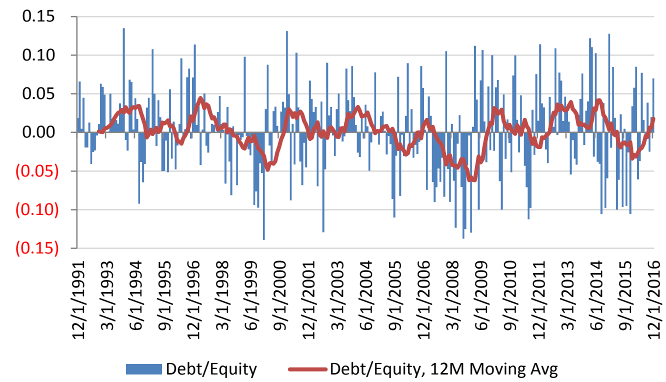
B) Earnings Yield, Rank IC



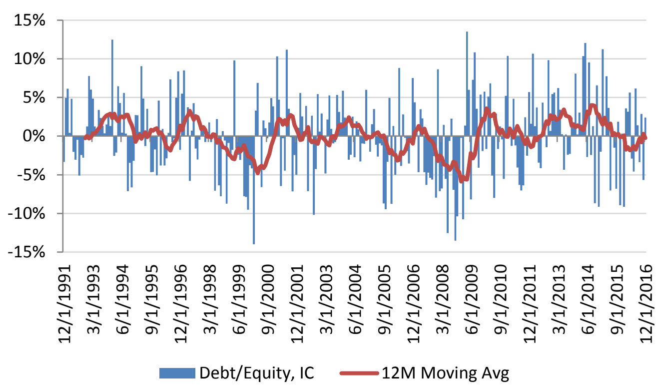
C) Earnings Yield, Long/Short Quintile Return



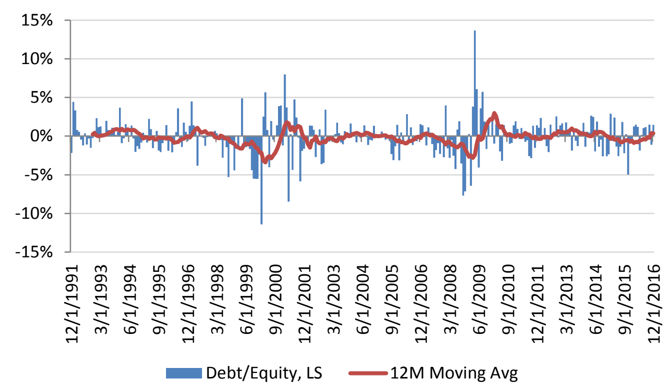
D) Debt/Equity Ratio, Fama-MacBeth Regression Coef



E) Debt/Equity Ratio, Rank IC



F) Debt/Equity Ratio, Long/Short Quintile Return



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

Using Fama-MacBeth as a Factor Selection Tool

Once a Fama-MacBeth regression is performed and T-statistics are computed, we can move on to traditional statistical decision making of whether to include or remove a regressor. For example, we can compute the p-value of each variable and remove any factor with a p-value greater than 10%.

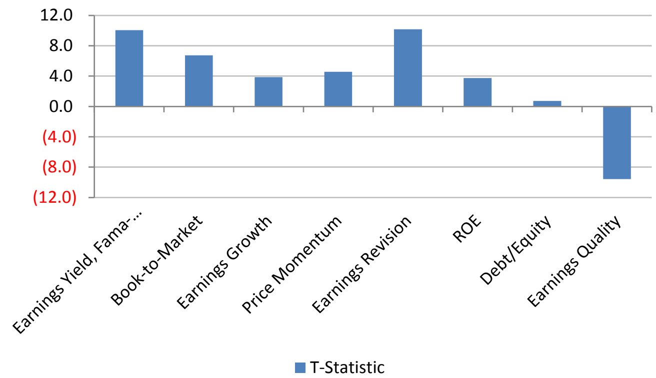
Figure 41 (A) shows the t-statistics for the eight factors in the US, using the Fama-MacBeth regression. Other than debt/equity ratio, all the other seven factors are statistically significant in predicting future stock returns. However, using the entire history to compute the t-statistic hides the fact that factor returns are time varying. Therefore, in practice, we often use a rolling window to compute the t-statistic.

Now, let's compute the t-statistic of each factor using a five-year rolling window. As shown in Figure 41 (B), the t-statistic of earnings yield is consistently above 2.0, while earnings growth is clearly cyclical. The predictive power of price momentum basically disappeared since 2004 (see Figure 41 C). The debt/equity ratio signal is not a significant pricing factor, other than the brief period around the 2008 financial crisis. The earning quality (Sloan's accruals) factor was highly persistent, but it seems to be losing its appeal in recent years (see Figure 41 D).

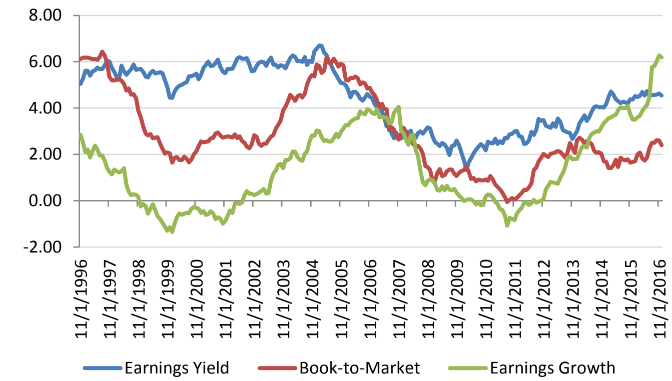
Figure 41 (E) shows that the earnings growth factor has been selected almost 40% of the time. On average, there have been around four to five factors in the model throughout the history (see Figure 41 F).

Figure 41 Using Fama-MacBeth Regression as a Factor Selection Tool

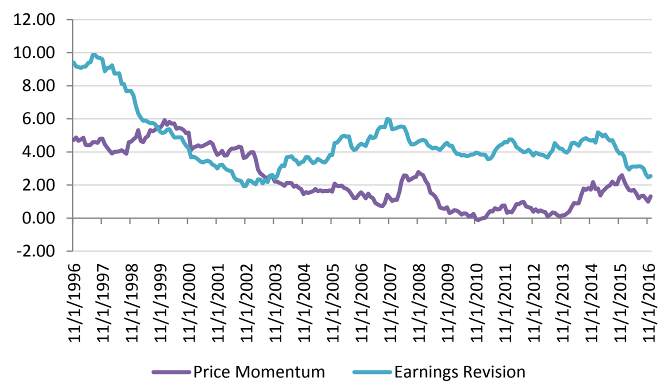
A) T-Statistics over the Entire History, US



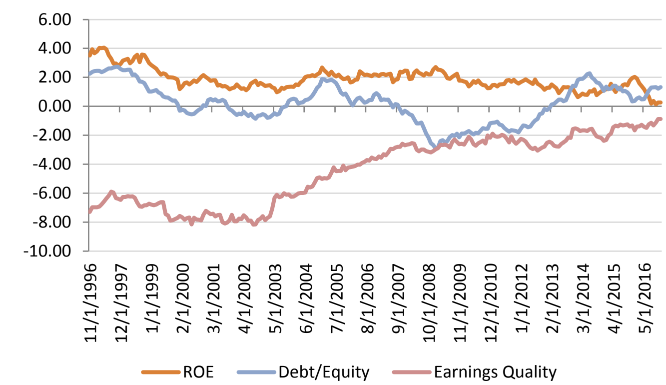
B) Rolling T-Statistics, Value and Growth



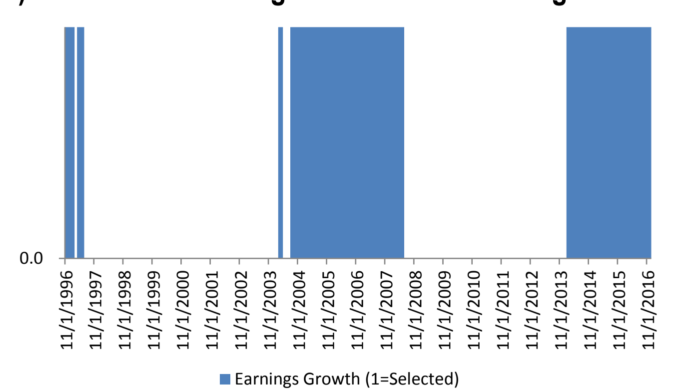
C) Rolling T-Statistics, Momentum and Sentiment



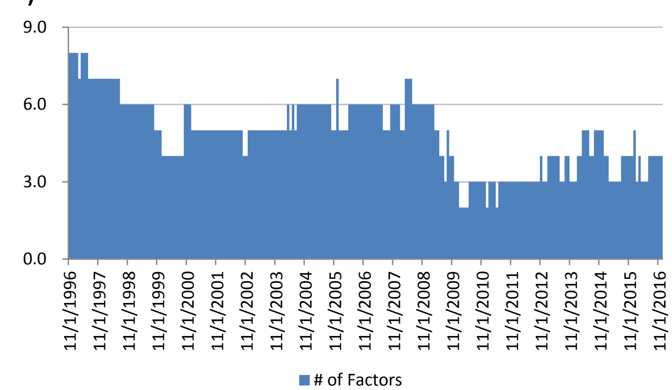
D) Rolling T-Statistics, Quality



E) When is the Earnings Growth Factor Being Selected?



F) Number of Factor Selected



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

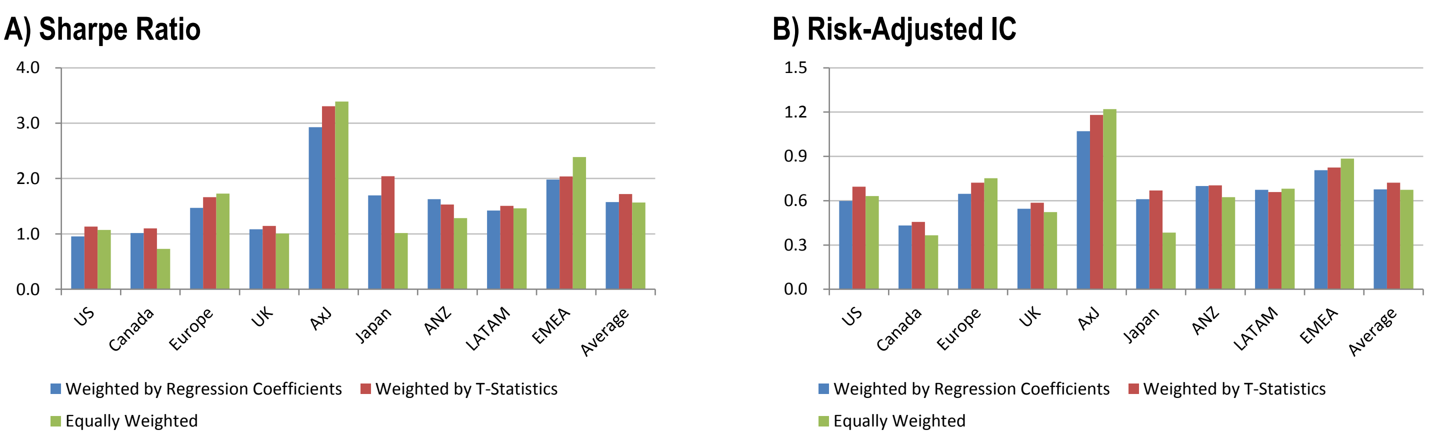
Using Fama-MacBeth as a Factor Weighting Tool

In addition to a factor selection tool, we can also use the results from the Fama-MacBeth regression to weight the chosen factors. For example, if five factors are all statistically significant today, how much weight are we going to assign to them? There are certainly a number of alternatives:

- **Equally weighting scheme.** This is definitely the easiest and arguably also fairly robust. We can assign equal weights to all factors: $\omega_k = \frac{1}{K}$.
- **Weighting by the regression coefficient.** The cross-sectional regression coefficient represents the orthogonal return from each factor ($\omega_k = \frac{|\beta_k|}{\sum_{k=1}^{K}|\beta_k|}$). Therefore, this resembles a weighting by (expected) return approach.
- **Weighting by the t-statistic.** Factor weighting, similar to portfolio construction, should take both expected return and risk into account ($\omega_k = \frac{|TStat_k|}{\sum_{k=1}^{K}|TStat_k|}$). If all correlations are zero (or the same), weighting by t-statistics is also similar to a full mean-variance optimization in the factor space. Therefore, weighting by t-statistics seem to be a better solution. However, since both risk and return are unknown variables and can only be estimated with data, potential estimation errors may overshadow the benefit.

As shown in Figure 42, The Fama-MacBeth procedure is highly robust. Once the factors are selected, the three weighting schemes yield similar performance. Weighting by t-statistics does appear to outperform the other two approaches slightly, based on both Sharpe ratio and risk-adjusted IC. Therefore, for the rest of this demonstration, we shall use the t-statistics weighting approach by default.

Figure 42 Comparison of Factor Weighting Schemes



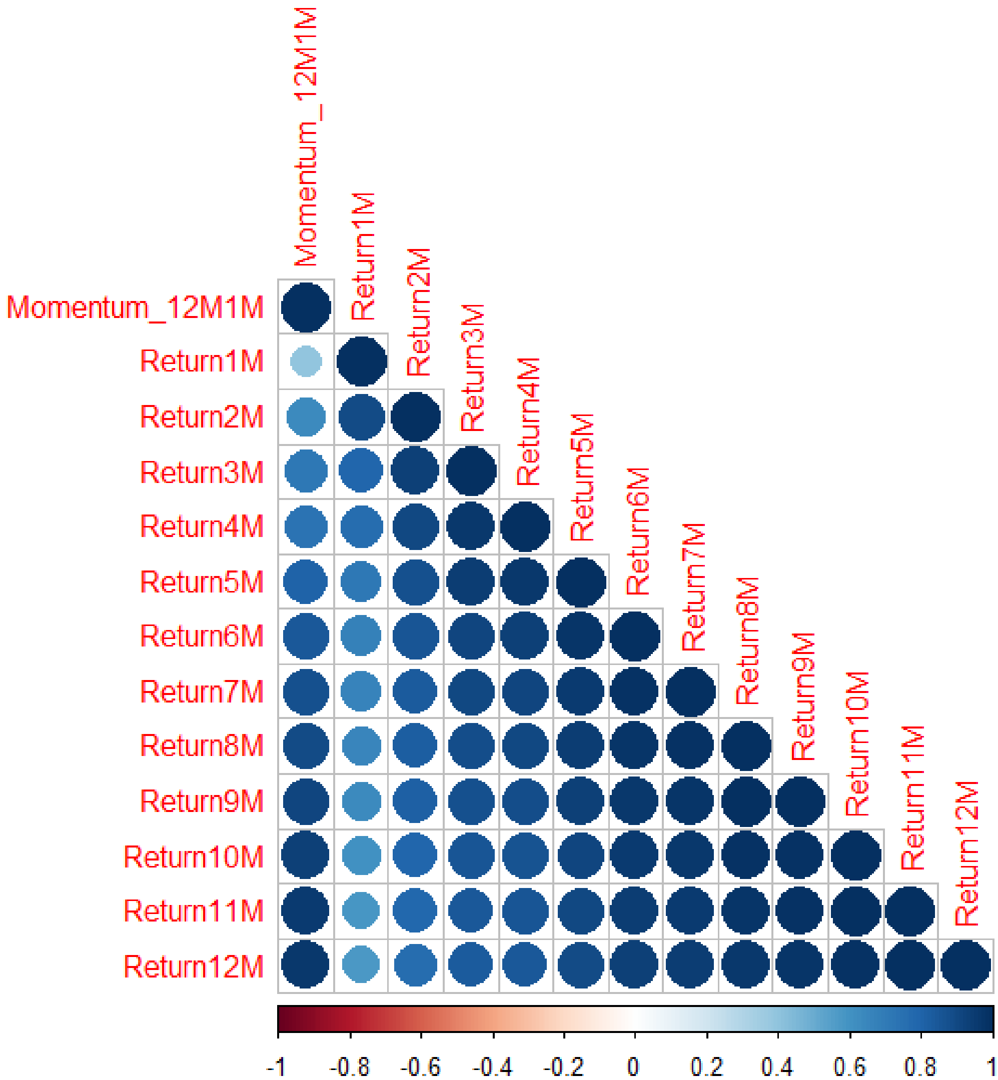
Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

Robustness to Multicollinearity

To measure the sensitivity to highly correlated factors, we add six more price momentum factors to the baseline model – past 6-, 7-, until 11-month returns. The six price momentum factors are closely related with our 12M-1M price momentum signal.

For example, Figure 43 shows the correlation matrix of 1-month, 2-month, until 12-month price momentum factors, along with our standard 12M-1M factor. The dark blue shades among the six-, seven-, to 12-month price momentum factors indicate high correlation of over 80%.

Figure 43 A Correlation Matrix of Price Momentum Factors



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

As we have learned in basic econometrics, multicollinearity (i.e., regressors are highly correlated) is addressed as a serious threat to regression. The argument is, if two or more variables are highly correlated, the coefficient estimates may change erratically, even for a small change in data. Multicollinearity is not supposed to impact the model's explanatory power too much, but it can greatly change the statistical inference of factors. In this section, we provide an empirical analysis on the true impact of multicollinearity in factor selection and model performance.

On the surface, adding multiple correlated factors to the Fama-MacBeth regression does not seem to cause much damage. As shown in Figure 44 (A), the overall t-statistics for the eight factors, even after we add the six highly correlated price momentum factors, are very similar to the original eight-factor model (see Figure 41 A). However, if we examine the difference in regression coefficients more carefully, the differences are considerable. As shown in Figure 44 (B), six out of the eight factors are statistically significant, with the largest impact on the original price momentum factor. Figure 44 (C) shows the time series difference in coefficient for the original price momentum factor, based on the two regressions (i.e., the baseline model and the multicollinearity model).

Moreover, as shown in Figure 44 (D), the t-statistics for the six newly added price momentum factors exhibit an odd pattern. The eight-month, nine-month, and 11-month momentum factors are statistically significant, but show different signs. The multicollinearity model essentially bets on the nine- and 12-month momentum factors, but offsets with eight- and 11-month momentum signals.

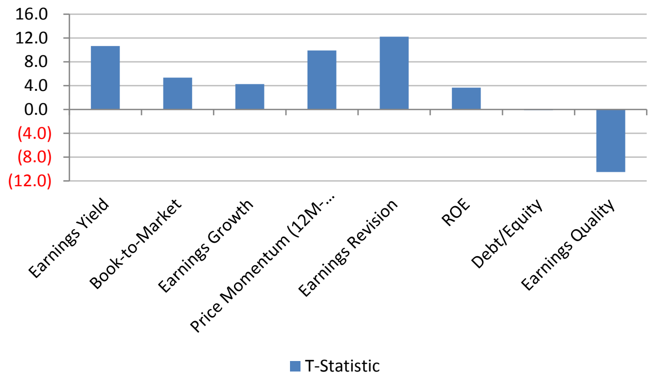
Lastly, we backtest the multicollinearity model, over the same period as our baseline case. We weight all the factors (the eight original factors plus the six additional price momentum factors) based on the t-statistic of each factor. Now we can formally assess the impact of having highly correlated factors in the model on performance. As shown in Figure 44 (E) and (F), multicollinearity does lower model performance, but the impact is modest, with the exception of Japan. In Japan, because price momentum is not a pricing factor, adding six more highly correlated (and non-informative) price momentum factors introduces a substantial noise.

Therefore, the overall conclusions on adding highly correlated factors are:

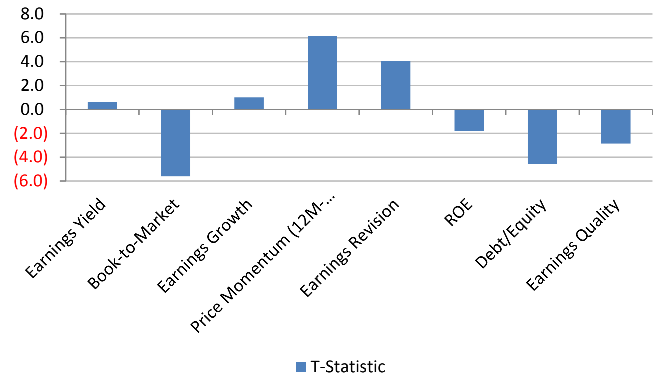
- As long as the factors all have reasonable predictive power, it reduces model performance, but the damage is modest.
- Adding highly correlated *uninformed* factors tends to weaken model performance notably.
- The statistical inference on whether a factor is important or not is small in the long term, but can be substantial at a given point-in-time.
- The regression coefficients on highly correlated factors are erratic and can be counter intuitive, which hurts a researcher's ability to interpret the model properly.
- Overall, the Fama-MacBeth regression is robust to multicollinearity. However, removing highly correlated factors should still improve model performance and interpretability.

Figure 44 The Impact of Multicollinearity

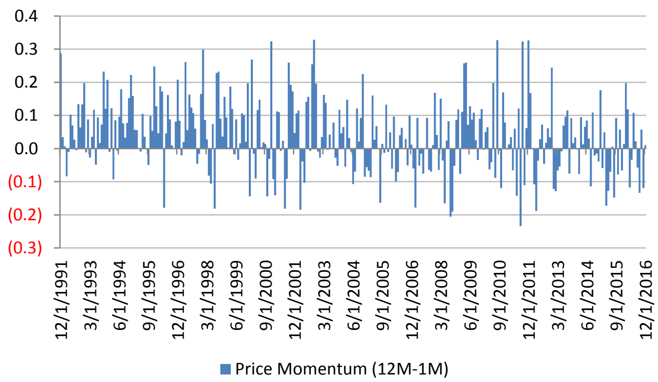
A) Multicollinearity Model, T-Statistics



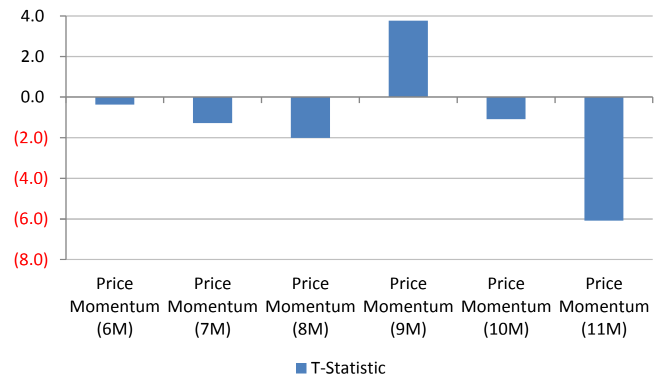
B) The Difference between the Multicollinearity Model and the Eight-Factor Model, T-Statistics



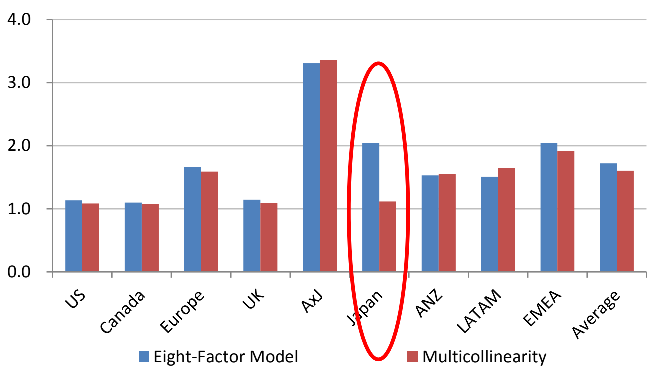
C) The Difference of Regression Coefficient for Price Momentum



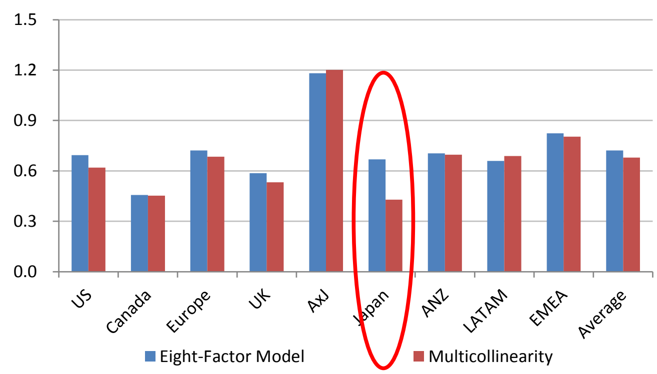
D) Correlated Factors, T-Statistics



E) The Impact of Multicollinearity on Performance, Sharpe Ratio



F) The Impact of Multicollinearity on Performance, Risk-Adjusted IC



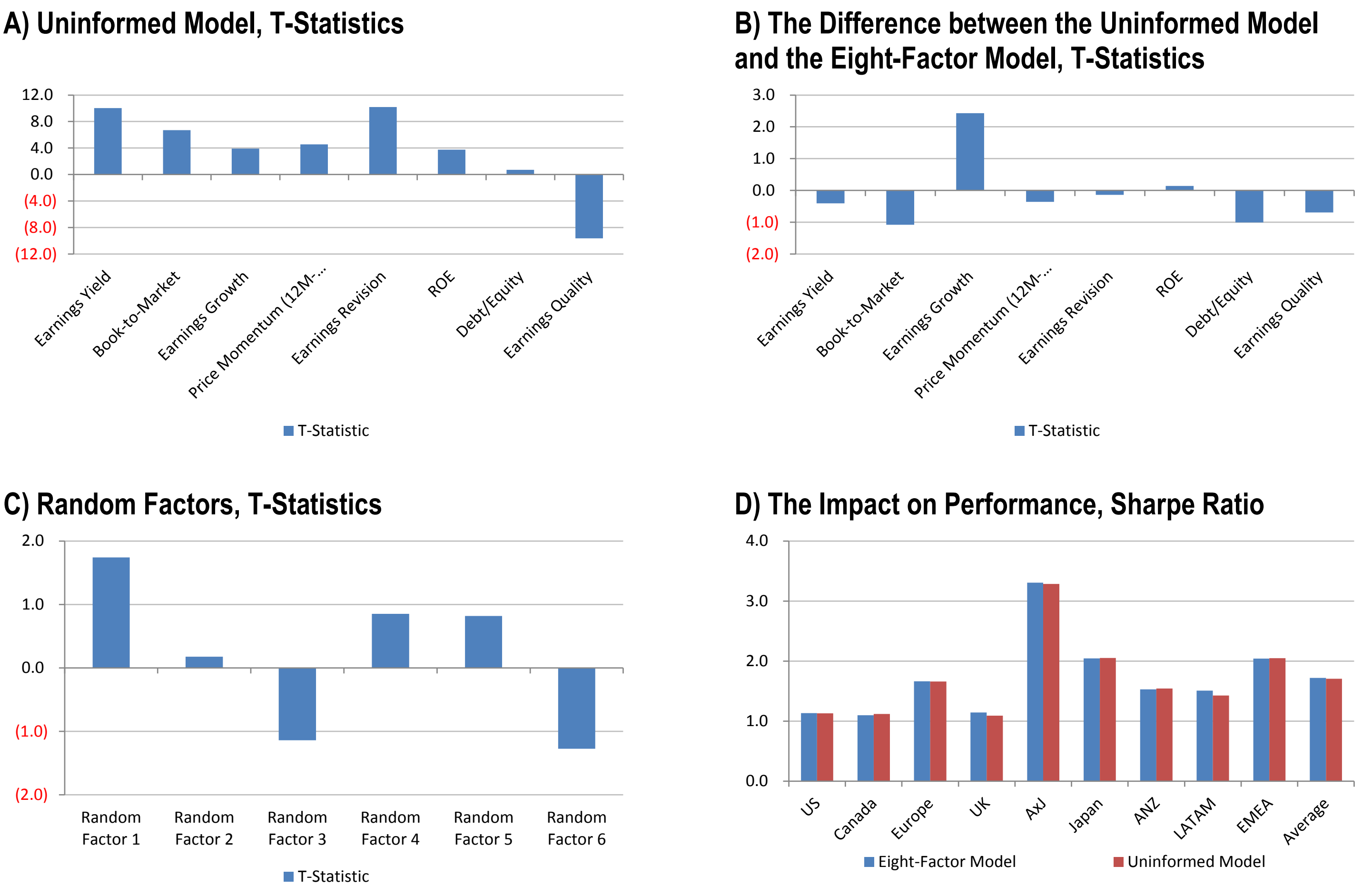
Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

Robustness to Pure Random Factors

In this section, we want to gauge the Fama-MacBeth procedure's ability to fend off noise or uninformed factors. For this purpose, we add six completely random factors, simulated independently from a standard normal distribution. Although for a given point-in-time, the random factor may show a statistically significant relationship with stock returns, we expect the Fama-MacBeth t-statistics to be able to identify the true relationship over time.

The Fama-MacBeth regression is fairly robust to random noise. Adding the six random factors does not appear to change the statistical significance of the original eight factors (see Figure 45 A). The difference in regression coefficient is minimal for seven out of the eight factors (see Figure 45 B). The model is also successful in identifying that none of the six random factor is meaningful in predict future returns (see Figure 45 C). Lastly, adding random noises does not materially impact model performance (see Figure 45 D).

Figure 45 The Impact of Random Noise



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

COMMON FACTOR WEIGHTING APPROACHES

In addition to Fama-MacBeth regression, there are a few other common factor weight approaches, e.g., equally weighting, risk parity, global minimum variance, and Grinold & Kahn.

Equal Weighting Scheme (EQW)

The simplest way to weight factors is to equally weight them. Surprisingly, it is actually not easy to beat this simple algorithm. Under this scheme, we make almost no assumption about factor return distribution; while in finance fewer assumptions often lead to more robust models.

Risk Parity or Equal Risk Contribution Weighting Scheme (ERC)

Another popular yet simple and robust factor weighting scheme is risk parity or equal risk contribution (ERC). Risk parity was originally developed and then populated in the asset allocation space[11]. Traditional pension funds tend to follow the so-called 60-40 allocation (i.e., 60% of the portfolio is allocated to equities and 40% is on fixed income). The problem is that equity is much more volatile then fixed income; therefore a 60-40 allocation portfolio is often dominated by equity risk. Risk parity allocation, on the other hand, focuses on allocation by risk. It is also known as equal risk contribution or ERC, in that each asset in the portfolio is forced to have equal risk budget (subject to other portfolio constraints). It is equally applicable in factor allocation, as we can treat factors the same way as asset classes. We will provide full empirical backtesting on the risk parity model in the next section.

Mathematically, the risk parity portfolio is constructed by the following optimization algorithm:

$$argmin_{\omega} \sum_{i=1}^{K} \sum_{j=1}^{K} \left[\omega_{i,t} COV\left(r_{i,t}, r_{p,t}\right) - \omega_{j,t} COV\left(r_{j,t}, r_{p,t}\right)\right]^2$$

Subject to:

$\sum_{i=1}^{K} \omega_{i,t} = 1$, and

$\omega_{i,t} \geq 0$

Where, $r_{i,t}$ and $r_{p,t}$ are the returns of asset i and portfolio p at time t, respectively; $\omega_{i,t}$ is the weight (to be determined by the algorithm) of asset i at time t; $COV\left(r_{j,t}, r_{p,t}\right)$ is the covariance between asset i and portfolio p; and K is the total number of assets.

Risk parity factor weighting is closely related to the smart beta allocation. In factor investing, each factor or risk premia is considered as a separate asset class. The common belief is that most risk premia factors will continue to generate returns, but it is difficult (or infeasible) to predict their returns. Therefore, risk parity allocation is naturally employed. A comprehensive introduction on risk parity can be found in Qian [2016].

Global Minimum Variance Weighting Scheme (GMV)

Similar to the Risk Parity Weighting Scheme, the Global Minimum Variance (GMV) algorithm only depends on the estimation of the covariance matrix among the factors, without requiring analysts to predict factor returns. As a result, it is straightforward to implement. The empirical performance of the

[11] Edward Qian at Panagora first coined the term "risk parity" in 2005 (see Qian [2005]).

GMV models also tends to be fairly competitive relative to more complex algorithms. The GMV portfolio weights can be solved by the following special case of Mean-Variance Optimization (MVO):

$$argmin_{\omega} \frac{1}{2} \omega' \Sigma \omega$$

Subject to:

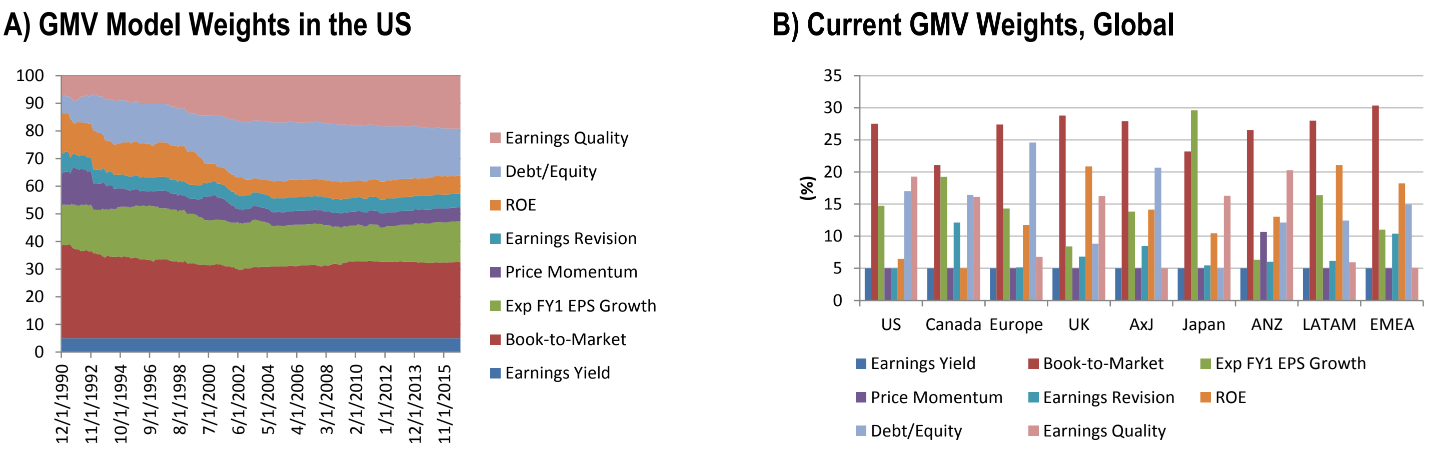
$\omega' i = 1$, i.e, the weights add up to 100%, and

$\omega \geq 0$, i.e., no shorting constraint

As a special case of MVO, GMV models are often dominated by few factors and therefore, in practice, we typically add a minimum weight constraint. For demonstration purpose, we set the minimum weight as 5% in our simulation (i.e., $\omega \geq 5\%$). We start the first GMV model once we have five-year of data. We use an expanding window to estimate the sample covariance matrix, for each of the nine regions.

As shown in Figure 46, the GMV model weights are relatively stable over the past 16 years in the US. Four factors (book-to-market, expected FY1 earnings growth, debt/equity ratio, and earnings quality) dominate the portfolio, while the weights of the other four factors mostly stay at the 5% minimum weight constraint. Globally, the weights from MV model are also somewhat surprising – book-to-market and debt/equity ratio, two factors that have unremarkable performance take the largest weights. In addition, the GMV models tend to be fairly concentrated and dominated by few factors. Lastly, two or more factors often reach the minimum weight constraint, suggesting that the GMV models are sensitive to portfolio constraints.

Figure 46 Global Minimum Variance Model Weights



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

Grinold & Kahn Weighting Scheme (G&K)

None of the above three approaches takes factor return efficacy into account. Equal weighting scheme ignores factor return distribution entirely, while ERC and GMV algorithms focus on risk allocation. The Grinold & Kahn [see Grinold and Kahn [1999] and Qian, Hua, and Sorensen [2007]) method is essentially a mean-variance optimization on the factor space. Mathematically, it is to maximize the expected IR (Information Ratio):

$$argmax_{\omega} IR = \frac{\omega' \widetilde{IC}}{\sqrt{\omega' \Sigma_{IC} \omega}}$$

Where,

ω is a $(K \times 1)$ vector of factor weights;

$\widetilde{IC}$ is a $(K \times 1)$ vector of expected factor IC; and

Σ_{IC} is a $(K \times K)$ covariance matrix of factor IC.

The above optimization problem has a closed-form solution, if we do not have any constraints:

$$\widehat{\omega} = \Sigma_{IC}^{-1} \widetilde{IC}$$

$\widehat{\omega}$ can be further normalized as $\widetilde{\omega} = \frac{\widehat{\omega}}{i' \cdot \widehat{\omega}}$, where i is a $(K \times 1)$ vector of 1's; then the weights add up to one.

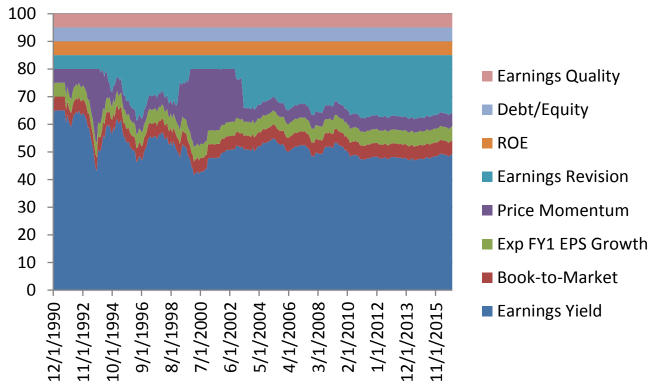
The Grinold & Kahn (G&K) weighting scheme critically depends on two input parameters – expected factor IC's and expected factor risk (covariance matrix). Given that in most applications, the number of factors K is small, estimating the factor covariance matrix is normally not too difficult – even sample covariance matrix is probably sufficient. On the other hand, estimating future factor IC's is likely to be challenging. As we will discuss in more details in the *Style Rotation, Machine Learning, and the Next Frontier of Systematic Investing* paper, factor returns are not stable, subject to structural breaks, and notoriously difficult to predict. In practice, many researchers use the long-term average IC's as a substitute for forecasting future IC's.

The G&K weighting scheme, similar to a typical mean-variance optimization, often produces extreme results. Compared to the GMV model weights (see Figure 46 A), the G&K model is even more concentrated (see Figure 47 A). Although we have eight input factors, the optimized model essentially bets on two factors – earnings yield and earnings revision. These two factors show the best performance in the past 30 years. Therefore, in practice, most managers add some constraints to the optimizer. In our case, similar to the GMV model, we set the minimum weight as 5%. As a result, six out of the eight factors stay at the minimum weight most of the time. Figure 47 (B) shows the weights of the G&K model for the nine regions. Similar to the US, there are a number of issues with the G&K model:

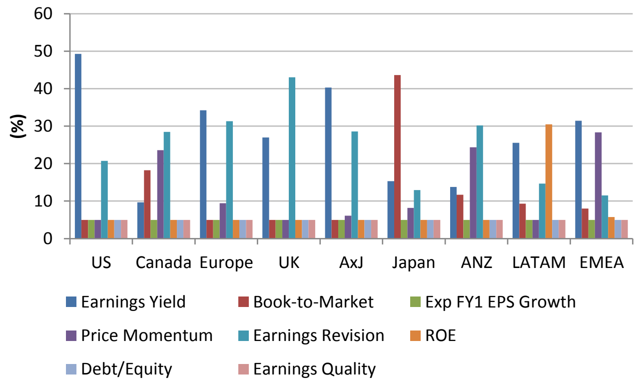
- Weights tend to be highly concentrated to two or three factors;
- Sensitive to the minimum weight constraint; and
- Easily affected by factor return prediction

Figure 47 Grinold & Kahn Model Weights

A) G&K Model Weights in the US



B) Current G&K Model Weights, Global



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

Representative Factor Models

There are unlimited ways to construct multifactor models. In this section, we use three factor models to represent the space of traditional multifactor stock selection models. In Part III of this research series, we will introduce our proprietary model – LEAP.

Global in Nature

All of our research is global in nature. Based on our previous research and experience, we divide the world into 10 regions: US, Canada, LATAM, Europe ex UK, UK, Emerging EMEA, Asia ex Japan, Japan, Australia and New Zealand, and China A (forthcoming). Currently, all of our models discussed in this paper include the first nine regional equity models.

Benchmark Model (BM)

The benchmark multifactor model is a proxy of typical multifactor models used by most quantitative managers. We choose eight popular factors from each main style category:

- Value: Trailing earnings yield – we prefer companies with high earnings yield
- Value: Book-to-market – we buy companies with high book-to-market, i.e., cheap stocks on valuation
- Growth: Consensus FY1/FY0 EPS Growth – we prefer companies with high earnings growth
- Price Momentum: 12M total return excluding the most recent month – we prefer companies with positive price momentum
- Analyst Sentiment: 3M EPS revision – we buy companies with positive earnings revisions
- Quality – Profitability: Return on equity – we like firms with high ROEs
- Quality – Leverage: Debt/Equity ratio – we prefer companies with low financial leverage
- Quality – Earnings Quality: Sloan's accruals – we buy companies with low accruals

Given the goal of representing the mainstream of quantitative investing, the BM has its own pros and cons:

Pros

- Intuitive – the underlying factors all have solid academic underpinning;
- Transparent – factors are static over time; and
- Simple and therefore, it has low risk of overfitting.

Cons

- The underlying factors have been widely studied and used in the investment community; therefore, they are more likely to be arbitraged away and have high risk of crowded trade; and
- The factors and their weights in the model are static, and ignore the underlying macroeconomic environment and factor turnover.

Factor Weighting

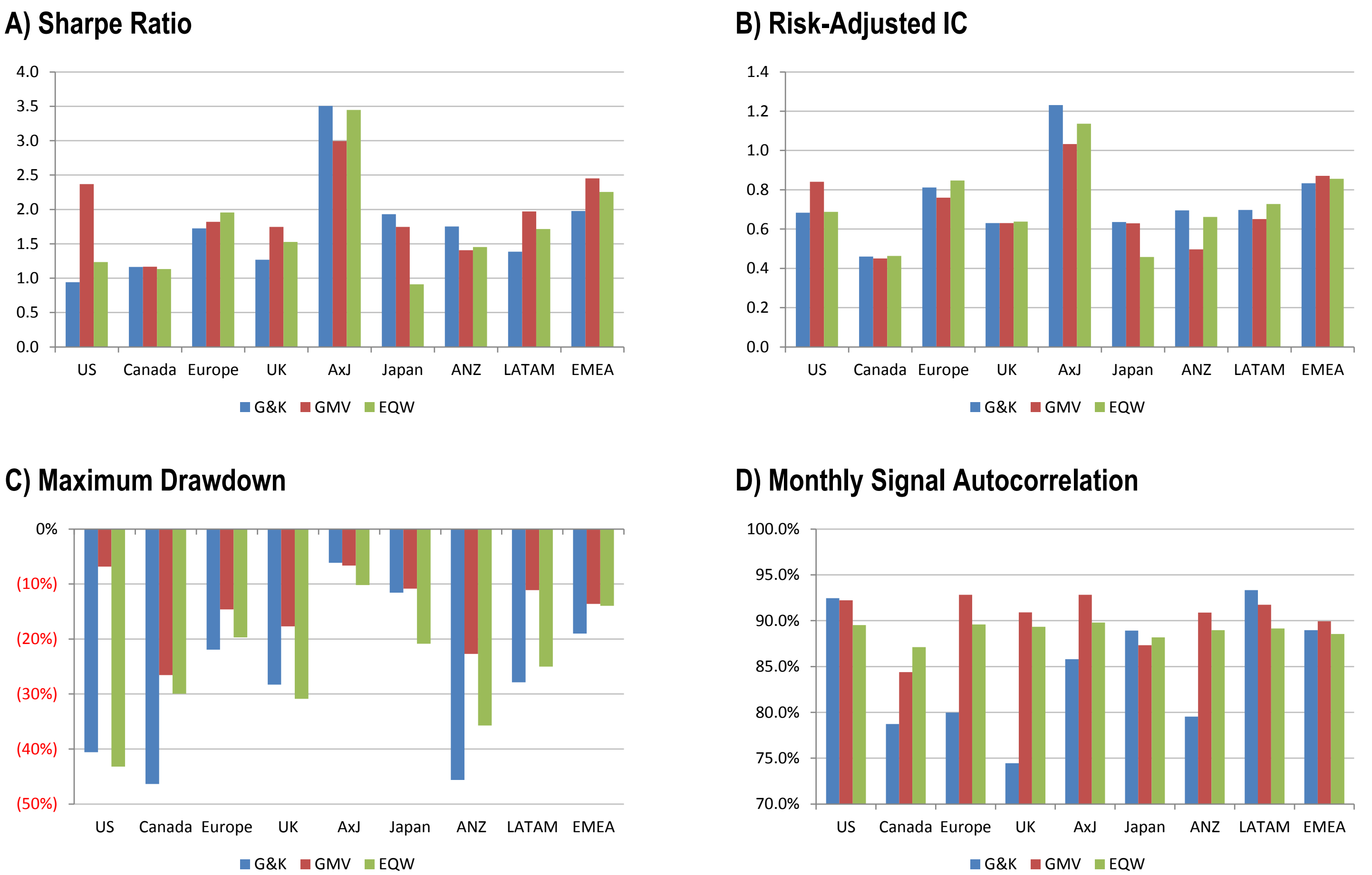
For the same philosophy as factor selection, we look for a simple yet robust factor weighting algorithm for the BM. Therefore, we choose three candidates weighting schemes: the equal weighting (EQW), the global minimum variance (GMV), and the Grinold & Kahn (G&K).

Figure 48 compares the performance of the three weighting schemes for each of the nine regions. There are a number of interesting observations:

- The performance of the three weighting schemes is generally comparable;
- There is no single model that dominates the performance in all regions;
- EQW model, despite of being extremely simple, is very competitive to the other optimization-based algorithms;
- GMV does lead to the lowest downside risk; and
- The G&K model does not quite meet its promise with the largest drawdown and turnover in most regions

Therefore, for simplicity, we choose the EQW scheme for the BM.

Figure 48 A Comparison of the three Factor Weighting Algorithms



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

Alternative Beta Portfolio (ABP)

The Alternative Beta Portfolio (ABP) serves a similar purpose as the BM. While the BM is the benchmark for multifactor models, the ABP is the benchmark for alternative beta portfolios. As discussed in earlier sections, the so-call alternative beta, smart beta, risk premia or factor investing has gained tremendous popularity in the past few years. For convenience, we use alternative beta throughout this research.

Alternative beta products can be constructed in many asset classes in equities, fixed income, currencies, commodity futures, etc. The traditional beta, i.e., CAPM beta based on the Capital Asset Pricing Model, refers to market capitalization weighted indices. A long list of research has been criticizing that the market cap weighted indices are highly inefficient. We are not getting into the debate on the merit or fault of these products in this research. Rather, we target from a practical point of view. In this research, we narrowly define alternative beta as factor portfolios.

There are many ways to construct factor portfolios. They can be either long-only or long/short. Stocks in the portfolio can be equally weighted in a naïve way, or an optimized portfolio can be built using a risk model. For demonstration purpose, we use the same eight factors in the BM and construct simple monthly rebalanced long/short quintile portfolios for each of the nine regions. The long/short quintile portfolios are country neutralized for Europe, AxJ, LATAM, and emerging EMEA; and sector neutralized for the US, Canada, UK, Japan, and ANZ.

Once the eight factor portfolios (i.e., alternative beta portfolios) are created, we apply the risk parity algorithm to weight these portfolios and form an overall asset allocation framework, for each of the nine regions.

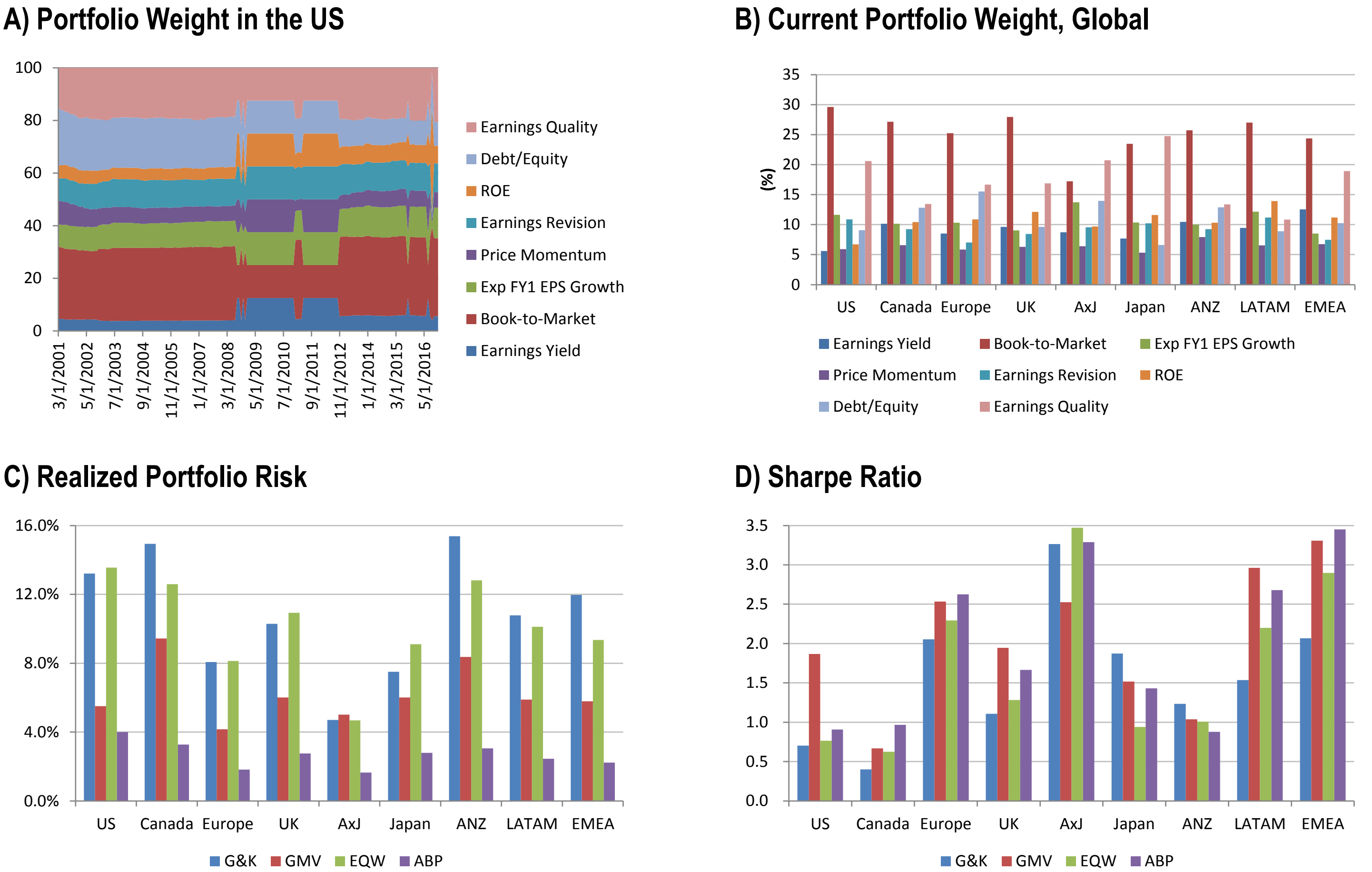
Therefore, active quantitative managers choose their own factors to build multifactor models, and then they use their models to construct portfolios. On the other hand, alternative beta products are essentially single factor portfolios. Asset owners then decide how they want to invest and mix these alternative beta portfolios. Risk premia products give asset owners and global macro investors more investable instruments for their asset allocation decisions.

The weights of the eight alternative beta portfolios in ABP are relatively stable over time (see Figure 49 A). Globally, the ABP portfolio mostly overweighs book-to-market and earnings quality products (see Figure 49 B)

The combined ABP has much lower realized volatility than the active multifactor BM model, in all regions (see Figure 49 C). ABP (along with GMV) tends to have higher Sharpe ratios than the two traditional multifactor models (EQW and G&K). However, we need to consider two additional points:

- The three traditional factor weighting algorithms either do not predict factor returns (i.e., EQW and GMV) or have naïve forecast (i.e., G&K). Active managers could have far more sophisticated factor timing and style rotation models. We will discuss this latter topic in our LEAP model.
- All returns are pre-transaction costs. In practice, active portfolios tend to be more efficient than alternative beta investing. In alternative portfolios, one stock might be held as a long position in one factor portfolio, while it could be shorted in another at the same time. Therefore, alternative beta portfolios do not engage netting and could incur higher costs.

Figure 49 Alternative Beta Portfolios



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

The Least Crowded Model (LCM)

In addition to the ABP, we further create a more sophisticated portfolio called the Least Crowded Model (LCM), using our proprietary minimum tail dependence algorithm. The LCM is dynamic, in the sense that factor weights are dynamically re-adjusted monthly. However, since we are using a 10-year rolling window, the weights only change modestly from month to month.

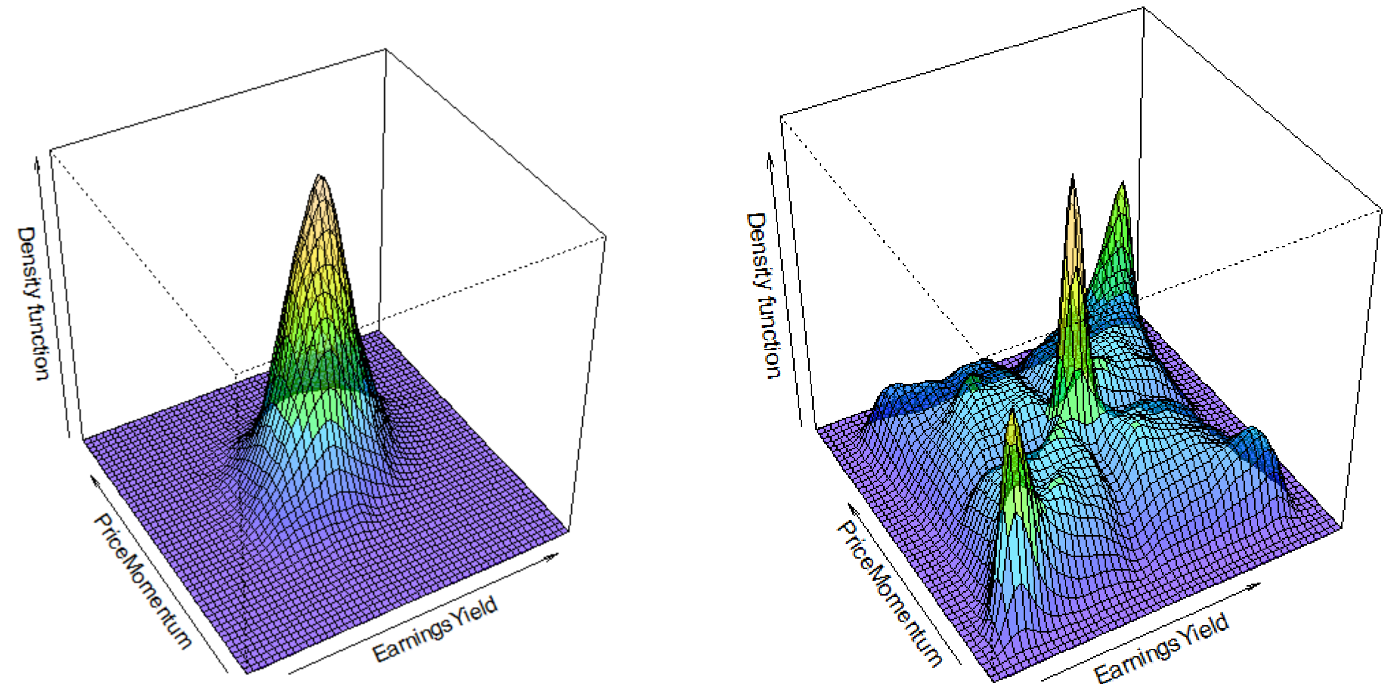
Introducing Tail Dependence

The tail dependence coefficient is similar to the correlation coefficient, but emphasizes on the statistical convergence at the tail. It is estimated via a Copula model (see McNeil, et al [2005] for a textbook treatment of the Copula model).

Figure 50 (A) shows the theoretical bivariate normal distribution between value and momentum factor portfolios. The distribution follows a classic bell-shaped curve. However, if we plot the real empirical distribution between value and momentum, we see something quite different (see Figure 50 B). The joint distribution is clearly tri-modal with three distinct peaks (or modes) at the center (implied by bivariate normal distribution) and two tails, i.e., the probabilities of these two styles both move higher and fall lower are much higher than what is implied by normal distribution.

Figure 50 The Importance of Using Tail Dependence

A) Theoretical Bivariate Normal Distribution **B) Empirical Distribution**



Sources: Wolfe Research Luo's QES

Extending beyond value and momentum, Figure 51 shows the difference between traditional correlation coefficients (upper triangle) and tail dependence (lower triangle) for the eight factors. Tail dependence tends to be higher than correlation, indicating that factors are more likely to be crowded in crises. For example, the correlation between value and momentum is at a modest 33% level, but the tail dependence is almost 30% higher at 42%. Therefore, traditional factor weighting algorithms that only consider mean and variance may be subject to crowded trades with large downside risk.

Figure 51 Correlation versus Tail Dependence

	EPSYield	B/P	Growth	Momentum	Revision	ROE	Debt/Equity	Quality
Earnings Yield	100%	(20%)	12%	33%	9%	86%	4%	13%
Book-to-Market	6%	100%	(56%)	(77%)	(69%)	(54%)	(45%)	5%
Earnings Growth	29%	0%	100%	61%	67%	33%	28%	(7%)
Price Momentum	42%	0%	47%	100%	71%	58%	41%	6%
Earnings Revision	28%	0%	60%	55%	100%	37%	39%	(2%)
ROE	66%	1%	38%	55%	43%	100%	21%	7%
Debt/Equity	23%	0%	31%	48%	42%	30%	100%	(11%)
Earnings Quality	14%	6%	9%	13%	6%	13%	2%	100%

Note: Pearson correlation coefficients on the upper triangle; tail dependence coefficients on the lower triangle

Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

Minimum Tail Dependence Optimization

Once we estimate the tail dependence coefficients among all of our factors, the minimum tail dependence portfolio can be constructed using a three-step optimization:

Step 1

$$argmin_{\omega} \frac{1}{2} \Psi_t{}' \mathrm{T}_t \Psi_t$$

Subject to:

$$\Psi_t' \iota = 1$$

$$\Psi_t \geq 0$$

Where,

Ψ_t is the first intermediate vector of asset weights to be solved at time t, and

T_t is the asset-by-asset tail dependence matrix at time t.

Step 2

We then rescale the first intermediate vector of asset weights Ψ_t by each asset's volatility $\sigma_{i,t}$:

$\Phi_t = \Delta_t^{-1/2} \Psi_t$, i.e., $\phi_{i,t} = \psi_{i,t} / \sigma_{i,t}$

Where,

Φ_t is the second intermediate vector of asset weights at time t, and

Δ_t is the diagonal matrix of asset variance at time t, with $\sigma_{i,t}^2$ at its (i, i) element and zero on all off-diagonal elements.

Step 3

Finally, we need to rescale the second intermediate asset weight vector (Φ_t), so the sum of the final weights equal to 100%:

$$\omega_{i,t} = \frac{\phi_{i,t}}{\sum_{j=1}^{K} \phi_{j,t}}$$

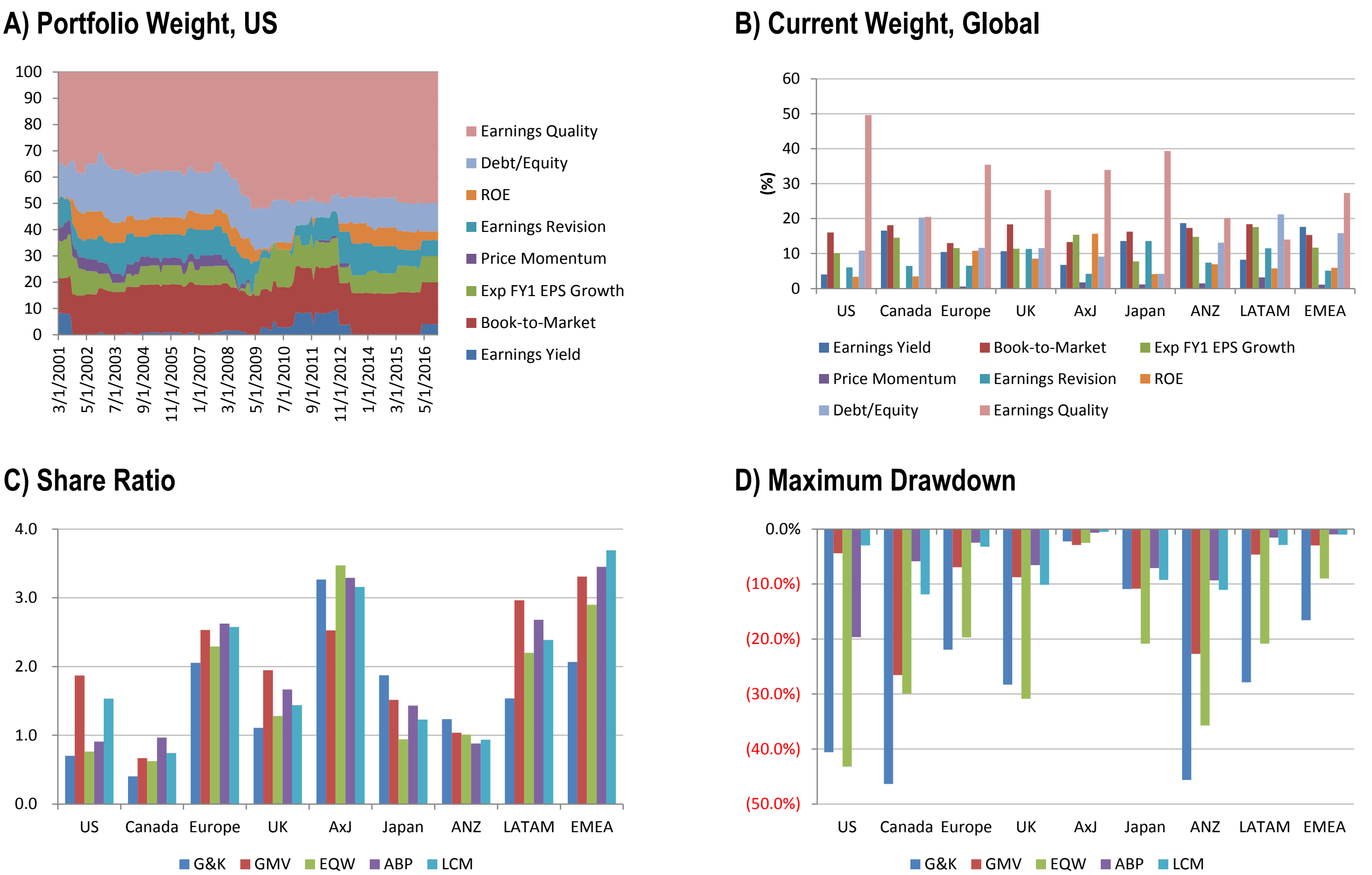
The intuition behind our minimum tail dependence strategy is that we overweight those assets least correlated to other assets at the tail level. By construction, the LCM is also the least crowded strategy, as it avoids crowded trades by other investors.

The Empirical Performance of the LCM

As shown in Figure 52 (A), the LCM takes a very different view from the other portfolios (e.g., EQW, GMV, G&K, ABP), with a sizable bet on earnings quality and minimal weight on earnings yield. Earnings quality portfolio takes the largest weight in almost all regions (see Figure 52 B).

The LCM portfolio is also very robust, showing highly competitive Sharpe ratio (see Figure 52 C), with almost the lowest downside risk in most regions (see Figure 52 D).

Figure 52 Least Crowded Portfolio



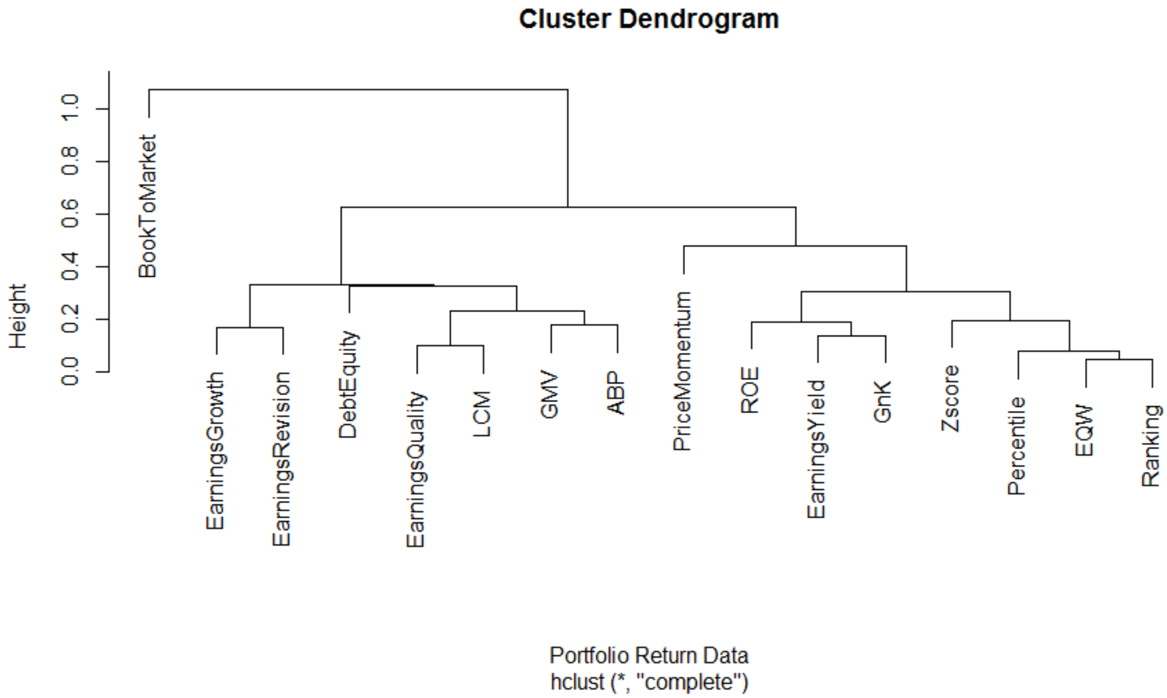
Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

Among all the factor weighting schemes and alternative beta portfolios, a cluster analysis reveals a few insightful observations (see Figure 53):

- The four data transformation techniques[12] (z-score, percentile, ranking, or our proprietary algorithm) are closely related.
- G&K factor weighting model forms a cluster with earnings yield and ROE.
- The two risk-based allocations – GMV and ABV fall into one cluster.
- Our LCM strategy is closely related to earnings quality.
- Book-to-market behaves differently from all other factors and strategies.

[12] The four data transformation techniques are discussed in Luo, et al [2017].

Figure 53 Cluster Analysis



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

FORTHCOMING RESEARCH

This paper forms the second part of our *QES Handbook of Active Investing* series. In the next few weeks, we will publish the other two key components:

- *Part III: Machine Learning, Style Rotation, and the Next Frontier in Systematic Investing*
- *Part IV: Risk, Portfolio Construction, Trade Execution, and Performance Attribution* – From Theory to Practice

In addition to the four-part introduction of Big Data and Machine learning in global equity investing, we are also working on a number of other issues:

- From Nowcasting to Forecasting – Economics and Portfolio Strategy in the New Age
- Industry-Specific Models in Global Banking and Insurance Industries
- Accounting Quality, Fraud Detection, and Corporate Governance
- Factors based on Alternative Data Sources
- Machine Learning in Global Stock Selection

BIBLIOGRAPHY

Amihud, Y. [2002]. “Illiquidity and Stock Returns: Cross-Section and Time-Series Effects”, *Journal of Financial Markets*, 5, 31-56.

Asness, C., Moskowitz, T., and Pedersen, L.H. [2012]. “Value and Momentum Everywhere”, SSRN Working Paper,

Bali, T.G., Cakici, N., and Whitelaw, R.F. [2011]. “Maxing Out: Stocks as Lotteries and the Cross-Section of Expected Returns”, *Journal of Financial Economics*, 99(2011), 427-446.

Bailey, D.H., Borwein, J.M., De Prado, M.L., and Zhu, Q.J. [2014]. “Pseudo-Mathmatics and Financial

Baltagi, B.H. [2005]. *Econometric Analysis of Panel Data*, Third Edition, West Sussex, England: John Wiley & Sons.

Basu, S. [1977]. “Investment Performance of Common Stocks in Relation to Their Price-Earnings Ratios: A Test of the Efficient Market Hypothesis”, *Journal of Finance*, Vol 32, No 3, 663-682.

Carhart, M.M. [1997]. “On Persistence in Mutual Fund Performance”, Journal of Finance, 52, 57-82

Cooper, M., Gulen, H., and Schill, M. [2008]. “Asset Growth and the Cross-Section of Stock Returns”, *Journal of Finance*, 63(4), 1609-1651.

Damodaran, A. [2012]. *Investment Valuation: Tools and Techniques for Determining the Value of Any Asset*, 3rd Edition, Wiley

De Bondt, W.F.M, and Thaler, R. [1985]. “Does the Stock Market Overreact?”, *Journal of Finance*, Vol 40, No 3

Ding, Z.X. [2010]. “The Fundamental Law of Active Management: Time Series Dynamics and Cross-Sectional Properties”, Working Paper

Fama. E.F., and MacBeth, J. [1973]. “Risk, Return, and Equilibrium: Empirical Tests”, *Journal of Political Economy*, May-June 1973.

Fama, E., and French, K.R. [1993]. “Common Risk Factors in the Returns on Stocks and Bonds”, *Journal of Financial Economics* 33(1): 3-56.

Fama, E., and French, K.R. [1996]. “Multifactor Explanations of Asset Pricing Anomalies”, *Journal of Finance*, Vol. 51, 55-84.

Fama, E., and French, K.R. [2014]. “A Five-Factor Asset Pricing Model”, Working Paper.

Graham, B., and Dodd, D. [1934]. *Security Analysis*, McGraw-Hill, 1934.

Grinold, R. [1989]. “The Fundamental Law of Active Management”, *Journal of Portfolio Management*, Vol 15, No. 3, 30-37

Grinold, R., and Kahn, R. [1999]. *Active Portfolio Management, A Quantitative Approach for Producing Superior Returns and Controlling Risk*, McGraw-Hill, 2nd Edition.

Harrell, F. [2001]. *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*, Springer, New York

Ilmanen, A. [2011]. Expected Returns, John Wiley & Sons

Jegadeesh, N. [1990]. “Evidence of Predictable Behavior of Security Returns”, *Journal of Finance*, 45: 881-898.

Jegadeesh, N., and Titman, S. [1993]. “Returns to buying winners and selling losers: implications for stock market efficiency”, *Journal of Finance*, Vol. 28, No. 1, 63, March 1993.

John, G.H., Kohavi, R., and Pfleger, K. [1994]. “Irrelevant Features and the Subset Selection Problem”, *Machine Learning: Proceedings of the 11th International Conference*, 121-129

Kroll, B., Trichilo, D., Braun, J. [2005]. “Extending the Fundamental Law of Investment Management”, JP Morgan Asset Management

Lakonishok, J., Shleifer, A., and Vishny, R.W. [1994]. “Contrarian

Li, X., and Sullivan, R.N. [2015]. “Investing in the Asset Growth Anomaly across the Globe”, *Journal of Investment Management*, Vol 13, No 4, pp. 87-107.

Luo, Y., Jussa, J., and Wang, S. [2016]. *Journal of Quantitative, Economics, and Strategy*, Wolfe Research Luo’s QES, December 12, 2016

Luo, Y., Jussa, J., and Wang, S. [2017]. “The Big and the Small Sides of Big Data”, Wolfe Research Luo’s QES, February 8, 2017.

McNeil, A. Frey, R., and Embrechts, P. [2005]. *Quantitative Risk Management: Concepts, Techniques, and Tools*, Princeton University Press, Princeton, NJ

Newey, W., and West, K. [1987]. “Hypothesis Testing with Efficient Method of Moments Estimation”, *International Economic Review*, 28, 777-787.

Qian, E. [2005]. “Risk Parity Portfolios”, Panagora, https://www.panagora.com/assets/PanAgora-Risk-Parity-Portfolios-Efficient-Portfolios-Through-True-Diversification.pdf

Qian, E.E. [2016]. *Risk Parity Fundamentals*, Chapman and Hall/CRC, February 23, 2016

Qian, E.E., Hua, R.H., and Sorensen, E.H. [2007]. *Quantitative Equity Portfolio Management: Modern Techniques and Applications*, Chapman & Hall, 1st Edition

Saeys, Y., Inza, I., and Larranaga, P. [2007]. “A Review of Feature Selection Techniques in Bioinformatics”, *Bioinformatics*, Vol 23, No. 19, 2507-2517

Sloan, R.G. [1996]. “Do Stock Prices Fully Reflect Information in Accruals and Cash Flows about Future Earnings”, *The Accounting Review* 71(3), 289-315.

DISCLOSURE SECTION

Analyst Certification:

The analyst of Wolfe Research primarily responsible for this research report whose name appears first on the front page of this research report hereby certifies that (i) the recommendations and opinions expressed in this research report accurately reflect the research analysts' personal views about the subject securities or issuers and (ii) no part of the research analysts' compensation was, is or will be directly or indirectly related to the specific recommendations or views contained in this report.

Other Disclosures:

Wolfe Research, LLC does not assign ratings of Buy, Hold or Sell to the stocks it covers. Outperform, Peer Perform and Underperform are not the respective equivalents of Buy, Hold and Sell but represent relative weightings as defined above. To satisfy regulatory requirements, Outperform has been designated to correspond with Buy, Peer Perform has been designated to correspond with Hold and Underperform has been designated to correspond with Sell.

Wolfe Research Securities and Wolfe Research, LLC have adopted the use of Wolfe Research as brand names. Wolfe Research Securities, a member of FINRA (www.finra.org) is the broker-dealer affiliate of Wolfe Research, LLC and is responsible for the contents of this material. Any analysts publishing these reports are dually employed by Wolfe Research, LLC and Wolfe Research Securities.

The content of this report is to be used solely for informational purposes and should not be regarded as an offer, or a solicitation of an offer, to buy or sell a security, financial instrument or service discussed herein. Opinions in this communication constitute the current judgment of the author as of the date and time of this report and are subject to change without notice. Information herein is believed to be reliable but Wolfe Research and its affiliates, including but not limited to Wolfe Research Securities, makes no representation that it is complete or accurate. The information provided in this communication is not designed to replace a recipient's own decision-making processes for assessing a proposed transaction or investment involving a financial instrument discussed herein. Recipients are encouraged to seek financial advice from their financial advisor regarding the appropriateness of investing in a security or financial instrument referred to in this report and should understand that statements regarding the future performance of the financial instruments or the securities referenced herein may not be realized. Past performance is not indicative of future results. This report is not intended for distribution to, or use by, any person or entity in any location where such distribution or use would be contrary to applicable law, or which would subject Wolfe Research, LLC or any affiliate to any registration requirement within such location. For additional important disclosures, please see.

The views expressed in Wolfe Research, LLC research reports with regards to sectors and/or specific companies may from time to time be inconsistent with the views implied by inclusion of those sectors and companies in other Wolfe Research, LLC analysts' research reports and modeling screens. Wolfe Research communicates with clients across a variety of mediums of the clients' choosing including emails, voice blasts and electronic publication to our proprietary website www.wolferesearch.com/disclosures.

This report is limited for the sole use of clients of Wolfe Research. Authorized users have received an encryption decoder which legislates and monitors the access to Wolfe Research, LLC content. Any distribution of the content produced by Wolfe Research, LLC will violate the understanding of the terms of our relationship.