GLOBAL VOLATILITY SUMMIT

# JULY 2016 NEWSLETTER

Dear Investor,

The Global Volatility Summit ("GVS") brings together volatility and tail hedge managers, institutional investors, thought-provoking speakers, and other industry experts to discuss the volatility markets and the roles volatility strategies can play in institutional investment portfolios. The GVS aims to keep investors updated on the volatility markets throughout the year, and educated on innovations within the space.

Deutsche Bank has provided the latest piece in the GVS newsletter series on behalf of BlueMountain Capital. Part I of the newsletter is enclosed. Please refer to the GVS website for Part II.

Cheers,
Global Volatility Summit

# EVENT

The 7th Annual GVS was held on March 16th in New York City. Joined by the other event sponsors, including banks and exchanges, ten volatility and tail hedge managers hosted a crowd of 350 attendees including senior investment representatives from the largest global pensions, sovereign wealth funds, endowments, foundations, and insurance companies.

**2016 MANAGER PARTICIPANTS**
Argentière Capital
BlueMountain Capital
Capstone Investment Advisors
Capula Investment Management
Ionic Capital Management
Man AHL
Parallax Volatility Advisors
PIMCO
Pine River Capital Management
True Partner Capital

**2016 KEYNOTE AND GUEST SPEAKERS**
The 2016 keynote speakers were Barney Frank and Marcus Luttrell. Barney Frank served as a US Congressman for over 30 years and most recently as the Chairman of the House Financial Services Committee from 2007 through 2011. He was a key author of the Dodd-Frank Wall Street Reform and Consumer Protection Act. Marcus Luttrell is a decorated Navy Seal and best-selling author of Lone Survivor. You can access their biographies and more information about the event on the website: www.globalvolatilitysummit.com.

Deutsche Bank
Markets Research

Global

Quantitative Strategy
# Signal Processing

Date
16 May 2016


Javed Jussa
javed.jussa@db.com

Gaurav Rohal, CFA
gaurav.rohal@db.com

Sheng Wang
sheng.wang@db.com

George Zhao
zheyin.zhao@db.com

Yin Luo, CFA
yin.luo@db.com

Miguel-A Alvarez
miguel-a.alvarez@db.com

Allen Wang
allen-y.wang@db.com

David Elledge
david.elledge@db.com

Binky Chadha
bankim.chadha@db.com

Parag Thatte
parag.thatte@db.com

Rajat Dua
rajat.dua@db.com

North America: +1 212 250 8983
Europe: +44 20 754 71684
Asia: +852 2203 6990


# Strategy Crowding

## An in-depth analysis of style crowding and investor herding

### The notion of crowding is inherently confusing

Crowding is a perplexing concept. Investors are constantly on the hunt for reliable crowding measures, but what does crowding really mean? Is it aggregate sentiment, market timing, or the inverse of trending? Or is crowding the tipping point when the strategy reverts? As a strategy becomes modestly crowded, it should outperform (to a point) as crowding takes hold. Alternately, can crowding measures simply be the inverse of a trending measure (which are widely known)? Undoubtedly, the concept of crowding can be baffling.

### We explore a diverse set of potential crowding measures

We explore a multitude of crowding measures including: valuation spreads, strategy expensiveness, market breadth, sentiment, shorting costs, institutional buying and selling, institutional ownership, stock pairwise correlation, tail dependence, concentration ratio, and funds flow, etc. Our research suggests that there is no silver bullet measure of crowding. There are series of useful measures that can hint as to which strategies are "crowded". The diversification ratio, for example, is a robust measure of investor crowding.

### How to read our crowding report

This report is an in-depth study of various "crowding" measures ranging from fairly simple metrics to more sophisticated ones. We have outlined the key findings within the first few pages. For those who want more details or clarity on our analysis, it may be useful to setup a meeting in person, over the phone, or attend one of our many webinars.

*Source: gettyimages.com*

Table Of Contents

# A letter to our readers

It is often said that quant strategies follow an evolutionary path: they are discovered, work well for a time, and then get crowded and arbitraged away. In this paper, we evaluate whether this is actually true by developing proxies for factor crowdedness. In theory, developing measures for strategy crowdedness should be fairly straightforward. In practice, however, it is rather difficult because we do not have a framework to compare the effectiveness of our crowding measures.

How does one measure crowding? How does one know if an indicator is a strong measure of investor crowding? Crowding is a concept that is often misunderstood. Investor crowding is often associated with poor future performance or a strong pullback in a strategy. While this may be true to some degree, our findings show that crowding can also be a healthy ingredient for stronger future performance. However, too much of one spice can potentially lead to distasteful returns. So where is the balance?

In this novel research, we explore the fine line between performance and crowding. We analyze a multitude of crowding measures from simple indicators, such as expensiveness and market breadth, to more cutting edge measures, such as short interest, stock pairwise correlations, and the co-movement effects in stocks' tail distributions. We also analyze more grassroots measures of investor herding using institutional holdings and ownership as well as funds flow data. Portfolio measures of crowding that we explore include the concentration and diversification ratio.

To test these measures, we introduce a fairly extensive analytical framework. We propose four methods to test for investor crowdedness: polynomial trend analysis, classification and regression trees or CART, multivariate regressions, and correlation analysis. Each algorithm has several benefits and drawbacks. However, the framework that we outline provides a sound structure for assessing strategy crowding.

As you will see, we believe that crowdedness is an important additional metric to consider when allocating to an investment strategy, selecting a factor (or factor weight) in an alpha model, or when deciding how to position a portfolio based on the current and future macroeconomic environment. We hope you enjoy the remainder of this report. Please contact us at dbeqs.americas@db.com for any questions.

Regards,

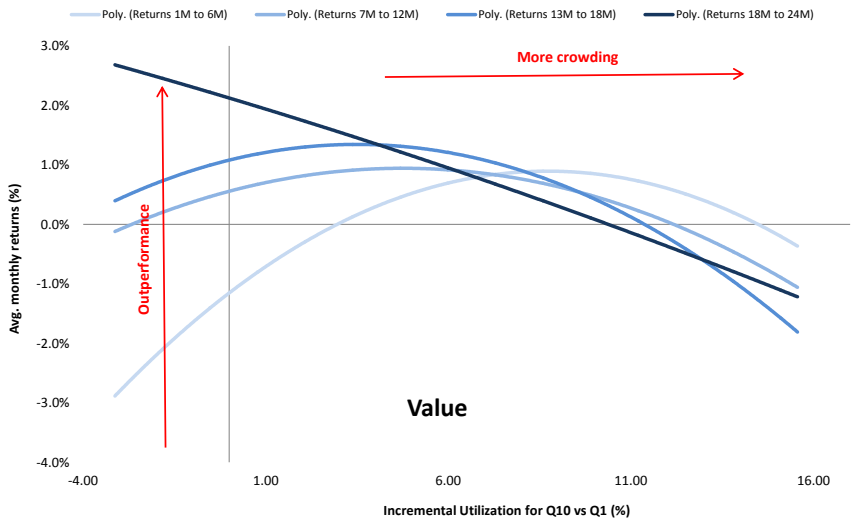
Yin, Javed, Gaurav, Sheng, and the quant team
**Deutsche Bank Quantitative Strategy**

# Five key findings

## 1. Crowding tends to be linked to future outperformance.

However, at extreme levels, crowded trades can lead to heighted risks of large drawdowns.

Figure 1: Crowding and the value portfolio



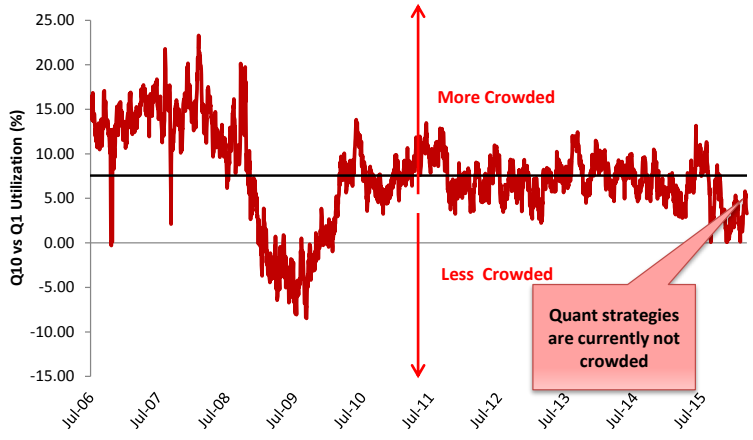

*Source: Bloomberg Finance LP, Compustat, Markit, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

*Figure 1 shows the forward returns of the value portfolio versus one measure of crowdedness. The value strategy tends to outperform as the strategy gets crowded. This outperformance tends to last for about 18 months. Thereafter, the value strategy tends to struggle rather significantly and with strong elasticity. Extreme levels of crowding (the very right side) are generally accompanied by future drawdowns and heighted risk.*

## 2. Generic quant strategies are currently not crowded.

Therefore, crowding is unlikely to be the reason for the challenging performance of most common factors in recent months.

Figure 2: Quant strategies and crowding



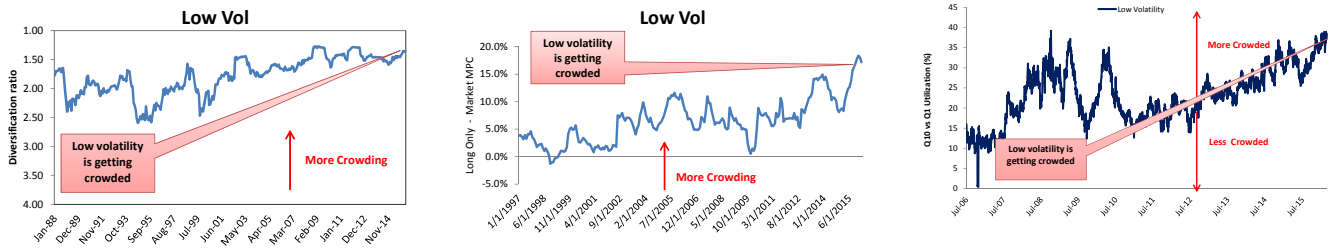

*Source: Bloomberg Finance LP, Compustat, Markit, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

*Figure 2 shows the crowdedness of a multi-factor model which is indicative of typical quantitative strategies. Currently, quant strategies do not appear to be crowded relative to historical levels*

## 3. However, low vol investing may be getting crowded

This is based on several measures we studied in this research.

Figure 3: Low volatility may be getting crowded based on the diversification ratio, pairwise correlations and utilization




*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Markit, Deutsche Bank*

## 4. Diversification Ratio is an effective crowding measure

In particular, it's a robust measure of crowdedness for low volatility and quality measures. Note that we explored over ten potential crowding measures.

*Figure 4 shows that the DR ratio is strongly negatively correlated to future annual returns for the low volatility and quality portfolios*

Figure 4: Correlation of DR and future annual returns

| | Div yield | Value | Momentum | Reversal | Growth | Quality | Low Vol | Sentiment |
|---|---|---|---|---|---|---|---|---|
| **DR Cap Weighted** | -4% | -20% | -19% | 35% | 11% | -49% | -64% | 35% |
| **DR Conviction Weighted** | 8% | -10% | -30% | 30% | -7% | -50% | -73% | 30% |

*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

## 5. See our crowding predictions over the next 12 months.

Figure 5: Risk and return prediction (12 month ahead)

**Return prediction (12 month ahead)**

| Factors | Expensiveness | Market breadth | Short interest | Buying power | Concentration ratio | Diversification ratio | Overall |
|---|---|---|---|---|---|---|---|
| Div yield | Neutral | Negative | Neutral | Neutral | Neutral | Neutral | Neutral |
| Value | Positive | Neutral | Positive | Positive | Positive | Positive | Positive |
| Momentum | Positive | Positive | Negative | Positive | Negative | Positive | Neutral |
| Growth | Positive | Neutral | Neutral | Neutral | Neutral | Negative | Neutral |
| Quality | Neutral | Negative | Positive | Neutral | Neutral | Neutral | Neutral |
| Low Vol | Negative | Neutral | Neutral | Neutral | Neutral | Negative | Negative |
| Sentiment | Neutral | Negative | Neutral | Positive | Neutral | Neutral | Neutral |

**Drawdown risk prediction (12 month ahead)**

| Factors | Expensiveness | Market breadth | Short interest | Buying power | Concentration ratio | Diversification ratio | Overall |
|---|---|---|---|---|---|---|---|
| Div yield | Modest | Modest | Severe | Modest | Severe | Severe | Modest |
| Value | Severe | Modest | Modest | Severe | Severe | Severe | Severe |
| Momentum | Severe | Severe | Severe | Severe | Modest | Modest | Severe |
| Growth | Modest | Modest | Modest | Modest | Modest | Severe | Modest |
| Quality | Severe | Severe | Modest | Severe | Severe | Modest | Severe |
| Low Vol | Severe | Severe | Modest | Severe | Severe | Severe | Severe |
| Sentiment | Modest | Severe | Severe | Modest | Modest | Modest | Modest |

*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

*Current prediction of future return is most consistent for value strategy with most proxies for crowdedness indicating an outperformance for the strategy over the next 12 months. On the other hand low vol remains the most crowded trade based on most proxies with a negative and volatile expectation on future returns*

# Classic measures

We begin by studying the conventional wisdom of strategy crowding. We explore whether strategy expensiveness (as measured by price to earnings multiples), historical performance, and sentiment conveyed by market breadth are possible crowding measures. Our analysis also provides some insight into the lead-lag effect of crowding and trending (i.e., does trending lead to crowding or vice versa).

## 1. Expensiveness

One method to potentially gauge investor crowding or herding is by looking at the relative expensiveness of a particular market, sector or strategy. If a strategy becomes overwhelmingly expensive, then this could be a potential crowding signal or indicate a potential turning point. To test this hypothesis, we simply compute the average price to earnings (P/E) for a particular market, sector, and strategy.

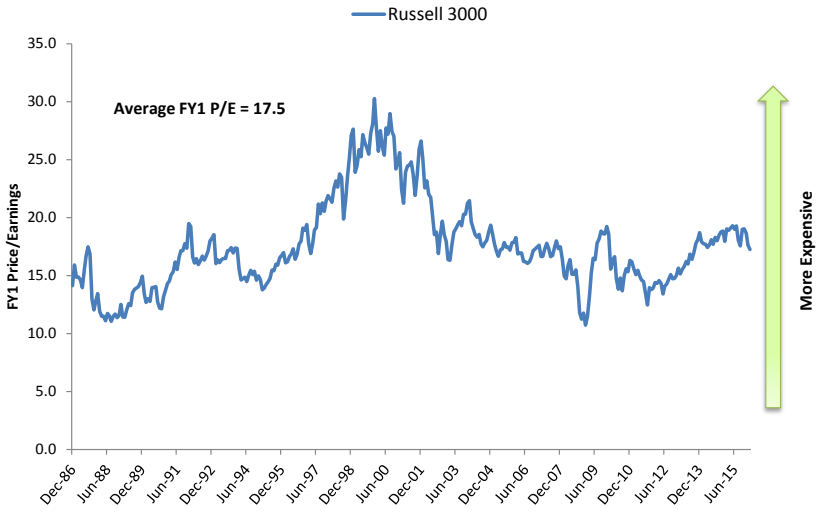
*One method to potentially gauge investor crowding or herding is by looking at the relative expensiveness of a particular market, sector, or strategy. If a strategy becomes overwhelmingly expensive, then this could be a potential crowding signal or turning point*

We can then analyze the correlation between the current P/E and future returns. If our indicator is negatively correlated to future performance, then this could suggest a valuation mean-reversal behavior. On the other hand, if it is positively correlated to future performance, then this would suggest a potential trending indicator. The strength of the correlation is also important. Note that correlation is only one approach we employ to understand the relationship between indicators and future returns.

We use the terms “mean-reversal” and “trending” loosely here. The term “mean-reversal” merely implies the likelihood of a future pullback, underperformance, or reversal in a strategy whereas “trending” implies potential future momentum or outperformance.

Figure 6 shows the Russell 3000 aggregate FY1 P/E ratio. The time-series commences from 1986. As P/E for the index increases, the broad equity market gets more expensive on a valuation basis. Note that the long-term average PE multiple during this entire period is 17.5x.

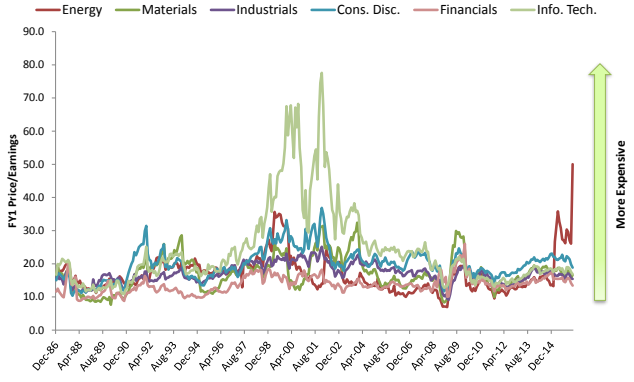
Figure 6: Russell 3000 aggregate FY1 price to earnings




*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

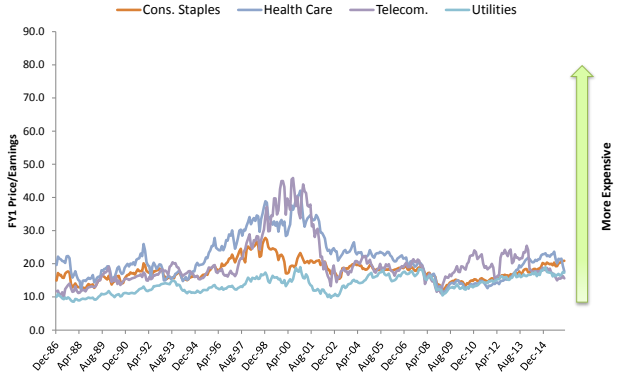
Figure 7 and Figure 8 show the PE at a sector level, delineated by cyclical versus defensive sectors. As expected the PE of cyclical sectors is more volatile than that of the defensive sectors.

Figure 7: FY1 PE for all cyclical sectors




*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

Figure 8: FY1 PE for all defensive sectors



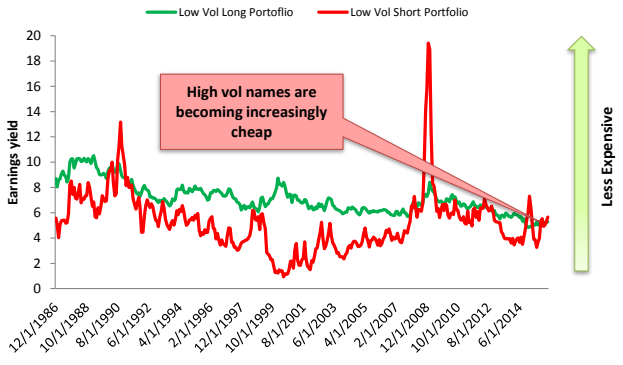

*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

It is obvious from Figure 6 to Figure 8, that valuation multiples do not strictly follow a mean-reversal pattern, e. g., the market can stay expensive (or cheap) for a long time. Therefore, there could be some mild crowding as sectors get expensive, which is not necessarily negative. We will revisit this notion later in the report.

*It is important to highlight that expensiveness is not a contemporaneous measure; meaning that sectors become expensive overtime*

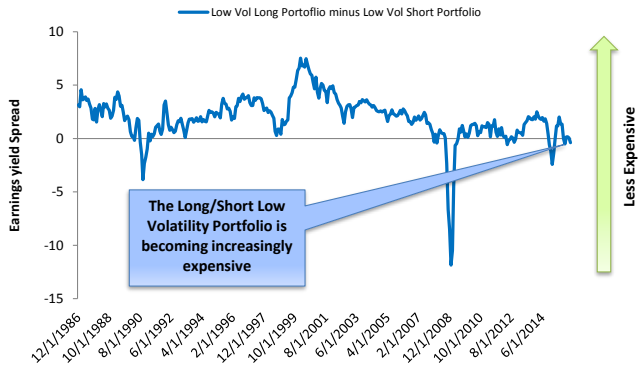
We also explore the expensiveness of a factor. Can expensiveness of a particular factor be a crowding indicator? To test this, we first compute the expensiveness of a broad set of quant factors. For example, take a long/short low volatility strategy. This is a portfolio that longs low volatility stocks and shorts high volatility stocks. We compute the average earnings yield (the inverse of P/E) for the long and short portfolios separately (see Figure 9). Then, we take the difference of the earnings yield for the long and short legs of the portfolio, which forms the valuation spread (see Figure 10).

Figure 9: Expensiveness of low volatility long and low volatility short




*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

Figure 10: Expensiveness of long/short low volatility portfolio

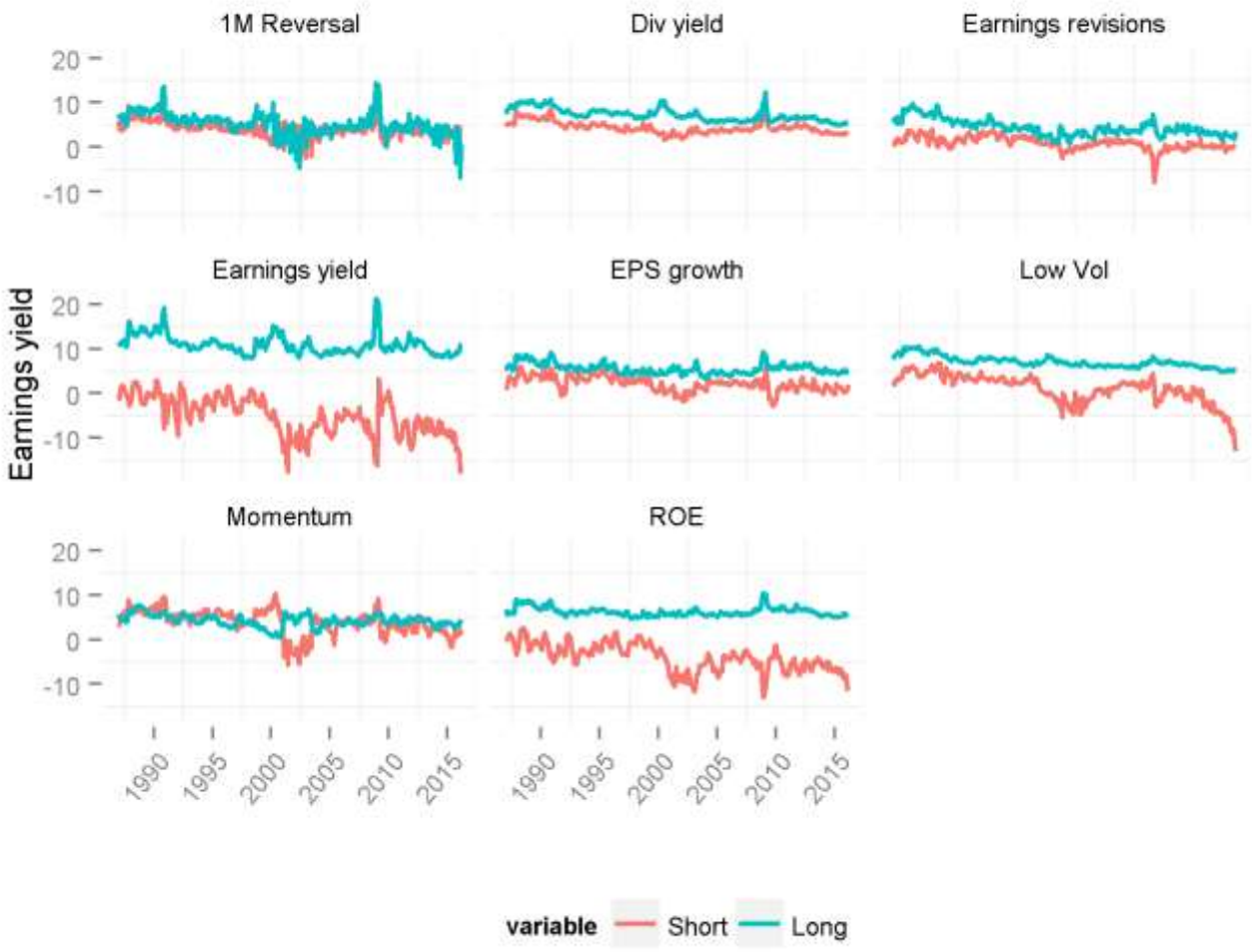



*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

If the earnings yield spread increases, this means that low volatility stocks have become cheaper (i.e., the long portfolio is becoming cheaper), and high volatility stocks are becoming expensive (i.e., the short portfolio is becoming expensive). Therefore, the net effect is that the low volatility strategy is becoming cheaper. Currently, we find that low volatility long/short portfolio has become more expensive relative to historical norms. The main driver is that high volatility stocks (that makeup the short leg) are becoming cheaper. We build upon this notion and compute the expensiveness of various quantitative strategies including value, growth, momentum, reversal, and sentiment (see Figure 11 and Figure 12). All and all, the expensiveness of a market, sector, or strategy could be a potential crowding indicator.

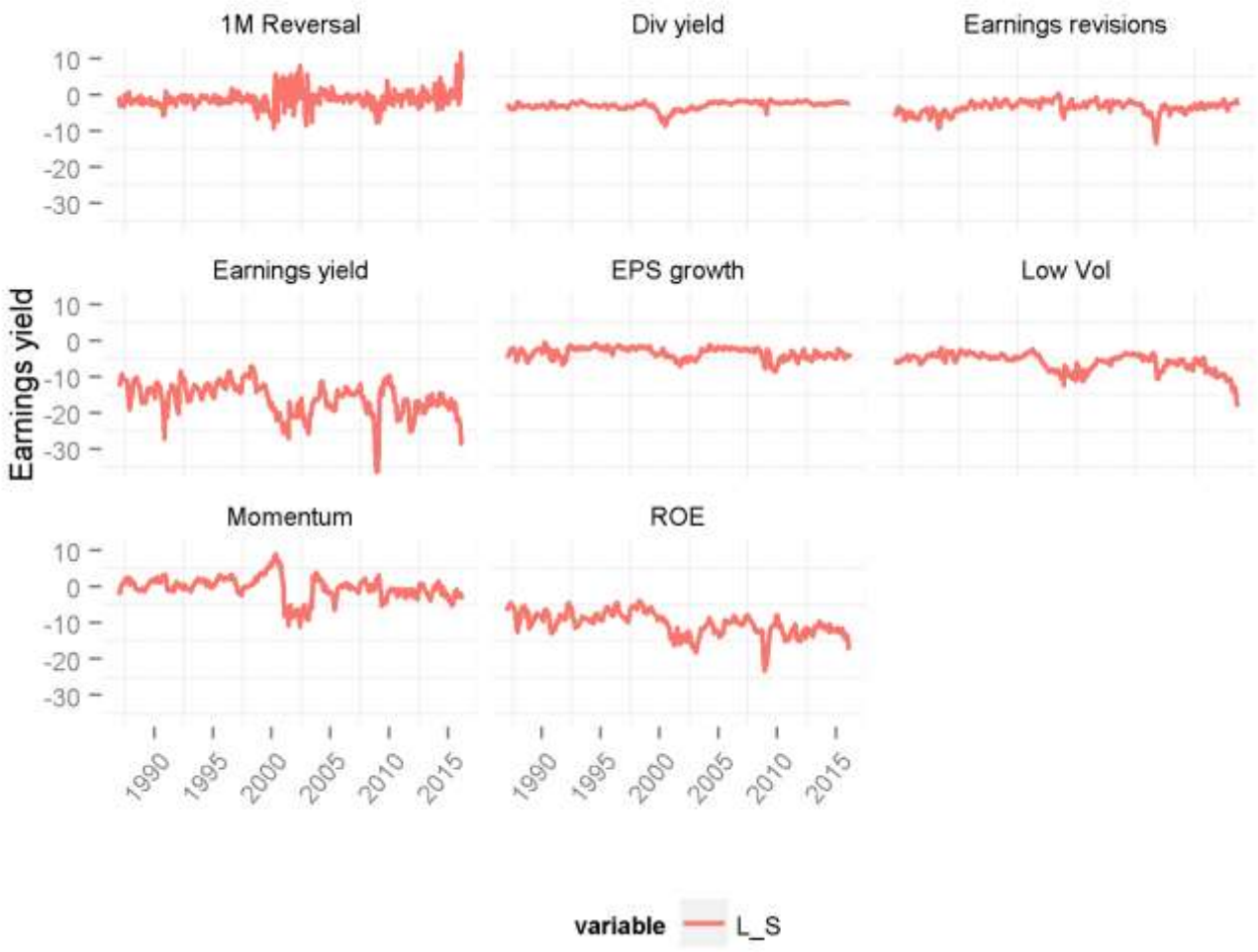
*Currently, we find that low volatility long/short portfolio is becoming expensive. The main driver is that high volatility stocks are becoming cheaper.*

Figure 11: Expensiveness of the separate long and short portfolios for various factors




*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

Figure 12: Expensiveness of common factors




*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

## 2. Historical performance

One often hears strategists announce that the performance of a particular market, strategy, or sector is overblown or has overshot. Watch out for a pullback or reversal. Indirectly, the idea is that the mere cumulative performance of a particular style or sector may be indicative of crowding. Can the historical performance of a factor, style or sector be a crowding indicator? Or, alternatively, is strong performance in a particular strategy indicative of future outperformance or trending (i.e., momentum or mean reversal)?

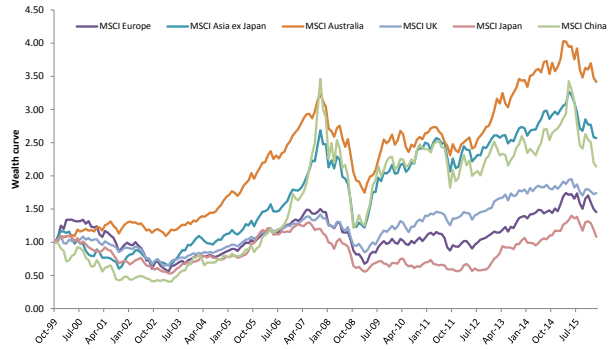
*One often hears strategists announce that the performance of a particular market, strategy, or sector is overblown or has overshot*

We start very simply, by observing the cumulative performance of various global indices. Figure 13 and Figure 14 show the wealth curve of North American and Global ex North American equity indices. At first glance, it seems that previous performance is more of a trending than crowding indicator. Simply analyzing the wealth curves show a steady trend in performance for most indices with only rare reversals or pull backs.

Figure 13: Wealth curve of North American stock market indices


*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

Figure 14: Wealth curve of global ex-NA stock market indices




*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

## 3. Market breadth

Market breadth is yet another indicator that is often widely discussed by investors. It is a sentiment indicator that may hint at the direction of the market. We test whether the sentiment conveyed by market breadth indicators is a crowding or trending measure?

*We test whether the sentiment conveyed by market breadth indicators is a crowding or trending measure or neither?*

Two widely used market breadth indicators are:[1]

1. Advance-Decline Ratio or AD ratio. The AD ratio is simply the number of stocks that went up minus # of stocks that were down over a given period. The metric is then divided by total number of stocks.

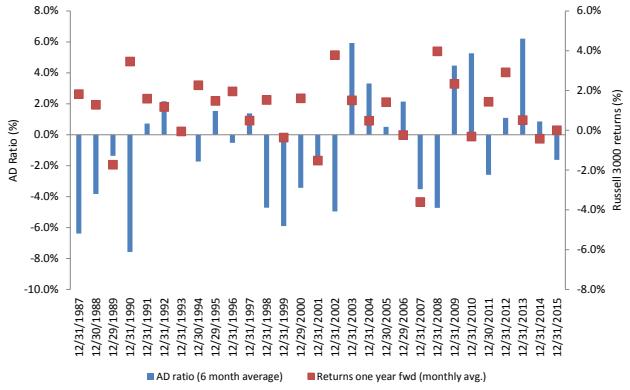
2. 52 Week High-Low Ratio or 52WHL ratio. The 52WHL ratio is the number of stocks near the 52 week high minus the number of stocks near the 52 week low.[2] The metric is divided by the total number of stocks.

To get a sense of these sentiment indicators, Figure 15 and Figure 16 show the AD and 52WHL ratio respectively, alongside the following 12-month Russell 3000 returns.

[1] See Rohal, et al [2016] for more details on market breadth indicators.

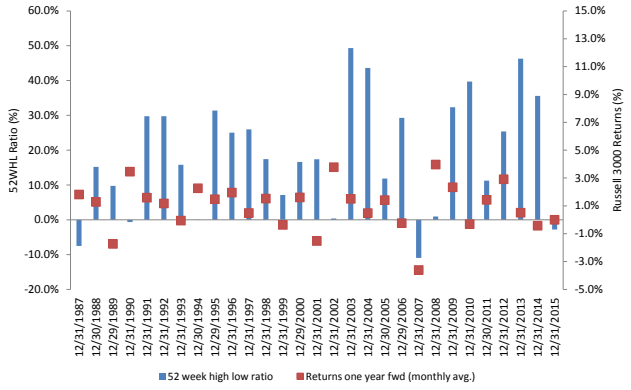
[2] Stocks near their 52-week high or low are those that are within 5% of their 52 week high or low.

Figure 15: AD ratio versus Russell 3000 forward returns




*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

Figure 16: 52WHL versus Russell 3000 forward returns




*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

We compute both sentiment metrics for the index, sectors, and style factors. To compute market breadth indicators for style factors, we simply compute each ratio separately for the long and short portfolios. Then, we subtract the market breadth indicator of the short portfolio from that of the long portfolio.

For example, for the long/short Low Volatility portfolio, we compute Low Vol AD Long minus Low Vol AD short. This way if the long portfolio indicates positive sentiment, while the short portfolio indicates negative sentiment, the long/short portfolio should show strong positive sentiment. We note that negative sentiment stocks are being shorted.

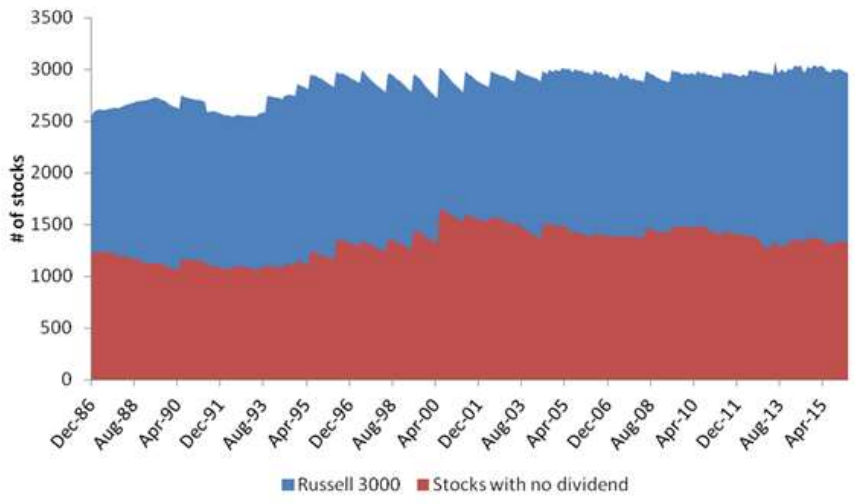
## Testing for crowdedness of the dividend/income strategy

*To test the efficacy of our crowding indicators, we analyze the relationship between our crowding indicators and future returns. If our indicator is negatively correlated to future performance, then this would suggest that it is a possible herding measure.*

To test the efficacy of all the aforementioned crowding indicators, we begin by simply analyzing the relationship between our crowding indicators and future returns for a particular strategy. We analyze the relationship between the dividend yield portfolio and market breadth. We chose to analyze the dividend yield portfolio because it is considered a defensive strategy, especially when risk aversion is high. As such, it may be more susceptible to a crowded trade now.

In the US market, many companies do not pay regular dividends (see Figure 17). Therefore, we form our dividend portfolio by buying the top 10% of companies with the highest dividend yield and shorting all companies that do not pay dividends.

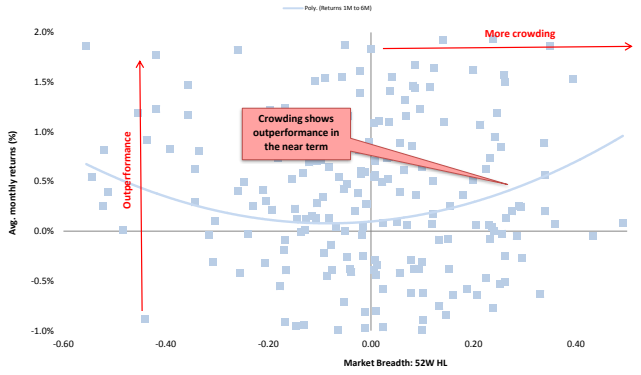
Figure 17: Coverage of companies not paying dividends in the Russell 3000




*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

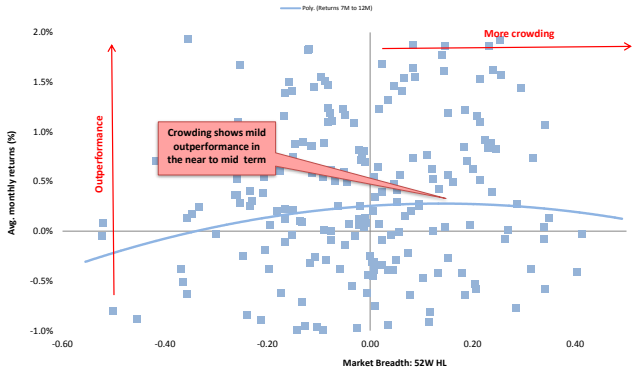
Figure 18 is a scatter plot between 52WHL of the dividend yield portfolio and the future one-to-six month average returns for the dividend portfolio.[3] To better understand the relationship between crowding and future returns, we overlay a polynomial fit on the scatter plot. Interestingly, we see a positive relationship between crowding and future one-to-six month return of the dividend portfolio. This reiterates our hypothesis that crowding can actually lead to stronger future performance.

Figure 18: 52WHL and the dividend portfolio future returns: 1-6 months




*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

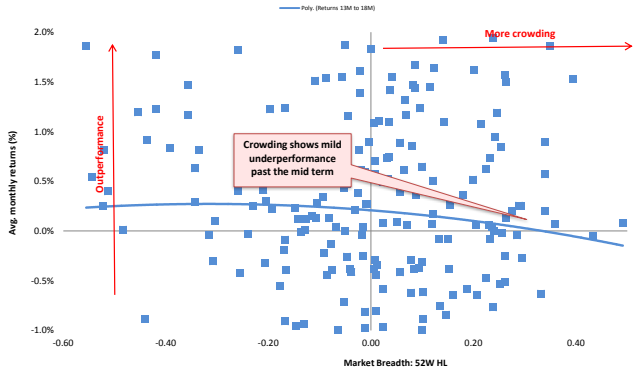
Figure 19: 52WHL and the dividend portfolio future returns: 7-12 months




*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

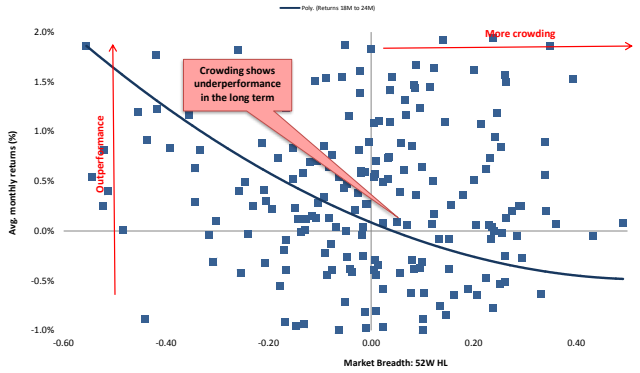
[3] The future returns are the average of the monthly returns using a range of one to six months.

Figure 20: 52WHL and the dividend portfolio future returns: 13 to 18 months




*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

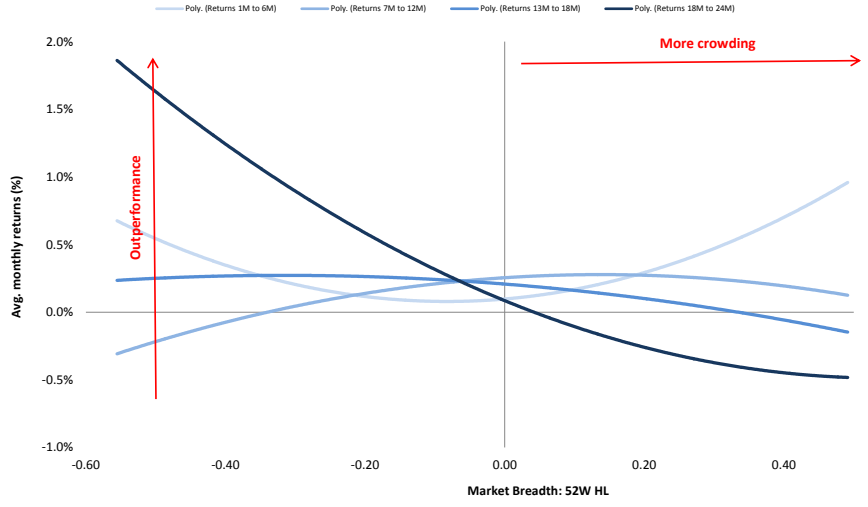
Figure 21: 52WHL and the dividend portfolio future returns: 19 to 24 months




*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

However, as we compare our crowding measure to returns further into the future, the positive association between crowding and future returns eases. For example, the relationship between market breadth and future 13-18 month returns (see Figure 20) and 19-24 month returns (see Figure 21) is clearly negative.

Figure 22: 52WHL and the dividend portfolio and all period future returns




*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

*In the long term, minor changes in crowding show underperformance. In fact, the elasticity of crowding is fairly strong. This means that small changes in expensiveness are associated with more significant levels of underperformance*

Combining all the charts from the various future return horizons (see Figure 22) yields some interesting findings:

- In the near to medium term, crowding in fact leads to stronger performance for the dividend yield portfolio.
- In the long term, the elasticity of crowding is fairly strong. This means that small changes in crowding are associated with more significant changes in performance.
- Crowding at extreme levels (the very right side of the chart) is associated with future drawdowns, irrespective of the periodicity of future returns.

By no means are these findings definitive, but they do provide a framework for analyzing and assessing potential crowding measures.

## Testing crowdedness for classic measures

*To get a better sense of how well all the aforementioned measures gauge investor crowding, we develop a crowding performance matrix*

To get a better sense of how well all the classic measures gauge investor crowding, we develop a crowding performance matrix. Essentially we want to test the relationship of our potential crowding measures with future performance. To do this we simply analyze the correlation between our crowding measures and future returns.

We run the crowding performance matrix for some common factors (see Figure 23). Generally speaking, crowding is positively correlated to near-term performance, but negatively associated with long-term returns.

Overall, the results for our classic measures are mixed yet promising. Albeit, the strength of the correlation measures are not that strong. Next, we reach into our quant toolbox and investigate more sophisticated measures of crowding.

Figure 23: Classic measures, crowding performance matrix for factors

| Factors | P/E | | | | Performance | | | | AD ratio | | | | 52 week HL | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Dividend yield | -43% | -34% | -27% | -14% | -13% | -16% | -7% | -14% | -1% | 0% | -6% | -18% | 2% | 5% | -4% | -24% |
| Earnings yield | -4% | -1% | 3% | 3% | 13% | -8% | -15% | -29% | 19% | -3% | -12% | -37% | 5% | -11% | -8% | -21% |
| Momentum | 4% | 18% | 25% | 30% | 17% | 9% | 0% | 11% | 33% | 7% | 4% | -9% | 26% | 8% | 4% | -7% |
| 1M Reversal | 2% | 1% | -12% | -4% | -13% | -10% | -4% | 7% | -23% | -28% | -23% | -25% | 5% | 1% | 5% | -4% |
| Growth | 17% | -1% | 3% | 4% | 16% | 5% | 5% | 1% | 30% | 4% | -2% | -12% | 23% | 7% | 9% | -9% |
| Quality | 19% | 14% | 12% | 16% | 0% | -17% | -13% | -24% | 26% | -7% | -5% | -21% | 14% | -8% | -6% | -14% |
| Low Volailtiy | -20% | -15% | -14% | 2% | -2% | -15% | -21% | -28% | 22% | 8% | -7% | -28% | 15% | 6% | -4% | -17% |
| Sentiment | 4% | -7% | -16% | -18% | 15% | 12% | 23% | 27% | 24% | 3% | 1% | 9% | 29% | 16% | 18% | 5% |
| Avg. Monthly Returns | 1 to 6 | 7 to 12 | 13 to 18 | 19 to 24 | 1 to 6 | 7 to 12 | 13 to 18 | 19 to 24 | 1 to 6 | 7 to 12 | 13 to 18 | 19 to 24 | 1 to 6 | 7 to 12 | 13 to 18 | 19 to 24 |

*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

# Novel crowding measures

## 4. Short interest

In our previous research, we have suggested other crowding measures (see Cahan, Luo, Alvarez, Jussa [2012a]). In this section, we update and extend this research.

Our methodology is based closely on the one proposed in Hanson and Sunderam [2011]. Suppose we have a proxy for investor appetite for a particular subset of stocks – like the short interest used in Hanson and Sunderam. Then at each point in time $t$, we regress that proxy cross-sectionally onto dummy variables that denote whether a stock falls into each quantile of a set of $j$ quant factors (e.g., value, momentum, and so on). We also include dummy variables for size and volatility as controls. More specifically, we have:

$$C_{i,t} = c + \sum_{j=1}^{J} \sum_{q=2}^{Q} \beta_{i,t,j,q} D_{i,t,j,q} + \sum_{q=2}^{Q} \beta_{i,t,size,q} D_{i,t,size,q} + \sum_{q=2}^{Q} \beta_{i,t,\sigma,q} D_{i,t,\sigma,q} + \varepsilon_{i,t}$$

*The advantage of their approach is that it is cross-sectional and therefore is less backwards looking. It does not rely on trailing regressions to compute style sensitivities*

where $C_{i,t}$ is the crowdedness score (e. g., short interest) for stock $i$ at time $t$ for a crowdedness proxy variable, $D_{i,t,j,q}$ is a dummy variable that indicates if stock $i$ is in the $q$-th quantile of factor $j$ at time $t$, and $D_{i,t,size,q}$ and $D_{i,t,\sigma,q}$ are dummy variables denoting if stock $i$ is in the $q$-th quantile of size and volatility at time $t$. In our analysis we set $Q$ =10, i.e. deciles, and omit the lowest decile from the dummy variables. We order our variables such that the least attractive decile is denoted $q$ =10. Hence by tracking the coefficient $\beta_{i,t,j,10}$ over time, we can measure the amount by which stocks that are unattractive, as measured by factor $j$, have a different score on the proxy variable compared to attractive stocks, after controlling for size, volatility, and potentially other quant factors.

For example, consider the value factor. If expensive (i.e. unattractive) stocks are more heavily shorted than cheap (i.e. attractive) stocks at a given point in time, then this might indicate that institutional investors (who are more likely to short) are heavily invested in value.

### Problems with using traditional short interest data

When looking for ways to measure factor crowding, the securities lending market seems a natural place to start. The idea is simple: if investors are, in aggregate, heavily shorting stocks that look unattractive on a particular factor, then that would suggest that more capital is chasing that particular strategy.

However, there are some limitations. First, we traditionally source short interest data from Compustat, which is available at a monthly frequency and is snapped on the 15th of each month. This means there can be up to a two week lag between when the data was captured and when it is available to the market.

Second, when we use short interest (i.e., the percent of total shares that are currently sold short) as the measure of shorting demand, we ignore the supply side. The problem with short interest is that it does not capture the supply side of the securities lending equation. For example, if two stocks both have 2%

short interest (based on total numbers of shares outstanding), but one stock has a lendable supply of 4% and the other has a supply of 10%, then clearly the first stock is much more heavily shorted relative to its lendable supply, even though both stocks would have the same ranking on the traditional short interest metric.

### Advantages of the Markit securities finance (MSF) dataset

To address these drawbacks, we leverage a unique database of securities lending activity from Markit (formerly called DataExplorers). We discussed this data set in considerable details, in our previous research, (see Cahan, et al [2012]). There are several advantages to using the MSF dataset. First, the MSF database is updated daily and is available on a T+2 basis. This negates the lag problem that comes with other datasets,

Second, the MSF database captures both the supply and demand sides of the securities lending market. In particular, a metric called Utilization is useful. It measures the shorting activity in a stock as a percentage of the pool of lendable supply in that stock, rather than as a percentage of the total shares on issue. This gives a more accurate depiction of the true level of shorting in a stock.

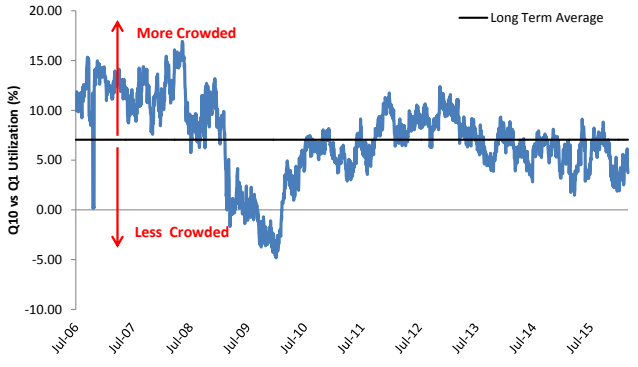
*Using the daily Utilization metric from Markit, we conduct the cross-sectional regression of the above equation on a daily basis from 2006 (when the DataExplorer database starts) to present*

### Factor crowding using utilization

Using the daily Utilization metric computed from the MSF database, we conduct the above cross-sectional regression on a daily basis from 2006 (when the MSF database starts). Figure 24 shows the coefficients for decile 10 stocks for the Value factor. Again, this chart is, loosely speaking, measuring the difference in Utilization for Q10 stocks (i.e. unattractive) stocks versus Q1 (i.e., attractive) stocks.

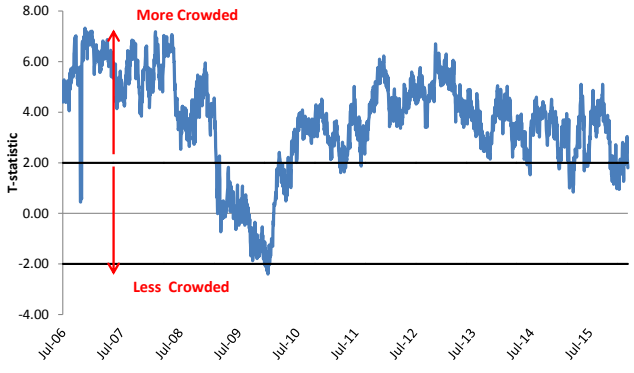
For example, the long-term average of the Value factor is around 7%. This indicates that over the long-run, expensive stocks tend to have Utilization that is 7% higher than cheap stocks, after controlling for differences in size and volatility. Obviously, expensive stocks tend to be more heavily shorted than cheap stocks. Furthermore, as shown in Figure 25, these differences tend to be statistically significant most of the time (roughly a t-statistic greater than 2.00 or less than -2.00).

Figure 24: Incremental Utilization for Q10 vs. Q1 – value




*Source: Bloomberg Finance LP, Compustat, Markit, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

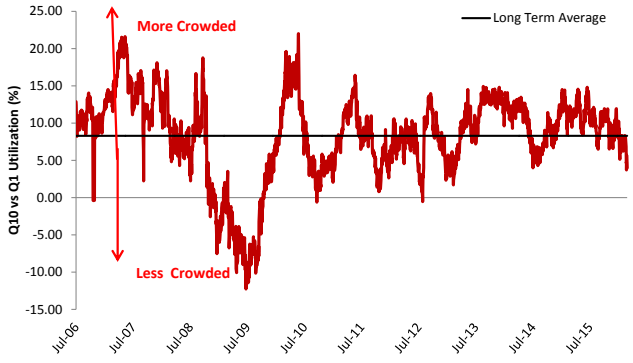
Figure 25: T-statistic for regression coefficients – value




*Source: Bloomberg Finance LP, Compustat, Markit, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

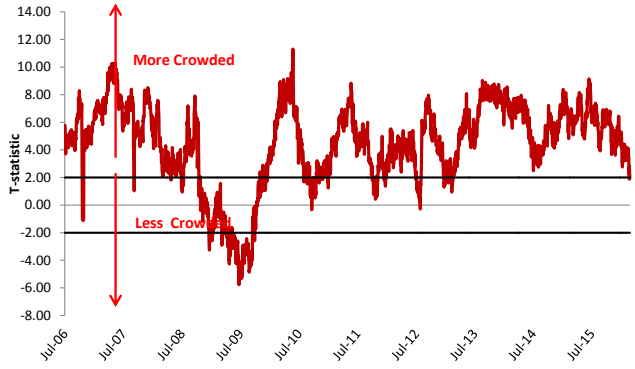
The same applies to Momentum – on average, past losers have higher Utilization than past-winners (see Figure 26 and Figure 27). Neither momentum or value appears to be currently crowded.

Figure 26: Incremental Utilization for Q10 vs. Q1 – momentum




*Source: Bloomberg Finance LP, Compustat, Markit, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

Figure 27: T-statistic for regression coefficients – momentum



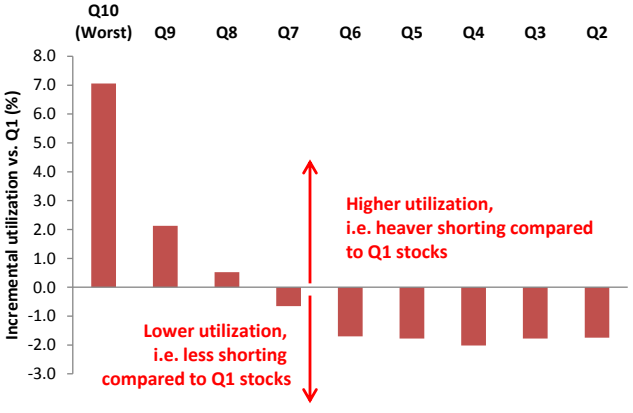

*Source: Bloomberg Finance LP, Compustat, Markit, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

## Robustness checks

So far we have been looking at the difference in Utilization between the best and worse stocks on a particular factor. However, if the intensity of shorting activity is a proxy for crowdedness, then we should observe a monotonic relationship between factor scores and Utilization.

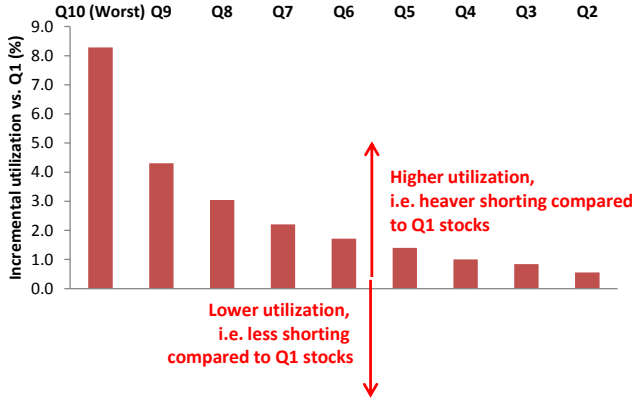
This is indeed the case. Figure 28 and Figure 29 show the difference in Utilization between decile one (most attractive) stocks and the other deciles, for Value and Momentum, respectively and measured over the whole history from 2006. In both cases, the results are quite monotonic; shorting is heaviest for expensive stocks and past-loser stocks, and then declines gradually as one moves towards more attractive stocks.

Figure 28: Incremental Utilization vs. Q1 – value




*Source: Bloomberg Finance LP, Compustat, Markit, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

Figure 29: Incremental Utilization vs. Q1 – momentum

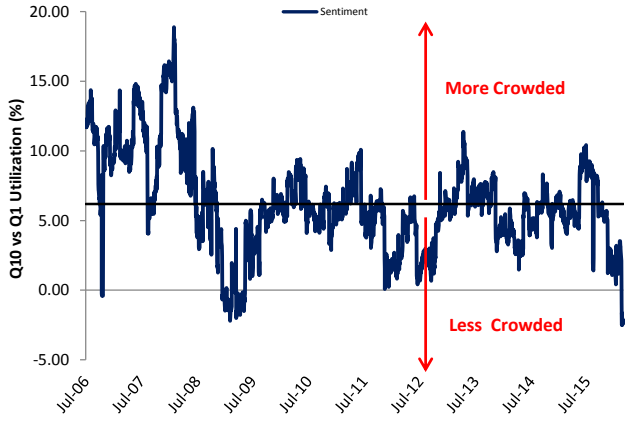



*Source: Bloomberg Finance LP, Compustat, Markit, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

Analyzing the other quant factors in isolation, Figure 30 to Figure 33 show the coefficients for common quant factors including: sentiment, quality, growth,

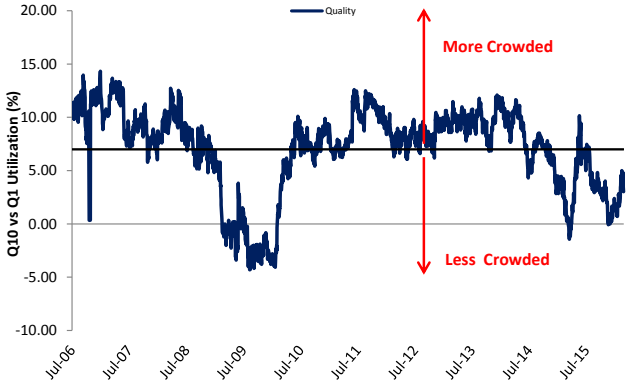
and reversal. Currently, other than reversal and low vol, most of these factors do not appear to be crowded, based on the utilization concept.

Figure 30: Incremental Utilization for Q10 vs. Q1 – sentiment



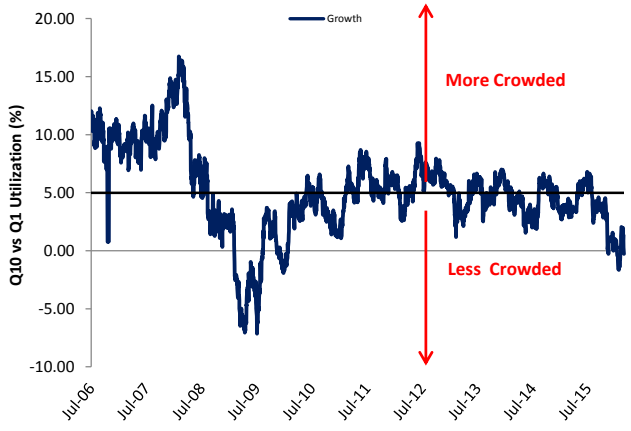

*Source: Bloomberg Finance LP, Compustat, Markit, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

Figure 31: Incremental Utilization for Q10 vs. Q1 – quality



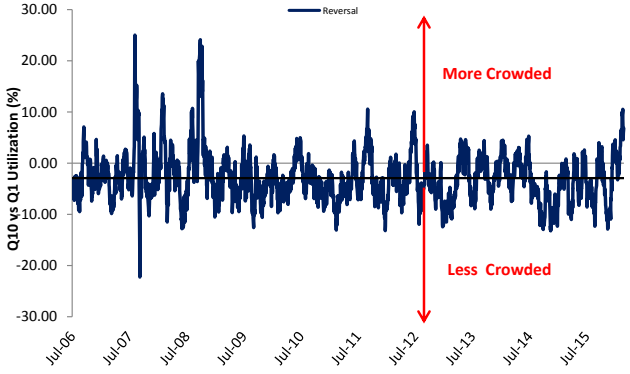

*Source: Bloomberg Finance LP, Compustat, Markit, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

Figure 32: Incremental Utilization for Q10 vs. Q1 – growth



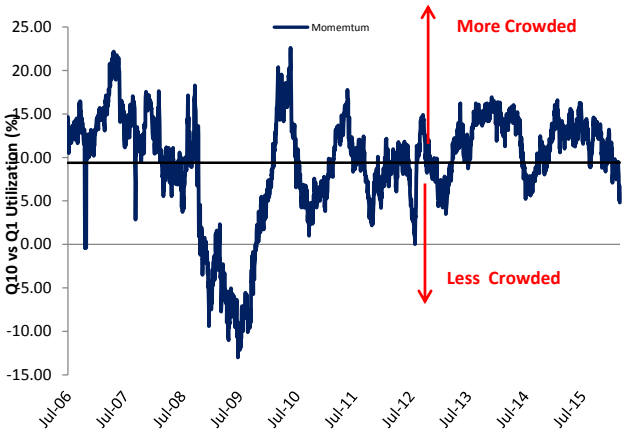

*Source: Bloomberg Finance LP, Compustat, Markit, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

Figure 33: Incremental Utilization for Q10 vs. Q1 – reversal



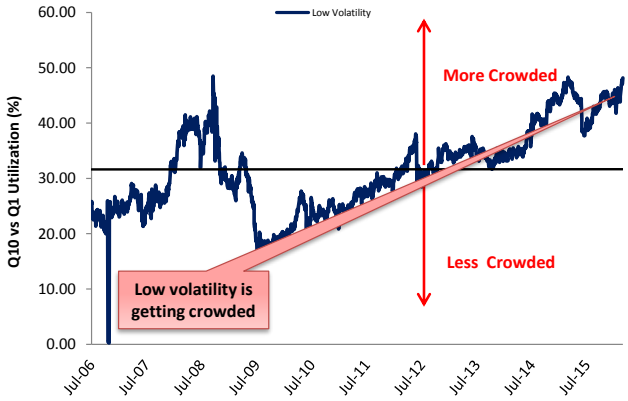

*Source: Bloomberg Finance LP, Compustat, Markit, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

Figure 34: Incremental Utilization for Q10 vs. Q1 – momentum




*Source: Bloomberg Finance LP, Compustat, Markit, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

Figure 35: Incremental Utilization for Q10 vs. Q1 – low volatility




*Source: Bloomberg Finance LP, Compustat, Markit, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

We highlight low volatility separately because it is a special case. Volatility is a dummy variable in our regression.[4] We did not add it in as additional quant factor. Low volatility appears to be expensive. However, we must also keep in mind that on average, high volatility stocks are expensive to short. As such, low volatility should be scrutinized against its own history. Based on this crowding measure, low volatility appears to be as crowded as it was during the financial crisis when investors simultaneously plunged into defensive, low volatility stocks. We will discuss more appropriate measures of crowding for low volatility strategies in the next few sections.

## Are quant strategies crowded on average?

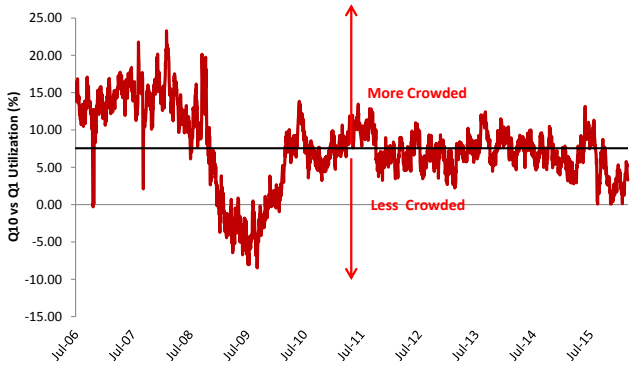
So far, we have concentrated on quant factors in isolation. Putting everything together, we can measure the crowdedness of quant strategies overall by analyzing a simple multifactor model with six standard quant factors. These factors are: value, growth, momentum, sentiment, quality, and reversal. Note we excluded low volatility because it is a control variable in the regression. We z-score each factor and form an equally-weighted alpha signal. Then we rerun our regression analysis, using one multifactor alpha signal instead of individual factors.

*Putting everything together, we can measure the crowdedness of quant strategies overall by constructing a simple multifactor model with six common quant factors*

Figure 36 shows the incremental Utilization for unattractive stocks (Q10) compared to attractive stocks (Q1), and Figure 37 shows the corresponding t-statistic. If we believe our six-factor model is somewhat representative of a typical quant strategy, then these charts should give us a good feel for how crowded the overall quant space is, at least in a relative sense (i.e. compared to history). The results do tend to support the commonly held thesis that quant crowding was high in the pre-crisis period, fell sharply through the worst of the financial crisis, and has now started to increase again albeit at lower levels than before. However, currently we observe no significant crowding within the quant space, based on our incremental utilization coefficient.

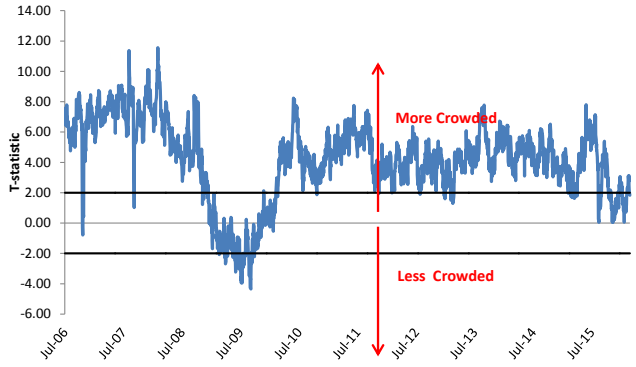
[4] We inverted the coefficient to make it comparable to the other quant factors

Figure 36: Incremental Utilization for Q10 vs. Q1 – multifactor




*Source: Bloomberg Finance LP, Compustat, Markit, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

Figure 37: T-statistic for regression coefficients – multifactor




*Source: Bloomberg Finance LP, Compustat, Markit, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

As a result of this evidence, we would suggest that concerns about crowding in the overall quant space are less than what they might have been pre financial crisis. At the individual factor level, there are some indications for concern. For example, the low volatility factor in particular is showing a level of crowdedness that is on par with pre financial crisis levels. In conclusion, we believe that the use of real-time securities lending data coupled with various crowding metrics are an effective way to monitor these trends.

## Crowding and performance – classification analysis

There are several analytical methods that we can employ to grasp the relationship between crowding and future performance. One method is to use a classification and regression tree or CART model.[5] The input to this model (or the independent variable) is our crowding indicator based on utilization.[6] The output to the model (or the dependant variable) is future returns.

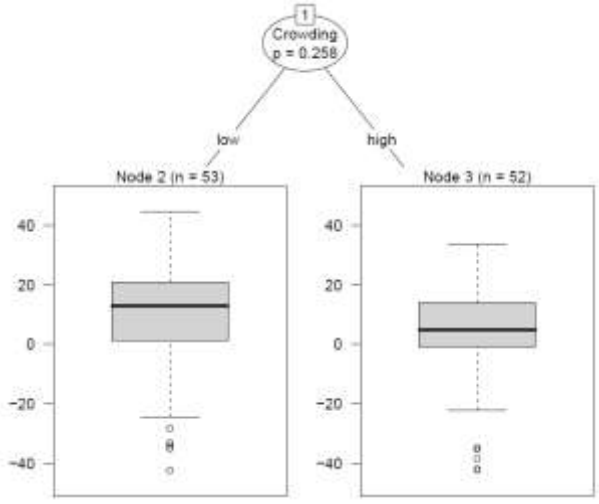
Figure 38 shows the results of the CART model on the value portfolio using our utilization measure as a crowding indicator. Note that we used 12-month future returns for the model. The results show that low levels of crowding are associated with higher 12-month future returns, in contrast to more crowded periods. Interestingly, higher levels of crowding are also associated with positive 12-month returns but with lower average returns. The p-value is also insignificant.

If we re-run our CART model using 13 month to 24-month future returns, we get even more interesting results, (see Figure 39) shows that low levels of crowding are still associated with modest positive returns. However, higher levels of crowding are associated with significantly negative and more volatile returns (with wider range of return distribution).

[5] We have used the CART model in many of our previous research, e. g. Luo, et al [2010], Wang, et al [2014].

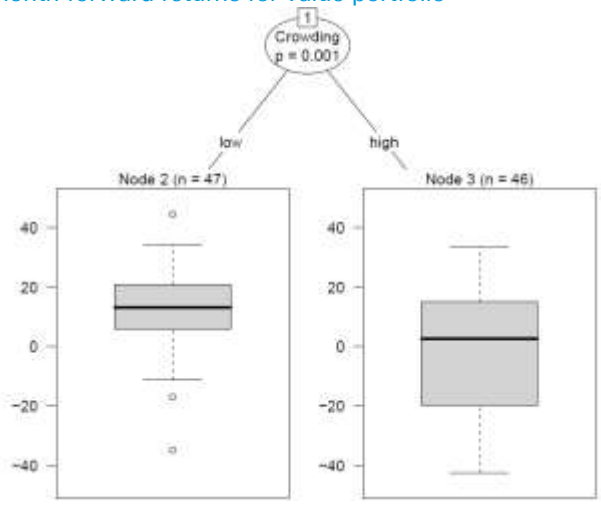
[6] To make the analysis more intuitive, we separate the crowding input into two measures, high crowding and low crowding. High crowding are values above the median and low crowding are values below the median. Note that the median is calculated over the entire history.

Figure 38: CART: incremental utilization and 12-month forward returns for value portfolio




*Source: Bloomberg Finance LP, Compustat, Markit, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

Figure 39: CART: incremental utilization and 13-24-month forward returns for value portfolio




*Source: Bloomberg Finance LP, Compustat, Markit, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

The results from our CART analysis are fairly interesting and intuitive. To test the robustness of the results, we can employ a simple multivariate regression framework to test for crowding and future performance.

## Crowding and performance – Multivariate regression analysis

We regress future factor returns against four explanatory variables: crowding, crowding squared (to capture extreme levels of crowding), change in crowding (to capture sudden shifts in crowding), crowding saturation (to capture level of crowding with respect to recent high). The regression takes the following form:

$$r_{t+n} = \beta_0 + \beta_1 crowding_t + \beta_2 (crowding)_t^{\ 2} + \beta_3 \Delta crowding_t + \beta_4 crowding\ saturation_t + \varepsilon_t$$

Where crowding saturation is defined as:

$$crowding\ saturation_t = Max\left(crowding_t, crowding_{t-1}, \ldots crowding_{t-36}\right) - crowding_t$$

We run two regressions using 12- and 24-month future returns. Our crowding measure is incremental utilization. Figure 40 below shows the results of the regression on the value portfolio. All the coefficients are statistically significant. The crowding and crowding squared measure both have negative coefficients, indicating that negative future returns can be explained by crowding and extreme levels of crowding.

Figure 40: Multivariate regression using incremental utilization and forward returns for value portfolio

| | 12 Month Forward Returns | | 24 Month Forward Returns | |
|---|---|---|---|---|
| | Beta | T-Stat | Beta | T-Stat |
| Crowding | (4.77) | (3.76) | (13.42) | (8.95) |
| Crowding squared | (2.98) | (2.02) | (5.99) | (3.61) |
| Crowding change | 1.62 | 3.14 | 3.76 | 6.94 |
| Crowding saturation | (3.40) | (3.05) | (7.27) | (5.7) |

*Source: Bloomberg Finance LP, Compustat, Markit, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

We ran the multivariate regressions for all the common factors. Figure 41 shows the average coefficients and test statistics. Again, for 24-month future returns, all the coefficients are statistically significant. The crowding and crowding squared measure both have negative coefficients.

Figure 41: Average coefficients for multivariate regression using incremental utilization

| | 12-Month Forward Returns | | 24-Month Forward Returns | |
|---|---|---|---|---|
| | Average Beta | Average T-stat | Average Beta | Average T-stat |
| Crowding | (3.71) | (3.04) | (8.77) | (5.40) |
| Crowding squared | 1.06 | 0.12 | (4.33) | (2.32) |
| Crowding change | 1.59 | 3.41 | 2.67 | 4.83 |
| Crowding saturation | (3.39) | (3.78) | (5.71) | (3.90) |

*Source: Bloomberg Finance LP, Compustat, Markit, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

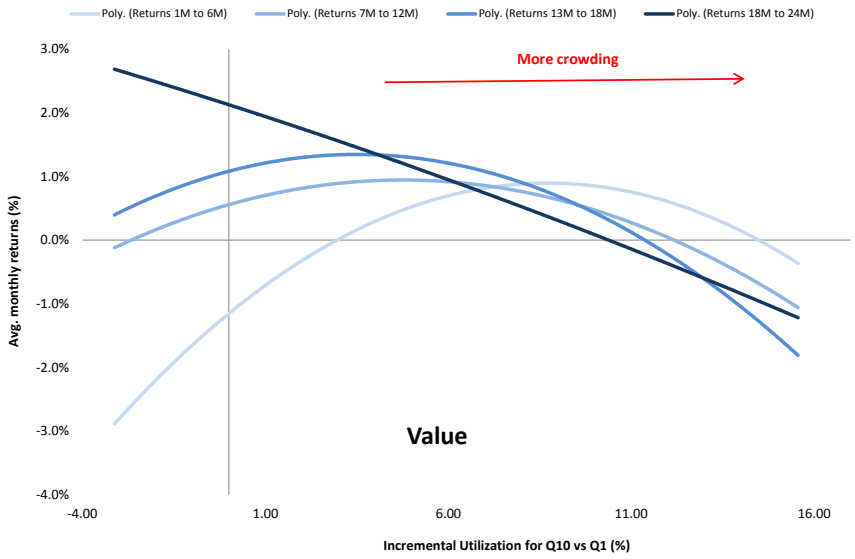
## Crowding and performance – Scatter plot analysis

Next we investigate the relationship between these factor crowding measures and future performance using scatter plot analysis. Figure 42 is a scatter plot between incremental utilization for the value factor and future monthly returns to the value portfolio (1 to 6 months, 7 to 12 months, 13 to 18 months, and 18 to 24 months). To gain more clarity on the relationship between incremental utilization and forward returns, we fit a polynomial function through each series of future returns. Analyzing the chart yields some insightful findings:

*In the near to medium term, mild crowding in fact adds to the performance of the value strategy. We see a positive relationship between incremental utilization and future performance*

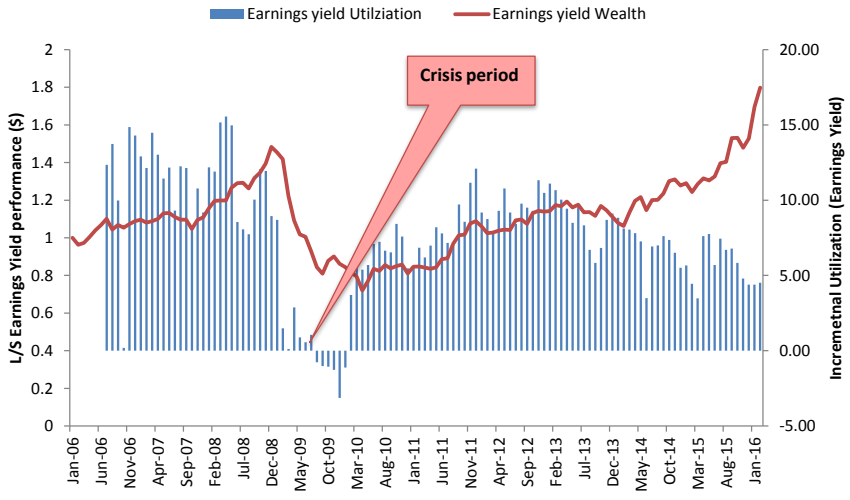
- In the near to medium term, mild crowding in fact adds to the performance of the value strategy. We see a positive relationship between incremental utilization and future performance.
- In the long term, minor changes in crowding show underperformance. In fact, the elasticity of crowding is fairly strong. This means that small changes in incremental utilization are associated within significant levels of underperformance.
- Crowding at extreme levels (the very right side of the chart) is associated with future underperformance irrespective of the periodicity of future returns. At extreme levels of crowding, we see underperformance in the near and long term.
- Extreme levels of crowding (or negative utilization) are typically associated with crisis periods. In a typical market regime, value investors would typically prefer to buy cheap stocks and short or underweight expensive stocks. During the onset of the financial crisis, investors likely bought safe haven, expensive stocks and shorted cheap stocks. This would cause a reduction in incremental utilization or it could even turn negative (see Figure 43).

Figure 42: Incremental utilization and future performance for value




*Source: Bloomberg Finance LP, Compustat, Markit, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

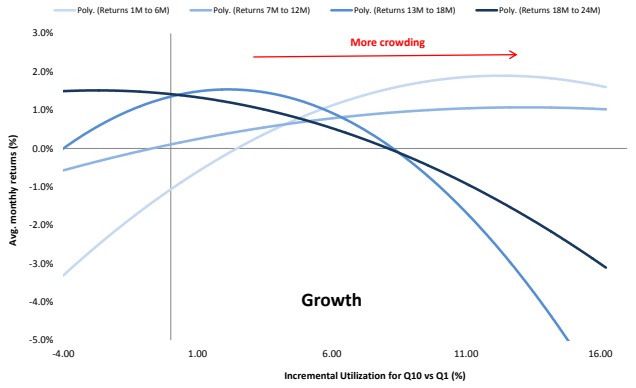
Figure 43: Time series of incremental utilization for value portfolio




*Source: Bloomberg Finance LP, Compustat, Markit, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

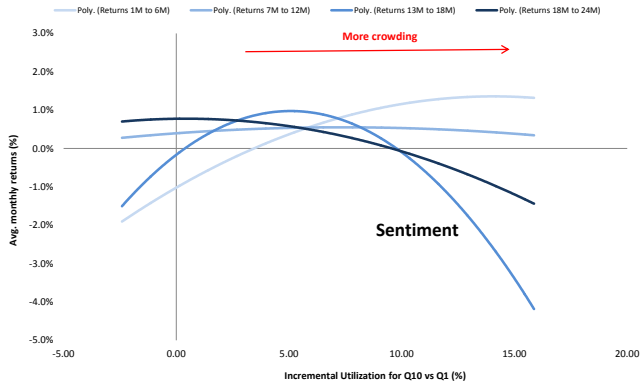
We ran a similar analysis for other standard quant portfolios (see Figure 44 to Figure 47). Most factors showed a similar polynomial pattern.

Figure 44: Utilization and performance – growth



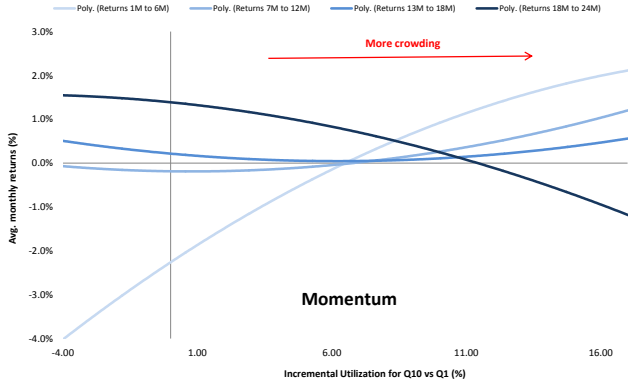

*Source: Bloomberg Finance LP, Compustat, Markit, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

Figure 45: Utilization and performance – sentiment



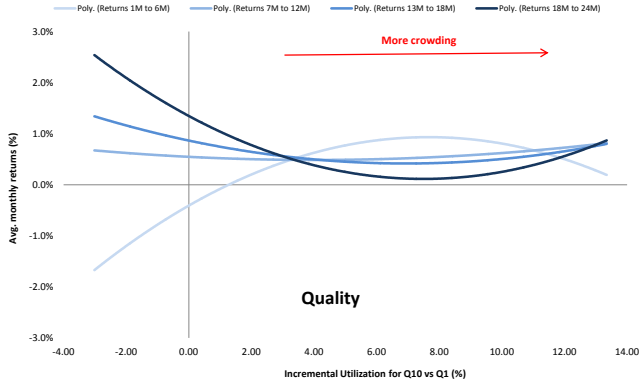

*Source: Bloomberg Finance LP, Compustat, Markit, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

Figure 46: Utilization and performance – momentum




*Source: Bloomberg Finance LP, Compustat, Markit, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

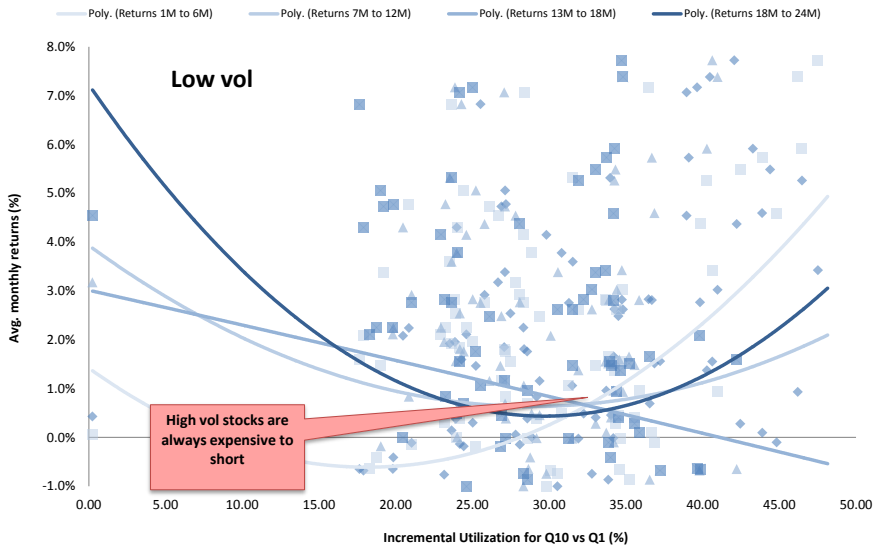
Figure 47: Utilization and performance – quality




*Source: Bloomberg Finance LP, Compustat, Markit, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

The one exception is low volatility (see Figure 48), the dummy variable in the regression. Highly volatile stocks tend to more heavily shorted than low volatility stocks. Additionally, high volatility stocks are also more expensive to short than low volatility stocks. The cluster of incremental utilization for low volatility is higher than any other factor.

Figure 48: Incremental utilization and future performance for low volatility



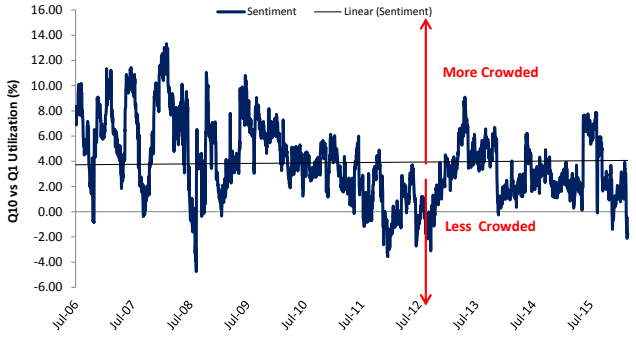

Source: Bloomberg Finance LP, Compustat, Markit, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank

*One of the drawbacks of using a single multifactor z-score within the regression is that there is no control for other quant factors*

## What's crowded after controlling for all quant factors?

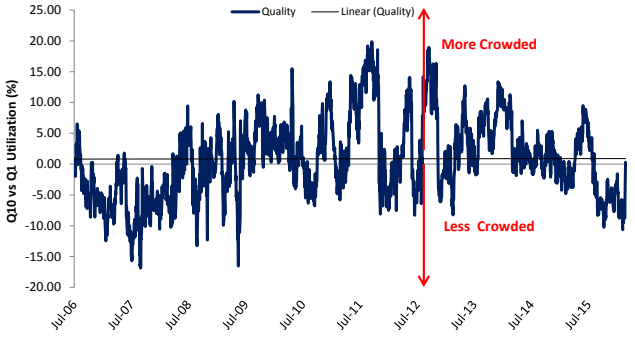
One of the drawbacks of using a single multifactor z-score within the regression is that there is no control for other quant factors. Therefore, we repeat our regressions with the factors as individual independent variables (plus Size and Volatility as controls). Figure 49 to Figure 54 show the coefficients of the regression. As we are using a cross-sectional regression framework, we need to keep in mind that there may be some correlation between the independent variables for which we have not accounted. However, the coefficients should give us a sense of which factors are crowded after controlling for other strategies. Other than low vol and short-term reversal factors, the other common systematic strategies do not appear to be crowded.

Figure 49: Incremental Utilization for Q10 vs. Q1 – sentiment



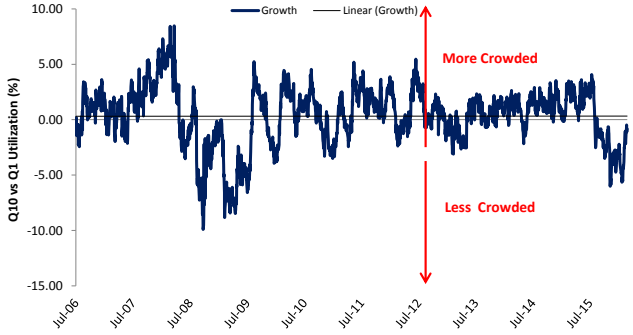

Source: Bloomberg Finance LP, Compustat, Markit, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank

Figure 50: Incremental Utilization for Q10 vs. Q1 – quality



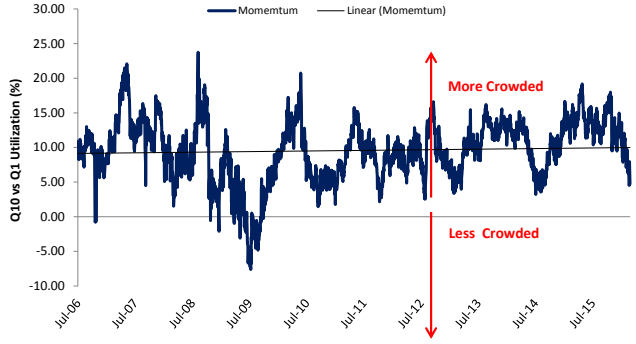

Source: Bloomberg Finance LP, Compustat, Markit, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank

Figure 51: Incremental Utilization for Q10 vs. Q1 – growth



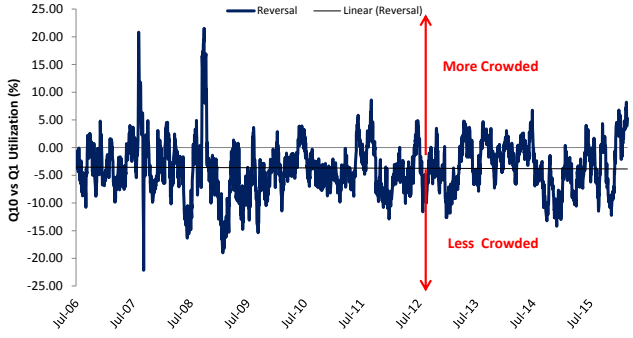

*Source: Bloomberg Finance LP, Compustat, Markit, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

Figure 52: Incremental Utilization for Q10 vs. Q1 – momentum



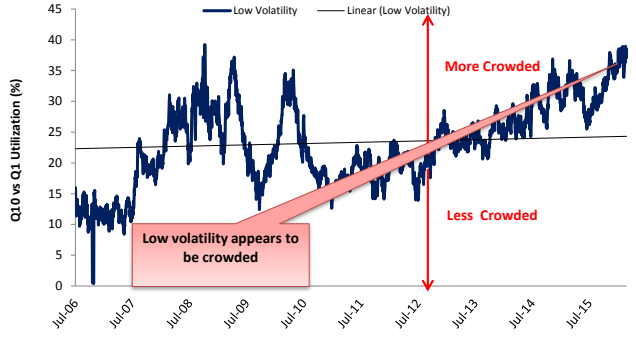

*Source: Bloomberg Finance LP, Compustat, Markit, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

Figure 53: Incremental Utilization for Q10 vs. Q1 – reversal




*Source: Bloomberg Finance LP, Compustat, Markit, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

Figure 54: Incremental Utilization for Q10 vs. Q1 – low volatility




*Source: Bloomberg Finance LP, Compustat, Markit, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

## 5. Mean pairwise correlation

Yet another potential crowding measure could be the pairwise correlation among stocks. In one of our previous research, we quantified crowdedness in low risk strategies using pairwise correlation and tail dependence (see Cahan, Alvarez, Luo, et al [2012b]). Recall that stock pairwise correlation is calculated by taking every possible pair of stocks and computing the correlation of their returns. Taking the mean value across all the pairs results in the mean pairwise correlation or MPC. Intuitively, this captures the tendency of stocks to move together or herd. If the MPC for a particular strategy, market, or sector is much higher than the market as a whole, then that could indicate that investors are trading these stocks as a group instead of individually. Additionally, if pairwise correlation had increased over time, that could indicate rising crowdedness in a strategy.

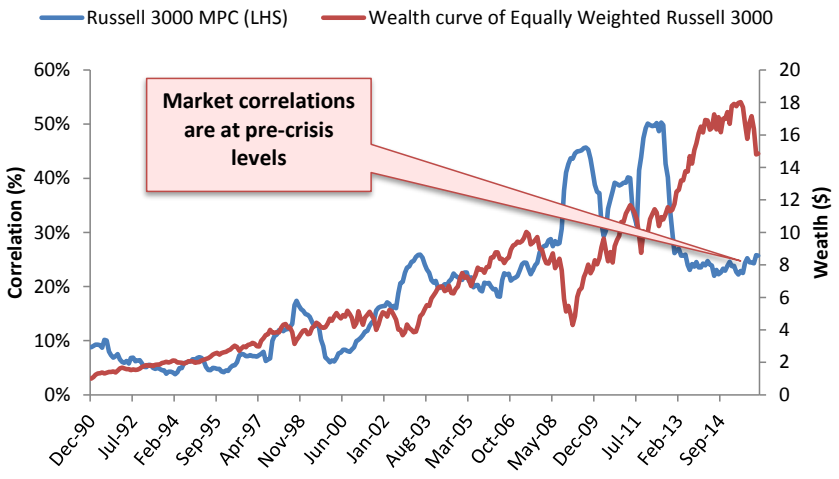
*If the MPC for a particular strategy, market, or sector is much higher than the market as a whole, then that could indicate that investors are trading these stocks as a group instead of individually*

To start, we plot the market MPC alongside the Russell 3000 index (see Figure 55). We note a significant uptick in MPC during the financial crisis as well as the European debt crisis. Currently, correlation levels have dropped to pre-crisis levels. Broadly speaking, Figure 55 shows that market stress episodes tend to be accompanied by sharp increases in MPC. This is consistent with the notion that during times of distress, investors sell first and ask questions later. This results in stocks being traded more as a basket rather than based on their individual characteristics.

*Currently, correlation levels have dropped to pre-crisis levels. Broadly speaking, the chart shows that market selloff episodes are accompanied by sharp increases in MPC*

Additionally, market-wide MPC was fairly low in the early 1990s when factor performance was strong. In recent years, MPC has been elevated, likely because of the proliferation of systematic and passive strategies while factor performance has been challenging.

Figure 55: MPC of Russell 3000




*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

To compute MPC for our factor portfolios, we compute the MPC of the long leg of the factor portfolio and subtract the MPC of the market.[7] As such, if the MPC of a factor portfolio is much higher than the market, this could be an indication of investor crowding. A value greater than zero would indicate that the factor portfolio is more correlated than the market.

To begin, we analyze the time-series MPC of the value portfolio (see Figure 56). We find that the value strategy peaked in mid-2007, which coincides with the 2007-quant crisis. This coincides with conventional wisdom that investors were chasing inexpensive, risky stocks during the bull market. However, during the onset of the financial crisis; investors steered clear of these risky stocks and hence the crowdedness of value drops significantly. This scenario aligns well with our utilization measure of crowding (see Figure 57). It showed that during the financial crisis, inexpensive stocks became cheaper to short than expensive stocks. This is because during the financial crisis investors preferred more stable, companies which tend to be more expensive. Currently,

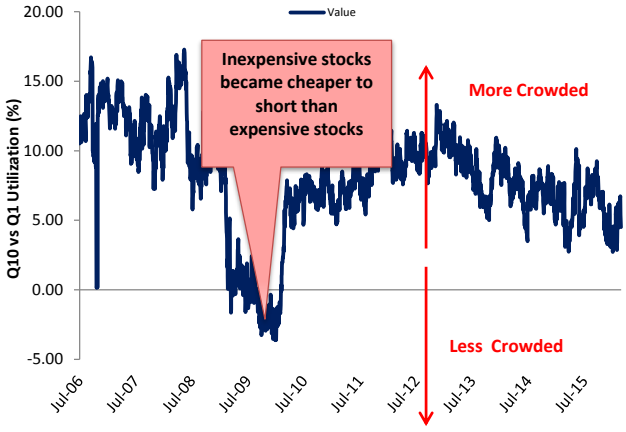
[7] The MPC is based on equally weighted factor portfolios.

the MPC measure seems to suggest that value factor is somewhat crowded and approaching the 2007 levels.

Figure 56: MPC of long value minus market


*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

Figure 57: Incremental utilization for value




*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Markit, Thomson Reuters, Deutsche Bank*

Analyzing the relative MPC of the quality portfolio during the financial crisis yields similar intuitive results (see Figure 58). In the bull market, leading up to the financial crisis, the quality portfolio became less crowded as inventors chased risky, low quality stocks. However, during the onset of the financial crisis, investors flocked to quality stocks. As such, quality became more crowded. Again, this aligns well with our intuition in the events surrounding the financial crisis.

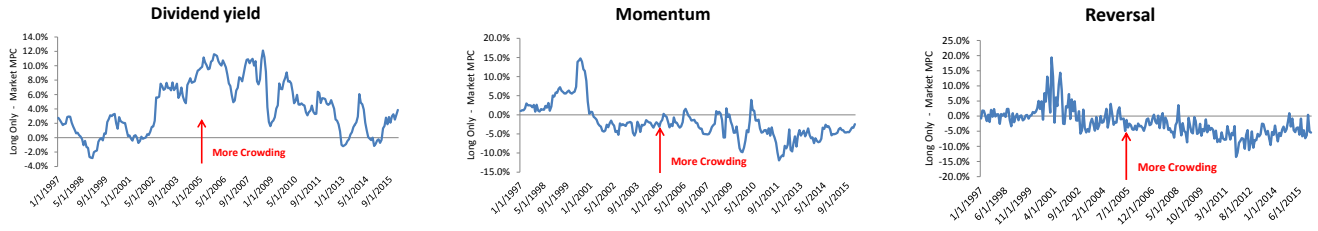
*In the bull market, leading up to the financial crisis, the quality portfolio became less crowded as inventors chased risky, low quality stocks.*

Figure 58: MPC of long quality minus market


*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

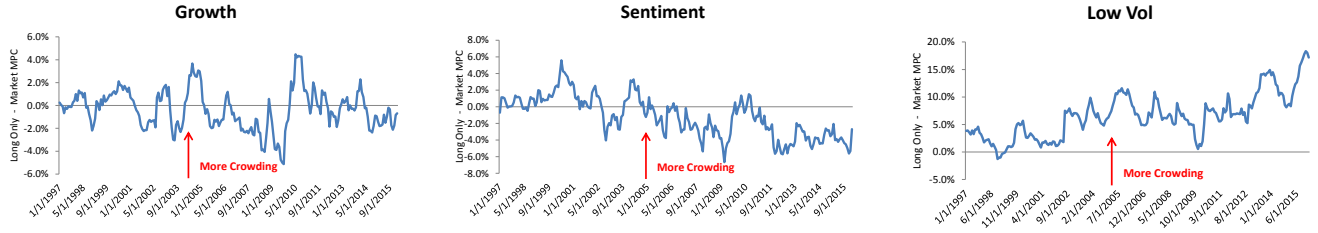
Since our MPC measure of crowdedness makes intuitive sense, we run it for all other standard quantitative portfolios (see Figure 59 and Figure 60). Interestingly, based on MPC, most factors do not show excessive levels of crowding relative to history. The one exception is the low volatility portfolio which shows a continued trend of crowdedness.

Figure 59: MPC of factor portfolios: dividend yield, momentum, and reversal




*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

Figure 60: MPC of factor portfolios: growth, sentiment, and low volatility




*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

## Performance of MPC

We further study the relationship between MPC and future returns (see Figure 61)[8]. The results are fairly mixed.

Figure 61: MPC and future factor returns

| Factors | MPC long only rel. to R3K | | | |
|---|---|---|---|---|
| Dividend yield | -7% | 2% | -8% | -1% |
| Earnings yield | -28% | -21% | -13% | -11% |
| Momentum | 12% | 17% | 5% | 14% |
| 1M Reversal | 2% | 7% | 8% | 27% |
| EPS growth | 25% | 21% | 6% | 8% |
| ROE | -13% | -13% | -2% | 14% |
| Low Vol | 2% | 10% | 4% | -7% |
| Earnings revisions | 9% | 16% | 22% | 14% |
| Avg. Monthly Returns | 1 to 6 | 7 to 12 | 13 to 18 | 19 to 24 |

*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

In fact, running the multivariate regression analysis (see Figure 88) shows that most coefficients are statistically insignificant.

[8] Future returns are based on the long only leg of the equally weighted factor portfolio excess to the Russell 3000 equally weighted index.

Figure 62: Average coefficients for multivariate regression using MPC

| | 12 Month Forward Returns | | 24 Month Forward Returns | |
|---|---|---|---|---|
| | Average Beta | Average T-stat | Average Beta | Average T-stat |
| Crowding | (0.34) | (0.18) | (0.38) | 0.08 |
| Crowding squared | (0.01) | (1.60) | 0.00 | (0.77) |
| Crowding change | (0.79) | (1.99) | (0.94) | (1.88) |
| Crowding saturation | (0.79) | (1.37) | (0.94) | (1.09) |

*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

### Mean downside pairwise correlation (MDPC)

We also test a similar measure to MPC called mean downside pairwise correlation or MDPC. The only difference between the two measures is that MDPC computes the stocks pairwise correlations only when the market was down. We found that MPC and MDPC are strongly correlated (i.e., greater than 90%) and as such the results did not differ significantly.

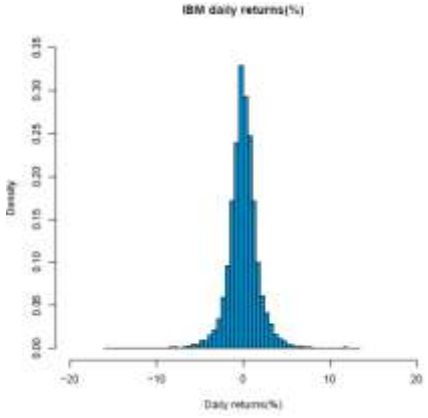
## 6. Minimum tail dependence

*We found that MPC and MDPC were strongly correlated (i.e. greater than 90%) and as such the results did not differ significantly*

A few years ago, we published and analyzed tail dependence as a measure for crowding (see Cahan, R., Alvarez, M., Luo, Y. et al. [2012a] and Luo et al [2014]). In this section, we update and enhance upon our previous research. We proposed a metric based on the likelihood that two stocks have large negative returns at the same time. After all, this is the co-movement that we really care about. Stocks can move for all kinds of reasons – supply/demand imbalances, news flow, etc. – which can disguise the true linkage between stock returns. However, in times of stress, the real relationships between stocks are laid bare, as we have seen all too often in recent years. A convenient way to capture this idea is a Copula. Without getting too bogged down in the technical details (interested readers can refer to Nelsen [1999] for example), a Copula is a function that describes the dependency structure between two random variables. One of the key features of a Copula is that the Copula is independent of the marginal distribution of each of the random variables.

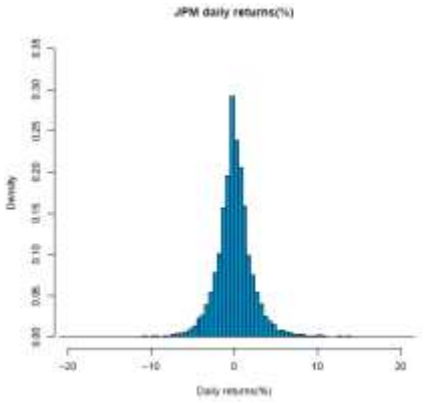
The Copula describes how likely it is that stock i has a return of $r_i$ at the same time as stock j has a return of $r_j$. In particular, we care mainly about how likely it is that stock i has a big negative return at the same time as stock j also has a big negative return. We can best visualize this idea with some pictures. Consider the marginal or independent distribution of three stocks: JP Morgan, IBM, and Citigroup (see Figure 63 to Figure 64).

Figure 63: IBM distribution



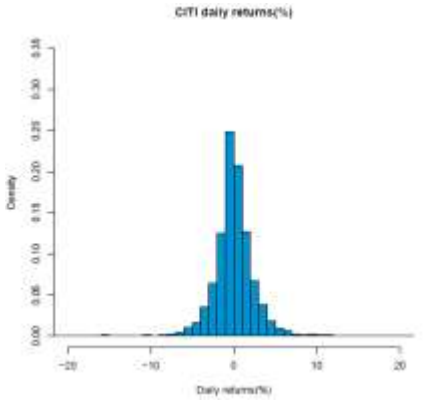

*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

Figure 64: JPM distribution




*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

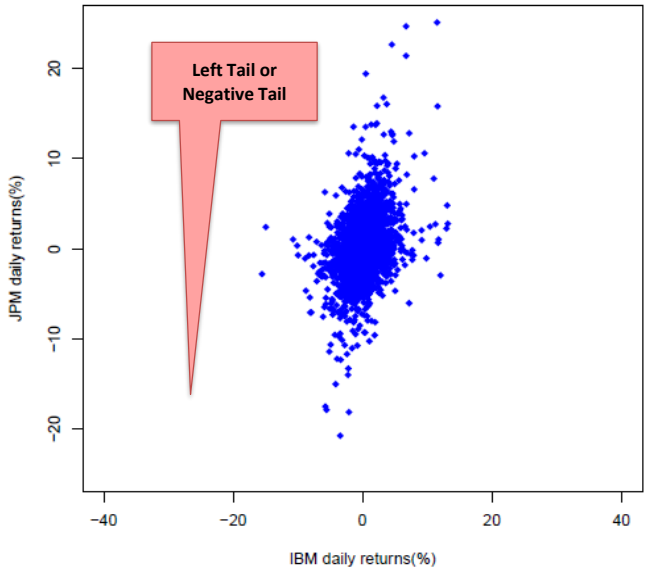
Figure 65: CITIGROUP distribution




*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

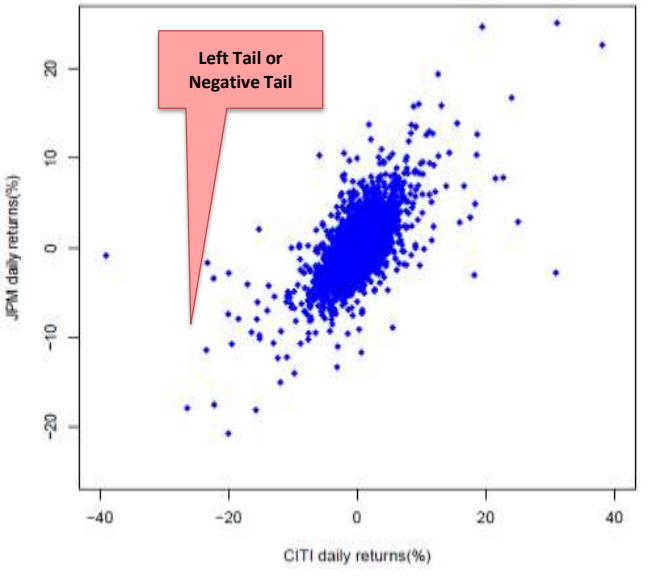
Since, JPM and CITIGROUP are within the same industry group, we would expect them to be more correlated than JPM and IBM or CITIGROUP and IBM. The correlation coefficient measures the overall strength of the relationship between the two stocks. However, the correlation coefficient gives little information about how stocks co-move across the tails of the distribution. For example, JPM and IBM have an overall return correlation of 39% whereas JPM and CITIGROUP have an overall correlation of 66%. However, this says little about the correlation within the tails. Figure 66 and Figure 67 hint that JPM and IBM are less correlated in the left tail (far left of the chart) than JPM and CITIGROUP. For JPM and CITIGROUP, it suggests that a large negative return for one stock will likely be accompanied by a large negative return of the other stock (i.e. JPM and CITIGROUP are more strongly correlated in the left tail than JPM and IBM).

Figure 66: JPM and IBM




*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

Figure 67: CITIGROUP and JPM



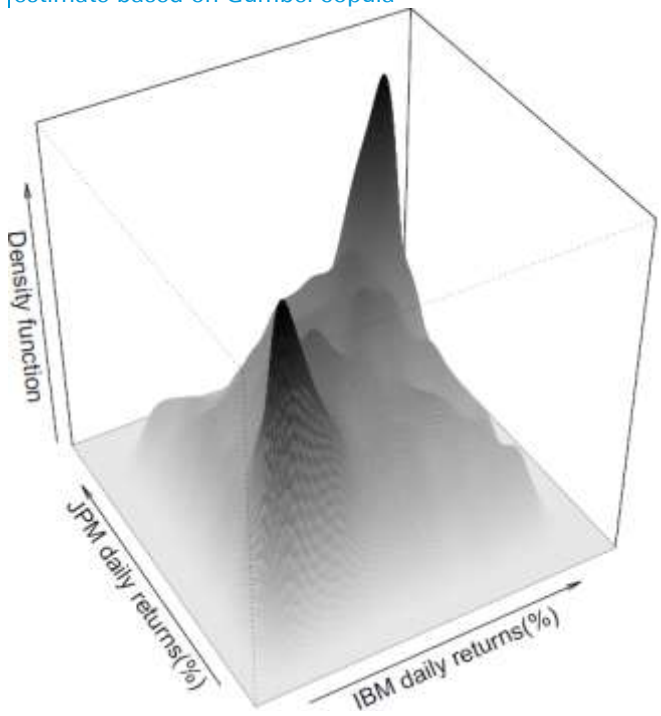

*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

This is where Copulae come into the picture. A Copula can provide a more complete codependence structure across the bivariate distributions. It can provide control and correlation structure over what parts of the distributions the variables are most strongly (or weakly) associated.

*A copula can provide a correlation structure across the distributions. It can provide control and structure over what parts of the distributions the variables are most strongly (or weakly) associated*

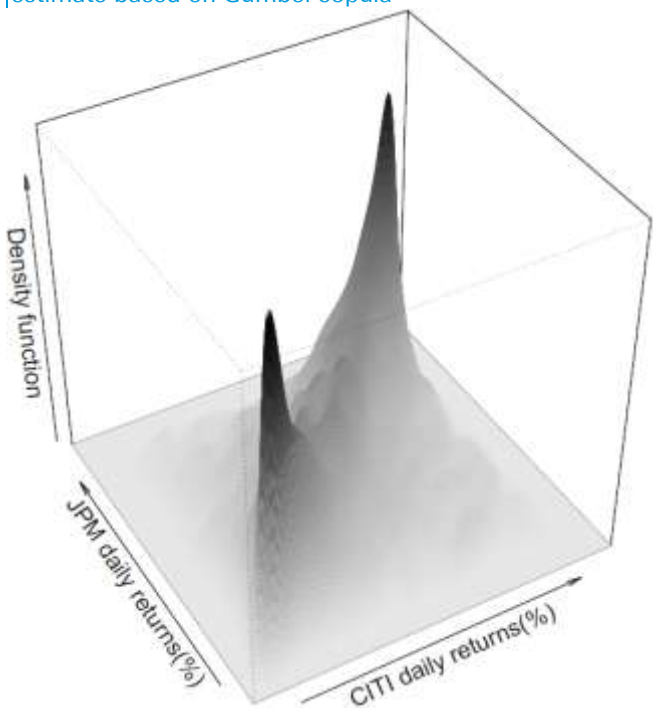
Consider two copulae, fit at a given point in time to two pairs of stocks: JPM-IBM (see Figure 68) and JPM-CITIGROUP (see Figure 69).The "surface" in these charts is the probability density functions, i.e., it tells us how likely we are to see a given pair of returns occurring together (note that the returns are normalized between 0 and 1, such that simultaneous large negative returns are in the back corner of each chart for easier visualization). The large spikes in the end of the charts show that in both cases, there is a much higher probability of having large negative returns than large positive returns, or any other possible combination of returns.

Figure 68: JP Morgan–IBM Tail Dependence, point estimate based on Gumbel copula




*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

Figure 69: JP Morgan–Citigroup Tail Dependence, point estimate based on Gumbel copula



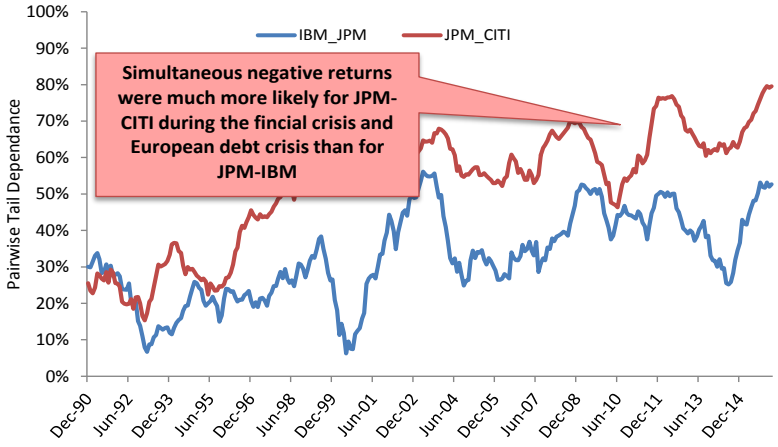

*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

More formally, we can measure a quantity called the asymptotic tail dependence which describes how variables co-move in the tails of the distributions. In rough terms, it measures how likely it is that we see large negative returns to both stocks at the same time, i.e. how "high" is the spike in the back corner of the charts. In this particular case, it is clear that there is a higher probability of JPM and CITIGROUP having concurrent negative returns, compared to JPM and IBM. This is probably not a surprise.

*More formally, we can measure a quantity called the asymptotic tail dependence which describes how variables co-move in the tails of the distributions. In rough terms, it measures how likely it is that we see large negative returns to both stocks at the same time*

Suppose we extend the example and consider how the tail dependence of our two pairs of stocks changes through time (see Figure 70). Once again, the results are intuitive. Through the financial crisis, the tail dependence of JPM-CITIGROUP increased dramatically, much more so than JPM-IBM.

Figure 70: Time-series of JPM–CITIGROUP and JPM–IBM Tail Dependence




*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

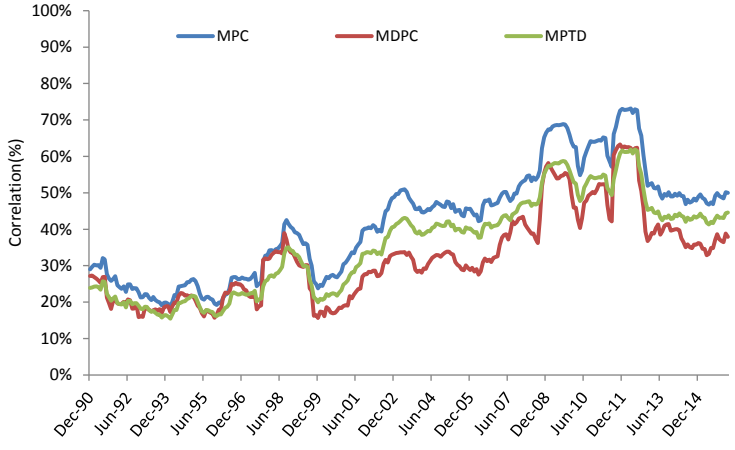
## Tail dependence as a herding measure

Next, we generalize this concept to the whole universe. At the end of each month, we obtain daily returns for the past year for each stock in the Russell 1000 at that point in time. Then, for each pair of stocks, we compute the negative tail dependence by using maximum likelihood to fit a Gumbel copula to the daily returns.[9] As one might imagine, this methodology is quite processor and data intensive.

*To make a long story short and avoid unnecessary analysis, our results find that MPTD is highly correlated to MPC and MDPC*

To make a long story short and, our results find that MPTD is highly correlated to MPC and MDPC (see Figure 71). In fact, MPTD is even more correlated to MDPC. This is somewhat expected and intuitive. Recall that MPC proved to be a descent crowding measure for quant factors.

Figure 71: Russell 3000 stock wise correlation using MPC, MDPC, and MPTD




*Source: Bloomberg Finance LP, Compustat, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank*

[9] There are a wide range of different copulas that have been tested on stock return data. We pick the Gumbel copula for a few reasons. First, it is asymmetric and allows greater tail dependence in the tail (note that for this reason we multiply all our returns by -1 before fitting the copula). Second, it is widely used in the academic literature on modeling the dependency structure of stock returns (see for example Ruenzia and Weigert [2012], also reviewed in our March 2012 edition of Academic Insights).