Standing out from the crowd

New metrics for measuring factor crowding

Blame the crowd?
Post the financial crisis there was much soul-searching among quantitative investors; the tried and true strategies that quants had come to rely on were not adding as much value as they used to. One of the oft-cited refrains was that quant strategies simply became too crowded, and as a result most of the alpha was arbitraged away. This argument has become almost dogma in the quant community, but is it actually true?

Quantifying factor crowdedness
In this report, we use three unique data sources to develop proxies for factor crowding: securities lending data, high frequency data, and institutional ownership data. Using these proxies, we find empirical evidence that factor crowding did indeed peak around 2007, then declined significantly through the financial crisis, and is now starting to climb back upwards.

Crowdedness and factor performance
We find the link between crowdedness and factor performance is not clear-cut. Our evidence suggests that in the long-run, crowdedness is bad for performance, but in the short-run it can actually help drive factor returns for some factors. However, in risk space the results are intuitive: higher crowding is positively correlated with larger drawdowns in factor performance.
# Table of Contents

A letter to our readers ........................................................................................................ 3
Don’t all run for the exits at once .................................................................................... 3

**Two’s a crowd** ............................................................................................................. 4
Academic evidence ........................................................................................................... 4
Measuring crowdedness: A cross-sectional approach ...................................................... 4

**Short stories** ............................................................................................................... 7
Detecting crowding with securities lending data ................................................................. 7
Regression analysis ......................................................................................................... 7
It’s a multifactor world ...................................................................................................... 14
Long-run crowdedness .................................................................................................... 16
Crowding and factor performance .................................................................................. 17
Month-end effects .......................................................................................................... 20

The tale of the tape ......................................................................................................... 21
What about high frequency data? ..................................................................................... 21

**Institutional ownership** ............................................................................................... 24
Can institutional ownership proxy for crowdedness ......................................................... 24

**Quantifying a crisis** .................................................................................................... 26
What happened in the summer of 2007? ....................................................................... 26

**Innovate to survive?** ................................................................................................... 29
Are new data sources really the answer? ......................................................................... 29
Disentangling the results ................................................................................................. 32

**Conclusion** .................................................................................................................. 34

**References** .................................................................................................................. 35
A letter to our readers

Don’t all run for the exits at once

Since we launched our quant research efforts here at DB two years ago, we have tended to eschew the usual “Outlook for 201X” reports. While it is interesting, and indeed useful, to debate which factors are likely to work in the coming year, we would argue that often such predictions are grounded more in speculation than concrete data.

Therefore, in this report we want to take a different road. Instead of jumping straight into pronouncements that factor X will beat factor Y in 2012, let’s start by thinking about the external forces that drive factor performance. First, we have macroeconomic forces. These we have studied in detail already – for example, on the alpha side we showed that macroeconomic variables can be used to predict future factor returns. At the same time, from a risk perspective we showed that many factors carry a hidden exposure to volatility that can savage a factor around turning points in risk appetite. Our conclusion – which will come as no surprise to anyone who has navigated the macro-driven environment of the past few years – is that macro forces play a powerful role in driving quant factor performance.

The second force is that of arbitrage. This inexorable force has been the catalyst for much of our alpha research. Every time we search for a new signal, or a fresh data source, we are implicitly acknowledging that markets are dynamic; factors are discovered, do well for a time, eventually become crowded, and finally get arbitraged out of existence.

Or do they? This Darwinian view of factor evolution has become close to dogma in the quant finance community in the post-2007 landscape. But is there really any empirical evidence to support this thesis? Do factors, even the bread and butter factors like value and momentum, really become crowded? And even if they do, what impact does that have on their performance? We are the first to admit we have been as guilty as anyone at jumping on the “quants need to innovate to survive” bandwagon. With this report we will hop off, if only temporarily, to seriously examine the veracity of the factor crowding argument.

Our first goal is to find a way to measure crowdedness. In this research we study how we can use securities lending data, high frequency data, and institutional ownership data to construct useful proxies for factor crowdedness. The next question is obvious but tricky to answer: if there is crowding, does it impact factor performance? Yes and no. Crowding can actually help some factors in the short term (think of momentum for example; if someone else buys a past-winner stock after you, it will probably help you by pushing the stock higher still). At the same time, crowding can increase the tail risk of a factor; if everyone runs for the exits at the same time in a crowded theater things will be worse than in an empty building.

As you will see, we believe that crowdedness is an important additional metric to consider when choosing factors (and factor weights) in an alpha model.

Regards,

Yin, Rocky, Miguel, Javed, John, and Sheng

Deutsche Bank North American Equity Quantitative Strategy

---


Two’s a crowd

Academic evidence

They say two’s a crowd, but when it comes to factors it is not that easy. Pinning down what crowding actually means is difficult. If we start with the academic literature, there is a small but growing body of literature on the topic of crowded trades. A notable recent paper by Pojarliev and Levich [2011] focuses on the currency market, but the methodology is applicable to other markets too. The authors use a simple and intuitive measure of crowdedness: they do time-series regressions of currency fund returns onto style factor returns, and then count the number of funds that have a significant positive exposure to each style less the number with a significant negative exposure. Tracking this metric over time allows them to build up a picture of the number of funds that are correlated with each style at each point in time, which is their proxy for crowding.3

A second useful paper is Hanson and Sunderam [2011]. Their approach is to use short interest data as a proxy for the directional views of institutional investors. They use cross-sectional regressions to measure the difference in short interest between stocks that look attractive on different factors and those that do not. For example, consider the value factor. If expensive (i.e. unattractive) stocks are more heavily shorted than cheap stocks at a given point in time, then this might indicate that institutional investors (who are more likely to short) are heavily invested against the value factor. The advantage of their approach is that it is cross-sectional and therefore less backwards looking; it does not rely on trailing regressions to compute style sensitivities.

Of course, stepping back a little it is easy to see that the approaches used to measure crowding are mainly applications of the much larger body of literature on performance attribution. The two approaches described above closely mirror the two common approaches to attribution: time-series analysis (in the spirit of Heston and Rouwenhorst [1994] for example) and cross-sectional techniques (for example Grinold [2006]). Since a common goal of attribution is to determine which underlying factors are driving portfolio returns, one can imagine how this literature might be useful for potentially measuring crowdedness.

Two other fields of study can also be helpful when exploring crowdedness. First, studies that try to measure market efficiency can suggest indirect techniques for measuring the strength of arbitrage for different factors – see for example Hogan, Jarrow, Teo, and Warachka [2004] who devise a test for statistical arbitrage and apply it to value and momentum strategies. Second, the literature on hedge fund replication can be instructive. The idea here is to measure the time-varying exposures that hedge funds take to underlying factors, and then use those exposures to replicate the payoff patterns of hedge funds without incurring the usual fees (see Hasan hodzic and Lo [2007] for a good overview). Again, it is easy to see that some of these techniques could be useful in measuring the number of funds that are chasing a given factor at a given point in time.

---

3 For a more detailed review of this paper, see: Cahan et al., 2011, “Academic Insights”, Deutsche Bank Quantitative Strategy, 20 January 2011
Measuring crowdedness: A cross-sectional approach

In our study, we focus mainly on the cross-sectional approach. As mentioned, the strength of this approach is that it does not rely on a backwards-looking, rolling time-series regression to estimate style sensitivities. It also avoids the need to have fund-level return data, something that most of our clients do not have access to. Of course, it comes with its own challenges, which we will discuss in the coming pages.

Our methodology is based most closely on that 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

\[
C_{i,t} = c + \sum_{j=1}^{J} \sum_{q=2}^{Q} \beta_{i,t,j,q} D_{i,t,j,q} + \sum_{q=2}^{Q} \beta_{i,t,\text{size},q} D_{i,t,\text{size},q} + \sum_{q=2}^{Q} \beta_{i,t,\text{vol},q} D_{i,t,\text{vol},q} + \epsilon_{i,t} (1)
\]

where \( C_{i,t} \) is the crowdedness score 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 \( D_{i,t,\text{size},q} \) and \( D_{i,t,\text{vol},q} \) 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 \( \beta_{i,t,\text{value},10} \) 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, suppose we use short interest percent as our proxy variable, and consider a case with one factor (\( J = 1 \)), let’s say a value factor. Then \( \beta_{i,t,\text{value},10} \) will measure the difference in short interest between stocks in the most expensive (i.e., unattractive) value decile (\( q = 10 \)) compared to those in the cheapest (\( q = 1 \)). If this coefficient is positive and significant, it would indicate that there is heavier shorting taking place in stocks that are expensive relative to stocks that are cheap. By tracking \( \beta_{i,t,\text{value},10} \) over time, we can build up a picture of how the shorting differential between cheap and expensive stocks (after controlling for other factors) evolves over time. We can then make inferences about whether that tells us something about the amount of money chasing the value factor.

Proxies for investors’ capital allocation

Of course, short interest is not the only proxy we can use to try to assess where investors are directing their capital. In this paper we will test a range of dependent proxy variables, including:

- **Utilization**: In our work on securities lending, we showed that utilization—which measures the percent of shortable inventory that is currently lent out (rather than the percent of total shares) – is a better measure of the true level of shorting in a stock.

- **Cost of Borrow**: The cost of borrowing a stock is jointly determined by the supply and demand for shorting a particular stock. Like utilization, we argue this factor is a better measure of the true level of shorting in a stock, and hence may give a more accurate picture of factor crowding.

- **Intraday Order Imbalance**: If there is a heavy buy or sell order imbalance for certain factors, this could be a short-term indicator that a factor is crowded. For example, if everyone is trying to sell expensive stocks and buy cheap stocks, perhaps around month-end, this could indicate crowdedness in the value factor.
- **Probability of Informed Trading:** This is a non-directional metric, but could also be a useful proxy for crowdedness. For example, suppose the top and bottom decile stocks based on the value factor have higher PIN; this could indicate that investors have stronger views about the direction of these two baskets, compared to the universe as a whole.

- **Institutional Ownership:** This metric has well documented drawbacks, the chief of which is its lack of timeliness given the lag with which funds are allowed to report holdings. Nonetheless, it could be useful to get a longer-term picture of which factors institutional investors hold most heavily. For example, if there is heavy institutional ownership for cheap stocks and little institutional ownership for expensive stocks, it could indicate value is becoming crowded.

In the rest of this report, we will examine each of these metrics in more detail, to see what they can tell us about factor crowdedness.
Short stories

Detecting crowding with securities lending 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. More importantly, it is likely that the majority of shorting activity is being driven by professional institutional investors who in turn are more likely to follow systematic quantitative strategies compared to retail investors.4

In their recent paper, Hanson and Sunderam [2011] do a good job of illustrating how short interest data can be used as a proxy for measuring factor crowding. However, their paper has two key limitations. First, they use 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 they rely exclusively on short interest (i.e. the percent of total shares that are currently sold short) as their measure of shorting demand. 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, 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.

A better way to measure the level of shorting

To address these drawbacks, we leverage a unique database of securities lending activity from a company called DataExplorers. We discussed this data set in considerable detail last year in our report “The long and the short of it”, so here we will focus on two key aspects: its greater timeliness and the more granular supply-demand metrics that it provides.5 First, with regards to timeliness, the Data explorers database is updated daily and is available on a T+2 basis. This negates the lag problem that comes with Compustat data, and also allows analysis on a daily instead of monthly frequency. Second, the DataExplorer database captures both the supply and demand side of the securities lending market. In particular, a metric called Utilization is useful. This is defined as the number of shares currently being shorted, divided by the number of shares available for lending. In other words, it measures the shorting activity in a stock as a percent of the pool of lendable supply in that stock, rather than as a percent of the total shares on issue. This gives a more accurate depiction of the true level of shorting in a stock. For example 100% Utilization indicates that the entire shortable supply in the stock is currently out on loan, hence such a stock is heavily shorted regardless of whether its short interest is 2% or 10%.

Regression analysis

Using the daily Utilization metric from DataExplorers, we conduct the cross-sectional regression in equation (1) on a daily basis from 2006 (when the database starts) to present. To keep things simple, for the first part of our study we only include two quant factors, Value (represented by trailing earnings yield) and long-term Momentum (represented by 12M-1M price momentum). We also include control variables for Size (market capitalization) and

4 For example, Hanson and Sunderam [2011] cite studies estimating that up to 85% of short positions in the U.S. market are driven by hedge fund investors.

Volatility (daily return volatility computed over a trailing one year window). For each factor, we order our stocks such that decile 10 contains the least attractive stocks (i.e., expensive stocks for value or past-losers for momentum) and decile 1 contains the most attractive. The omitted category in our regression is decile 1, hence the coefficient for decile 10 represents how much more Utilization unattractive stocks have relative to attractive stocks, holding all other variables constant.

We find that both Value and Momentum have a statistically significant difference in Utilization between attractive and unattractive stocks

Figure 1 shows the coefficients for decile 10 stocks for the Value and Momentum factors. 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 for the blue line, which represents the Value factor, is around 5%. This indicates that over the long-run, expensive stocks tend to have Utilization that is 5% higher than cheap stocks, after controlling for differences in size and volatility. In other words, expensive stocks tend to be more heavily shorted than cheap stocks. The same applies for Momentum – on average past-losers have higher Utilization than past-winners. Furthermore, as shown the Figure 2, these differences tend to be statistically significant most of the time (roughly a t-statistic greater than 2 or less than -2).

We find that both Value and Momentum have a statistically significant difference in Utilization between attractive and unattractive stocks

Figure 1: Incremental Utilization for Q10 stocks versus Q1 stocks (%)

Figure 2: T-statistic for regression coefficients

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

Interpreting the time trend

The Utilization differential for Value is consistent with intuition: crowding was highest pre-2007, then declined through the financial crisis, and is now climbing again

Crowdedness in Momentum is more cyclical, as we would expect for a factor that implicitly relies on following the crowd

The time trend is probably the most interesting aspect of the above charts. If we focus on Value first (the blue line) there was a relatively steady Utilization differential (around 10%) between expensive and cheap stocks before the start of the financial crisis. However, through the depths of the crisis, this relationship actually reversed; in fact for a period of time cheap stocks were actually being shorted more than expensive stocks (hence a negative differential). This is fairly intuitive if we think about it. During the crisis, most deep value stocks were cheap for a reason (i.e., they were highly distressed, often financial stocks) and hence risky for managers to short. As well, there were restrictions on shorting certain financials, which probably helped to reinforce the negative relationship. However, as the crisis abated, we note that the original positive relationship slowly returned, and indeed the Utilization differential has been steadily climbing over the past year and is now nearly back to pre-crisis levels. Is this a cause for concern?

The Momentum factor (the red line) shows a broadly similar trend. In the crisis it too reversed. Again, this is fairly intuitive. With the massive reversal in risk appetite that was precipitated by the crisis, the low-beta, defensive names that had been underperforming (i.e., had bad momentum) suddenly became the new outperformers. Hence investors who had been shorting defensive names reversed their positions and began to short high-beta past winners instead.
The time-trend of crowdedness mirrors the amount of shorting in the market

But does it measure crowdedness?
The time trend in Figure 1 is interesting because it is consistent with the generally accepted prior that there was more crowding in quant factor space in the lead up to the “quant crisis” in the summer of 2007. Once the full blown financial crisis hit, the picture is consistent with significant deleveraging occurring relative to the standard Value and Momentum strategies. Another way to see this is to overlay market aggregate Utilization with the Utilization differential for each factor. Figure 3 shows the result for Value and Figure 4 shows Momentum. In both cases, the gray line shows market-wide aggregate Utilization.

![Figure 3: Value coefficient overlaid with Aggregate Utilization](source)

![Figure 4: Momentum coefficient overlaid with Aggregate Utilization](source)

From these charts, it is clear that there was indeed a significant decline in market-wide shorting activity during the financial crisis, and since then the level of shorting activity has failed to recover to its pre-crisis level. Therefore, it would be easy to argue that the apparent decline in factor crowding after the crisis was really just a byproduct of a market-wide fall in leverage and tells us nothing about the relative crowdedness of factors before and after the crisis. However, this argument misses a key point: the regression analysis is designed to measure the difference in shorting between attractive and unattractive stocks. In fact, the intercept term in equation (1) will account for changes in the average, market-wide level of Utilization. So by construction our metric should be relatively immune to changes in the level of market-wide shorting activity.

The only caveat is that we are measuring the Utilization differential in absolute terms. Clearly a difference in Utilization of 10% between attractive and unattractive stocks is more meaningful in a market where the average Utilization is 20%, compared to a market where it is 40%. Therefore, our charts may understate the crowdedness of factors in today’s environment compared to the past. With this in mind, we run a robustness test where we replace the dependent variable in our regression with the z-score of Utilization instead of raw Utilization. This means the coefficients for the Q10 dummy variables are measuring how much more shorting there is in unattractive stocks, in units of standard deviation instead of raw percent. Figure 5 compares the results from the two regressions for the Value factor, and Figure 6 shows the Momentum factor.

It turns out that it doesn’t make much difference. As a result, we prefer to use our original formulation, since it is more intuitive to say that “Utilization is around 10% higher for expensive stocks compared to cheap stocks” rather than “Utilization is around 0.4 standard deviations higher for expensive stocks compared to cheap stocks.”
Another important question is whether Utilization is really a better measure than Short Interest. In Figure 7 and Figure 8 we compare the Value and Momentum regression coefficients obtained using Utilization to those we would get if we used Short Interest. Interestingly, there are some significant differences. For example, with the Value factor we would have got a major spike in apparent crowdedness in 1H 2010, had we used Short Interest instead of Utilization. On the other hand, the Utilization regression shows two jumps in Momentum crowdedness, around July 2010 and July 2011, which are missed by the Short Interest regression.

Utilization versus Short Interest

Figure 7: Incremental Utilization versus incremental Short Interest for Q10 Value stocks

Figure 8: Incremental Utilization versus incremental Short Interest for Q10 Momentum stocks

We find that Cost of Borrow is a close proxy for Utilization

Which is more accurate? One way to make a determination is to look at the cost of borrow. To our mind, cost of borrow is the ultimate arbiter of how hard a stock is to borrow. After all, the cost of borrow is determined by the pull and push of shorting demand and the lendable supply in each stock. If we re-run our regressions using cost of borrow as the dependent variable, and compare the results to those we got with Utilization, we get Figure 9 and Figure 10.\(^6\) It turns that the results are almost identical, for both Value and Momentum.

\(^6\) We use the so called DCBS score as a proxy for cost of borrow. This is a score provided by DataExplorers that ranges from 1 to 10, with 1 being cheap to borrow and 10 being expensive to borrow. For more details, see: Cahan, et al., 2011, “Signal Processing: The long and the short of it”, Deutsche Bank Quantitative Strategy, 18 January 2011
Our view is that Cost of Borrow is the ultimate arbiter of where the supply-demand balance lies. This confirms our view that Utilization is a better proxy for where the supply-demand balance lies in each stock. A stock may have high short interest, but this is fairly irrelevant if that stock also has ample supply available for shorting. The result also suggests that one can think of Utilization and costs interchangeably. If expensive stocks have higher Utilization than cheap stocks, it also means that investors are willing (or being forced) to pay more to borrow expensive stocks. Ultimately, we argue that this makes Utilization a more accurate measure of true level of shorting in a stock, and hence a more accurate barometer of the amount of money chasing a particular strategy.

We find a monotonic relationship between how stocks rank on a factor and their level of shorting. So far we have been looking at the difference in Utilization between the best and worse stocks on a particular factor. But if our thesis that 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 11 and Figure 12 show the difference in Utilization between decile 1 (most attractive) stocks and the other deciles, for Value and Momentum respectively and measured over the whole history from 2006 to now. 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.
Of course, this monotonicity applies on average, but there have been times when it broke down. In Figure 13 we plot the difference in Utilization between Q10 versus Q1 Value stocks (the blue line) and Q5 versus Q1 Value stocks (the red line). In the fallout from the credit crisis, there was a noticeable reversal in the normal relationship. As mentioned previously, this was a time when value stocks were generally cheap for a reason, and as a result most investors steered away from basic valuation-based strategies. In other words, the factor became less crowded, or even in a sense “negatively” crowded (i.e. on average investors were taking a contrarian view towards value).

**Figure 13: Normal monotonic relationship between Utilization and Value broke down post the financial crisis**

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

The tail wagging the dog?

Figure 13 raises another important question: are we really measuring factor crowdedness? Currently, big macro themes are dominating stock performance. So it is plausible that what we are observing here is just a second derivative effect, or a case of the tail wagging the dog. Let’s take Momentum as an example of what we mean by this. In our paper “Reviving momentum: Mission impossible?” we showed that momentum can take on a significant positive or negative exposure to beta at different points in time (Figure 14).7

**Figure 14: Expected correlation: 12M-1M momentum and beta**

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

---

Supposed we are at January 2009, in the depths of the financial crisis. Then clearly following a momentum strategy (buying past winners) is almost exactly opposite a strategy of buying high beta stocks and selling low beta stocks (around -90% correlation). So the fact that the momentum strategy appears “uncrowded” at that point in time (see Figure 4) might have more to do with the market’s view on beta than it does on momentum per se. If the market in general wants to buy low-beta, defensive names and short high-beta names, then doesn’t the low crowding just reflect investors’ aggregate view of the market, rather than a lack of money chasing the momentum strategy? In other words, we already know that not all shorting is quant-driven. So could it be that certain stocks that look “bad” on a particular factor are being shorted for other reasons by non-quants, and the apparent crowdedness is just the relic of a factor’s time-varying exposure to other themes the market cares about?

Our first counter to this argument is that we have been careful to control for two potential macro-themes: size and volatility. Recall in equation (1) we included dummies representing size and volatility deciles. In Figure 15 and Figure 16 we show the regression coefficients and t-statistics for the Q10 dummy for the Volatility and Size factors.

![Figure 15: Incremental Utilization for low volatility stocks and small cap stocks (%)](source)

![Figure 16: T-statistic for regression coefficients](source)

It turns out that both are quite significant in explaining the cross-section of Utilization so it is definitely important to control for these two variables (low volatility stocks tend to have lower Utilization than high volatility stocks and small cap stocks tend to have higher Utilization than large caps). More importantly, Value and Momentum show significant Utilization differentials between attractive and unattractive stocks, even after controlling for these two variables. In other words, Value and Momentum are important in explaining the level of shorting in stocks, above and beyond Size and Volatility.

**It’s all about beta**

However, the single biggest theme of all recently has been beta, which has been serving as a proxy for the risk-on/risk-off trade that has dominated market psyche through the financial crisis and subsequent sovereign debt crisis. Therefore, as a second robustness test, we rerun our regressions after replacing the volatility factor with a beta factor (since they are highly correlated, we cannot have both at the same time). Figure 17 shows the Q10 coefficient for Beta and Figure 18 shows the corresponding t-statistic. We find that Beta on average is not statistically significant, and including Beta in the regression only changes the coefficients on Value and Momentum marginally. This gives us some comfort that our crowdedness metrics are indeed picking up the ebb and flow of capital allocation towards Value and Momentum as stand-alone strategies, as opposed to proxies for bigger market themes.
**It’s a multifactor world**

So far we have concentrated on what are probably the two most common quantitative strategies: Value and Momentum. But these two strategies are also quite popular with non-quants. Valuation in particular is present in most investment portfolios in some form – it is rare to find a manager, quant or fundamental, who intentionally buys expensive stocks. As a result, Value and Momentum might appear crowded, but in fact most of the capital invested against these strategies might come from non-quant investors.

Does it matter? We could argue that it doesn’t. After all, dollars are fungible; a dollar of quant AuM looks the same as a dollar of fundamental AuM. Or we could argue that crowding by quants increases the downside risk of a factor in a forced deleveraging scenario. We will explore these ideas in more detail later in this report. In the meantime, we will focus instead on extending our analysis beyond Value and Momentum into other common quant strategies.

In Figure 19 we pick a set of six common quant factors to use in our analysis. In addition to Value and Momentum, we include 1-month Reversal, 3-month Earnings Revisions, ROE, and year-on-year EPS Growth.

### Figure 19: Average cross-sectional rank factor correlations

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
<th>Momentum</th>
<th>Reversal</th>
<th>Revisions</th>
<th>ROE</th>
<th>Growth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Momentum</td>
<td>0.01</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reversal</td>
<td>-0.04</td>
<td>-0.03</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Revisions</td>
<td>0.05</td>
<td>0.37</td>
<td>0.06</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ROE</td>
<td>0.59</td>
<td>0.16</td>
<td>0.04</td>
<td>0.15</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Growth</td>
<td>0.09</td>
<td>0.33</td>
<td>0.06</td>
<td>0.38</td>
<td>0.18</td>
<td>1</td>
</tr>
</tbody>
</table>


Because we are using a cross-sectional regression framework, it is important to avoid having factors that are highly correlated. As shown in the table above, the correlations are fairly moderate in most cases – the highest being the 59% correlation between ROE and Value.

Rerunning our regressions with these six variables (plus Size and Volatility as controls) we get the charts in Figure 20 through Figure 22. Figure 25 shows the average absolute t-statistics for the Q10 regression coefficients, over time.
Even with six factors, we still get a similar picture for Value and Momentum

In Figure 22 we find that Growth and ROE show a little more cyclicity, however, on average the difference in Utilization between attractive and unattractive stocks based on these factors is not statistically significant over time (Figure 25).

Avoiding the crowds

Overall, the results are quite useful. Of the six factors, Value shows the most worrying signs; crowdedness has been increasing steadily since the financial crisis, and is now at levels last seen prior to the quant crisis in August 2007, albeit with a slight downtick in the most recent months. In contrast, most of the other factors do not show consistent signs of increasing crowding. The one exception is Momentum, where the crowdedness tends to be quite cyclical, with recent peaks around mid-2010 and mid-2011. However, Momentum is, by construction, a bit of an odd beast. Whenever the market is expecting the current trend in
equity performance to continue, then we would expect crowdedness to rise as investors continue to short past-losers relative to past-winners.

Putting everything together, we can measure the crowdedness of quant strategies overall by constructing a simple multifactor model with these six factors. We z-score each factor and form an equally-weighted alpha signal. Then we rerun our regression analysis, using the multifactor alpha instead of the six individual factors. Figure 22 shows the incremental Utilization for unattractive stocks (Q10) compared to attractive stocks (Q1) and Figure 25 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 crisis, and has now started to increase again albeit at lower levels than before.

On the back of this evidence we would suggest that concerns about crowding in the overall quant space are less than what they might have been pre-crisis. Having said that, at the individual factor level there are some reasons for concern. The Value factor in particular is showing a level of crowdedness that is on par with pre-crisis levels. We think the use of real-time securities lending data is a good way to monitor these trends.

Long-run crowdedness

It is also instructive to step back and look at the bigger picture. While our daily Utilization data only goes back to 2006, it is possible to use monthly, exchange-provided short interest data to conduct an analysis that goes back to 1987. This data suffers from two weaknesses that we have already discussed: (1) it has a two week lag because it is snapped mid-month but is only available towards the end of the same month; and (2) short interest is not a perfect measure of the true supply-demand picture in each stock. Nevertheless, since we are mainly interested in the long-term trend over a 25 year period, we will ignore these two shortcomings temporarily.

Figure 26 shows the difference in short interest for Q10 (unattractive) versus Q1 (attractive) stocks, for Value and Momentum since 1987. Figure 27 shows the corresponding t-statistics.
We find a similar story over the long-run: there was a steady increase in crowding through the bull market leading up to the crisis.

Over this long horizon, the steady increase in crowdedness from the early 2000s to 2008 is readily apparent. In fact, on this expanded time scale it is clear that crowdedness, while well below pre-crisis levels, is still a lot higher than it has been in the long-run for both factors. The rise in crowdedness has largely mirrored a general increase in the level of shorting in the market overall, as seen in Figure 28 and Figure 29. Again it is worth reiterating that our regression framework already controls for a market-wide increase in shorting (via the intercept term) so these charts are saying that as shorting activity increased over time, this shorting was disproportionately directed into unattractive stocks, as defined by the Value and Momentum factors.

Can we link crowdedness to factor performance, or vice versa?

Crowding and factor performance

So far we have focused on devising a proxy for measuring crowdedness. But that is only the first step. Once we can measure crowdedness, the real question is whether crowdedness actually impacts factor performance.

In Figure 30 and Figure 31 we overlay a 12-month rolling average of our long-term Value and Momentum crowdedness metric with a 12-month rolling average of each factor’s performance. In both cases there is a negative correlation between crowdedness and contemporaneous factor performance. This would seem to support the argument that crowding can be detrimental to factor performance on average, particularly for Value.
As a more rigorous test, we conduct a time-series regression analysis where we regress monthly factor performance onto lagged (and leading) values of factor crowdedness. Figure 32 and Figure 33 show the results for Value and Momentum respectively. In the case of Value, we find little compelling evidence to suggest a significant relationship between past crowdedness and current returns (i.e. lags -6 to -1), or current returns and future crowdedness (i.e. leads 1 to 6). For Momentum, there is a statistically significant positive relationship between crowdedness from five months ago and this month’s factor performance. At the same time, there is a negative relationship between this month’s factor return and crowdedness five months from now.

As a robustness check, we also repeat this analysis for the month-on-month change in crowdedness – see Figure 34 and Figure 35. As before, we find little evidence of a lead/lag relationship for Value. For Momentum we see hints that past increases in crowdedness, at about a six month lag, are associated with positive factor performance today.

As in further unreported results, we also repeat our analysis using the monthly decile spread return instead of the IC to measure factor performance. We find qualitatively similar results, and hence only report our IC analysis.
Momentum seems to benefit in the short run from higher crowdedness, which is somewhat intuitive

This suggests that for Momentum at least, crowdedness in the short term is actually not a bad thing – if you are short past-loser stocks, you want others to come in and short them too, to help drive them down further. This reconciles with our past research on securities lending factors, where we found that in the short term one wants to “follow the shorts” because heavily shorted stocks on average tend to underperform in the following month.9 So if unattractive stocks are heavily shorted relative to attractive stocks (i.e. the factor is more crowded, based on our definition) then this is actually a good sign in the short term.

While these results are far from conclusive, we do think they gel with intuition. In the long run we find evidence that crowding is bad; witness the negative correlations between 12-month factor performance and crowdedness in Figure 28 and Figure 29. But for some factors – Momentum in particular – crowding can actually be beneficial in the short term. In other words, in the long run if everyone is chasing the same signal the mispricing will tend to get arbitrated away, but in the short term running with the crowd can be helpful.

The results in risk space are more clear-cut

We find that higher crowding is positively correlated with larger drawdowns and higher factor volatility

The downside to factor crowding

Of course, factor performance is not determined by the mean alone, and in our recent research we have been paying particular attention to the tail risk of factors.10 Our analysis in the previous section seems to indicate that crowdedness can have a confounding impact on factor returns – negative in the long run, but possibly positive in the short run. But what about factor risk? To answer this question, we plot the maximum drawdown (measured as most negative daily IC) for the six factors from our multifactor analysis as a function of the average absolute crowdedness of each factor (Figure 36). For crowdedness we use our daily Utilization-based measure because the higher frequency daily data gives us more data points to estimate factor risk with. As the chart shows, we find a strong negative relationship, i.e. factors with higher average crowdedness tend to have larger drawdowns.

We also plot the same chart for factor volatility (measured as the time-series standard deviation of daily factor ICs), as shown in Figure 37. Here we see a similar story – factors with higher crowdedness also tend to have more volatile returns. Admittedly our sample size is small – only six factors – but these results are consistent with the view that crowding can lead to large drawdowns when investors try to exit a common trade at the same time.

Month-end effects

An oft-cited hypothesis is that most quants rebalance around month-end, and hence there is a crowding effect in the first few days of each month. Anecdotally, we don’t believe this is necessarily true anymore – many quant managers have moved to a more continuous rebalancing process, where their portfolios are rebalanced opportunistically when the expected alpha exceeds the expected cost of the rebalance, rather than waiting for a fixed date. Nevertheless, our crowdedness proxy gives us a way to examine this phenomenon. In Figure 38 and Figure 39 we show event studies in the +/- 8 trading days around the last business day of each month (day 0).

Interestingly the results are opposite for the two factors. For Value, crowdedness seems to be a little higher in the trading days immediately after month-end, while for momentum crowdedness appears to peak just before month-end. Note though that the economic significance of these results is marginal; in the case of Value the difference between the pre- and post-month-end is only 12 bps and for Momentum it is -38 bps. In percent of a percent terms, this represents a 2% increase and a -5% decrease from pre-month-end levels. Recall this coefficient represents the difference in Utilization between unattractive and attractive stocks, so this is just saying this difference widens by 2% after month-end for Value, and declines by -5% for growth. These are fairly moderate in magnitude relative to the general volatility in the crowdedness time series.
The tale of the tape

What about high frequency data?

The second potential proxy for factor crowdedness that we consider is intraday data. In our report “Frequency arbitrage” we showed that high frequency data can be useful even for lower-frequency quantitative investors. Here we borrow some of the ideas in that paper as potential proxies for factor crowdedness.

Intraday order imbalance

Intraday order imbalance is designed to measure, over the course of the day, whether more orders are buyer initiated or seller initiated. Of course, directly observing which orders are buyer or seller initiated is not possible, so one must devise a statistical algorithm for guessing which side trades where generated by. In the academic literature the Lee & Ready algorithm is frequently used, however in our research we tend to favor the computationally more efficient Tick Test (see “Frequency arbitrage” for more detailed analysis on the pros and cons of each technique).

Once we have Order Imbalance for each stock on each day, our methodology is identical to that described in the previous section; we use equation (1) except we use Order Imbalance as the dependent variable instead of Utilization. Note that because Order Imbalance is a lot more volatile compared to Utilization (which doesn’t change much from day to day) we smooth the lines in the charts below using a 42-day moving average. Also note that we invert the y-axis to keep the directionality consistent with the previous section (we expect unattractive stocks to have lower, or negative, imbalance on average whereas for Utilization we expect unattractive stocks to have higher Utilization).

If we overlay the charts with the Utilization results from the previous section, it is easier to see the similarities and differences (Figure 42 and Figure 43). Broadly speaking, we get somewhat similar patterns over time.

We find the story is similar to what we found with securities lending data.

PIN is potentially a non-directional crowdedness metric.

Overall, given the results are roughly in line with what we found using securities lending data, we would tend to prefer the latter, for two reasons. First, Order Imbalance is much more volatile, so one a day to day basis it is hard to spot the trend, unless once applies some sort of moving average (which has the downside of introducing stale data). Second, intraday data is hard to store and manipulate, which means we always prefer a lower-frequency data set unless there is some compelling reason to go to a higher frequency.

What about informed trading?

Another factor of interest that derives from intraday data is Probability of Informed Trading, or PIN. Again, see our previous work for more details on this factor. The basic idea is that PIN tries to infer the proportion of a stock’s trading that is being driven by informed traders versus passive liquidity providers. In simple terms, one can think of informed traders as those trading with a strong view on a particular stock. In the context of factor crowding, it is easy to see how this might be a useful concept. If stocks with high or low factor scores have a higher PIN score, this might indicate that investors have a strong view on those stocks which in turn might suggest crowding.

Overall, securities lending data and intraday data yield similar results: for two reasons. First, while Order Imbalance is volatile, it is hard to spot the trend day to day unless a moving average is applied. Second, intraday data is hard to store and manipulate, so we prefer lower-frequency data unless there is a compelling reason to go to a higher frequency.

What about informed trading?

Another factor of interest derived from intraday data is Probability of Informed Trading (PIN). Again, see our previous work for more details on this factor. The basic idea is that PIN tries to infer the proportion of a stock’s trading that is being driven by informed traders versus passive liquidity providers. In simple terms, one can think of informed traders as those trading with a strong view on a particular stock. In the context of factor crowding, it is easy to see how this might be a useful concept. If stocks with high or low factor scores have a higher PIN score, this might indicate that investors have a strong view on those stocks which in turn might suggest crowding.

Figure 44 and Figure 45 show the results if we rerun our analysis with PIN as the dependent variable, compared to the previous results we obtained with Utilization. Note that PIN is non-directional, so instead of setting dummy variables for Q10 to Q2, we set our dummies to capture the difference in PIN between Q10 & Q1 compared to Q5 & Q6.
We find PIN gives us a similar picture for Value, but is somewhat different for Momentum.

Again, the results are actually surprisingly close to what we found with Utilization. The most noticeable difference is a divergence in Momentum crowdedness around July 2009. At that point in time, the PIN metric said Momentum was crowded, whereas the Utilization metric said it was uncrowded. On reflection, this divergence is partly because of the fact that PIN is non-directional. If investors have a strong view about stocks with extreme factor scores, this will show up as a positive coefficient. However, this coefficient doesn’t tell us the direction of those views.

The fact PIN is non-directional makes it less useful.

Therefore, similar to Order Imbalance, we argue that PIN, while interesting, is an inferior crowdedness proxy compared to Utilization. For most quants, the direction that the crowd is moving in is crucial, and this is something we don’t get when using PIN.
Institutional ownership

Can institutional ownership proxy for crowdedness

On face value, using the holdings data reported directly by money managers (via 13F and other regulatory filings) would seem to be the most obvious way to measure crowdedness. If most managers hold deep value stocks then we might infer Value factor is becoming crowded. Unfortunately, the biggest problem with ownership data is the lag between when a fund actually holds a position and when they have to report that position (roughly 2 months). This means that any information gleaned from ownership data will be somewhat backwards looking.

Nonetheless, it is certainly worth investigating. We use the Thomson Reuters ownership database as our source for fund holdings. Using this data, we compute the percent of institutional ownership for each stock on each day, using the most recently reported regulatory filings for each stock. We then use this Institutional Ownership metric as our dependent variable in the regression given by equation (1). The results are shown in Figure 46 and Figure 47. Again we reverse the y-axis for consistency with Utilization; we expect decile 10 stocks (the worst stocks) to have the lowest institutional ownership.

In Figure 48 and Figure 49 we compare the results for Value and Momentum to those obtained using Utilization. It turns out the charts match to extent. Having said that if we look at Value currently, we can see some divergence – the Utilization-based metric says crowdedness is rising, while the Institutional Ownership-based metric says it is declining. Because of the lag in the institutional holdings data, we still prefer Utilization as our primary means of determining factor crowdedness. However, the close match between the charts does give us comfort that we are on the right track.

We find results that are largely consistent with those from Utilization, despite the lag inherent in the ownership data.
Figure 48: Incremental Utilization versus incremental Institutional Ownership for Q10 Value stocks

Figure 49: Incremental Utilization versus incremental Institutional Ownership for Q10 Momentum stocks

Source: Bloomberg Finance LP, Compustat, DataExplorers, IBES, Russell, S&P, Thomson Reuters, Deutsche Bank
Quantifying a crisis

What happened in the summer of 2007?

If we are looking for a test-bed to evaluate whether our factor crowding proxy actually measures crowdedness, we need to look no further than the turmoil of August 2007. It has been well documented that this so-called quant crisis was precipitated and then fueled by a sharp deleveraging scenario – see Khandani and Lo [2008] for an excellent summary.

In Figure 50 to Figure 52 we show how our Utilization-based crowdedness proxies evolved on a daily basis over July and August 2007, for the six factors we considered in the previous sections. Figure 53 shows the Value and Momentum results using an alternative metric – Order Imbalance.

The most notable results are for Value, Momentum, and Reversal. There was a sharp decrease in Value crowdedness (i.e. investors suddenly stopped shorting expensive stocks) between the 6th and 13th of August. Momentum showed a slightly more moderate decline over that window. Interestingly, 1-month Reversal had a sharp increase in crowdedness.
around the same time. Keep in mind the directionality here; this is saying that stocks that did well last month were suddenly being shorted more heavily (since in a reversal strategy we buy stocks that did poorly last month and short those that did well).

We think all three results are consistent with the deleveraging story. First, the five-day fall in Value and Momentum crowdedness was highly significant. For Value (Figure 54) the decrease in the level of shorting of expensive stocks was a 4 standard deviation event. For momentum, it was less dramatic, but still a 1.4 standard deviation event (Figure 55). In the case of Reversal, the fact that investors rapidly turned negative on short-term winners also makes sense – if you are faced with a liquidity crunch, it is natural to try to take profits on your winners first. As an aside, the results for our alternative Order Imbalance proxy (Figure 53) clearly illustrate the drawback of this metric – it is much too volatile from day to day to spot short-term movements.

The decline in crowdedness corresponded to the major factor drawdowns over the worse days in August

Next, we overlay our crowdedness proxies for Value (Figure 56) and Momentum (Figure 57) with daily factor performance (the bars in the charts). Now we can clearly see the chain of events in terms of performance and crowding. The sequence of poor Value returns in July and early August came to a head in the second week of August when investors savagely delevered their Value strategies, exacerbating the drawdown in Value performance.

We find a large decrease in factor crowding for Value and Momentum over the most intense days of the crisis
The Momentum chart shows a similar spiral of drawdowns, deleveraging, and more drawdowns. Having said that, our results are also consistent with the existing literature that argues that in fact most strategies bounced back quite quickly. It is often said that those managers who held their nerve over August and maintained their positions were largely unaffected by what turned out to be a temporary deleveraging episode. Indeed, we can see from the charts above that crowdedness, at least based on our proxies, was largely back to where it started by the end of August. The real fall in crowdedness didn’t come until much later, when the credit crisis exploded in the following year.
Innovate to survive?

Are new data sources really the answer?

A final question we want to address in this paper is whether new, innovative factors really help alleviate crowding. Having devised a proxy for crowdedness and convinced ourselves it has some merit by tracking it through the 2007 quant crisis, we now have the tools to more directly address this question. Overall the results are a little surprising.

In the following charts, we take some of what we consider to be our unique factors – our Composite Options Factor\(^\text{12}\), RPIN\(^\text{13}\), Adjusted Bond Momentum\(^\text{14}\), our Securities Lending Factor (DBSLX)\(^\text{15}\), and even our QCD Model\(^\text{16}\) – and run our crowdedness analysis. In each case, we also include Value and Momentum and the control variables Size and Volatility in the regression, in addition to the factor in question. This is to make sure we are truly measuring crowdedness against the new factor, and not just picking up on a time-varying loading on the more traditional factors.

For example, Figure 58 shows the results for our Composite Options Factor.

Figure 58: Incremental Utilization for Q10 stocks versus Q1 stocks (%), Value, Momentum, and DB Composite Options Factor

Similarly, Figure 59 shows the results for RPIN and Figure 60 shows Adjusted Bond Momentum.

---


From these charts it is difficult to draw any firm conclusions on whether these factors do indeed show less crowding than the basic factors like Value and Momentum. There are hints that some factors, like Adjusted Bond Momentum for example, have a consistently lower level of crowdedness compared to Value and Momentum. But it is difficult to generalize this given the short time-series and the limited number of factors we are considering.

There is also another important point to consider. These charts say nothing of payoff from incurring a unit of crowdedness. As an illustration, consider Figure 61. This shows the results for a factor we call DBSLX, which is designed to buy lightly shorted stocks and short stocks.
that are already heavily shorted, after adjusting for the cost of borrowing such stocks. On face value the chart suggests that the factor is horrifically crowded. But as we showed in our report, this factor actually generates attractive returns above and beyond the cost of implementation, notwithstanding the fact it actually seeks out crowding (i.e. it wants to short stocks that everyone is already heavily short).

One problem is that crowdedness needs to be considered in the context of the expected payoff or drawdown from incurring that crowdedness.

Another good example of why we can’t automatically conclude that crowded equals bad is Figure 62. Here we consider our QCD model. Through most of the financial crisis, our model would appear to be quite crowded relative to Value and Momentum. But again, the performance of our model suggests this position was more than justified in return space.
Disentangling the results

A little further analysis is warranted to better understand the dynamics of using shorting as a crowdedness proxy. As mentioned, our past research has demonstrated that there is a premium to shorting stocks that are already heavily shorted, if one is careful to adjust for the cost of taking those short positions. In other words, one wants to short stocks where the expected alpha (i.e. further underperformance) is greater than the (often high) cost of shorting stocks that are already heavily shorted. Our DBSLX factor is designed to do exactly that.

This raises the important question that we touched on before. Perhaps the crowdedness we are observing is justified by the positive payoff that investors are getting from shorting stocks that are already heavily shorted. We seek to better understand this issue in two ways. First, in Figure 63 we conduct the following regression:

\[
Utilization_{i,t} = c + \sum_{j=1}^{J} \sum_{q=2}^{Q} \beta_{j,t,j,q} D_{j,t,j,q} + \sum_{q=2}^{Q} \beta_{j,t,DDBSLX_{,q}}D_{j,t,DDBSLX_{,q}} + \sum_{q=2}^{Q} \beta_{j,t,\sigma,q} D_{j,t,\sigma,q} + \epsilon_{i,t}
\]  

(2)

where all the terms are the same as in equation (1) on page 5, except for the addition of a new set of dummy variables, \(D_{j,t,DDBSLX_{,q}}\), which capture the deciles of the DBSLX factor. Second, in Figure 64 we conduct the following regression:

\[
-DBSLX_{i,t} = c + \sum_{j=1}^{J} \sum_{q=2}^{Q} \beta_{j,t,j,q} D_{j,t,j,q} - \sum_{q=2}^{Q} \beta_{j,t,\sigma,q} D_{j,t,\sigma,q} + \sum_{q=2}^{Q} \beta_{j,t,DDBSLX_{,q}}D_{j,t,DDBSLX_{,q}} + \epsilon_{i,t}
\]  

(3)

where again all terms are the same as in equation (1), except now the dependent variable, \(DBSLX_{i,t}\), is our DBSLX factor. We place a negative sign in front because our DBSLX factor is sorted such that a lower score corresponds to a stock that is more heavily shorted. This makes the direction of our results consistent with Utilization.

What do these results tell us? First consider Figure 63, which shows the t-statistics for the crowdedness coefficients for Value and Momentum from equation (2). In this regression, we...
are effectively adding in the DBSLX factor as a control variable. The idea is to see if Value and Momentum still show significant crowding, after controlling for what we might call “good shorting”. If we find the difference in shorting between attractive and unattractive Value and Momentum stocks is no longer significant, then this tells us that these two variables are actually just proxies for DBSLX. Think about it this way: if expensive stocks (we’ll use Value as an example) are heavily shorted, some of this shorting might be investors trying to short “good shorts”, i.e. shorts they expect to go down by more than the cost of taking the position. We can argue that such shorts don’t really represent crowdedness in Value per se, since they are not being taken by investors who are chasing Value strategies. However, the results in Figure 63 show us that the statistical significant of Value and Momentum crowdedness has not changed that much with the addition of DBSLX as a control variable (c.f. Figure 2 on page 8). This tells us that there is actually not much overlap between “good shorts”, as represented by DBSLX, and Value and Momentum. In other words, investors do seem to be shorting Value and Momentum because they have a view that expensive stocks will go down and past-losers will go down, and not because these baskets happen to contain stocks that have attractive downside relative to the cost of shorting them.

We can also look at this from the other direction, which is that implied by equation (3). Here we are looking at whether Value and Momentum have any exposure to “good shorting”, again measured using our DBSLX factor. Based on our hypothesis above, we would expect to find insignificant results with this regression. This appears to be somewhat borne out. Momentum crowdedness becomes insignificant most of the time, whereas Value skirts the edges of significance for most of the history.

Overall this gives us some comfort that our crowdedness proxy is indeed picking up on investors who are specifically chasing a given strategy, rather than those who are seeking to opportunistically find heavily shorted stocks that have more expected downside than justified by their cost of borrow.

It turns out that crowdedness is not just an artifact of investors seeking “good shorts”.

We again find evidence that higher crowding is associated with larger drawdowns.

Does crowding increase risk?

We saw previously (see Figure 36 and Figure 37 on page 20) that factors with higher crowding do seem to have larger downside risk and greater factor volatility. In Figure 65 and Figure 66 we repeat this analysis for the DB proprietary factors that we have considered. In terms of downside risk, we still see a negative (albeit weaker) relationship between absolute crowdedness and maximum drawdown (Figure 65). For volatility, there is little relationship between crowdedness and factor volatility (Figure 66).
What can we make of all these results? We think the key is to recognize the difference between “normal” factor performance and the downside risk to factors. On average, we have seen that crowding does not automatically mean a factor will underperform. In fact, in the short term we showed that crowding can actually help some factors, assuming you are not the last one into a particular trade. A case in point was our “crowd-seeking” DBSLX factor, which actually looks to jump in and short the most heavily shorted stocks, after adjusting for the cost of borrow. A more common example is the momentum factor.

However, at a more strategic level, paying attention to crowding makes sense, because crowdedness can have an impact on the size of the drawdown in a deleveraging scenario or indeed when investor sentiment shifts (say from risk-seeking to risk-averse). In each case, the impact of the collective unwinding of a factor tilt is going to be stronger in a more crowded factor. For example, we saw in this report that our DBSLX factor is more exposed to a sharp deleveraging compared to, say, our less crowded RPIN factor. Our analysis of the quant crisis in the summer of 2007 offers a useful lesson on how quickly such deleveraging can occur.

The bottom line is that crowdedness is another metric that should be considered in the factor selection and weighting decision. In choosing and weighting factors, quantitative investors consider a plethora of metrics – expected risk, expected return, turnover, alpha decay, higher moments, to name but a few. We argue they need to add crowdedness to the toolbox. We hope this report has offered some useful ideas on how to do this.
References


Appendix 1

Important Disclosures
Additional information available upon request

For disclosures pertaining to recommendations or estimates made on a security mentioned in this report, please see the most recently published company report or visit our global disclosure look-up page on our website at http://gm.db.com/qer/disclosure/DisclosureDirectory.eqsr.

Analyst Certification

The views expressed in this report accurately reflect the personal views of the undersigned lead analyst(s). In addition, the undersigned lead analyst(s) has not and will not receive any compensation for providing a specific recommendation or view in this report. Rochester Cahan/Yin Luo/Miguel-Alvarez/Javed Jussa/Zongye Chen/Sheng Wang

Hypothetical Disclaimer

Backtested, hypothetical or simulated performance results have inherent limitations. Unlike an actual performance record based on trading actual client portfolios, simulated results are achieved by means of the retroactive application of a backtested model itself designed with the benefit of hindsight. Taking into account historical events the backtesting of performance also differs from actual account performance because an actual investment strategy may be adjusted any time, for any reason, including a response to material, economic or market factors. The backtested performance includes hypothetical results that do not reflect the reinvestment of dividends and other earnings or the deduction of advisory fees, brokerage or other commissions, and any other expenses that a client would have paid or actually paid. No representation is made that any trading strategy or account will or is likely to achieve profits or losses similar to those shown. Alternative modeling techniques or assumptions might produce significantly different results and prove to be more appropriate. Past hypothetical backtest results are neither an indicator nor guarantee of future returns. Actual results will vary, perhaps materially, from the analysis.
Regulatory Disclosures

1. Important Additional Conflict Disclosures

Aside from within this report, important conflict disclosures can also be found at https://gm.db.com/equities under the "Disclosures Lookup" and "Legal" tabs. Investors are strongly encouraged to review this information before investing.

2. Short-Term Trade Ideas

Deutsche Bank equity research analysts sometimes have shorter-term trade ideas (known as SOLAR ideas) that are consistent or inconsistent with Deutsche Bank’s existing longer term ratings. These trade ideas can be found at the SOLAR link at http://gm.db.com.

3. Country-Specific Disclosures

Australia & New Zealand: This research, and any access to it, is intended only for "wholesale clients" within the meaning of the Australian Corporations Act and New Zealand Financial Advisors Act respectively.

EU countries: Disclosures relating to our obligations under MiFiD can be found at http://www.globalmarkets.db.com/riskdisclosures.

Japan: Disclosures under the Financial Instruments and Exchange Law: Company name - Deutsche Securities Inc. Registration number - Registered as a financial instruments dealer by the Head of the Kanto Local Finance Bureau (Kinsho) No. 117. Member of associations: JSDA, Type II Financial Instruments Firms Association, The Financial Futures Association of Japan, Japan Securities Investment Advisers Association. Commissions and risks involved in stock transactions - for stock transactions, we charge stock commissions and consumption tax by multiplying the transaction amount by the commission rate agreed with each customer. Stock transactions can lead to losses as a result of share price fluctuations and other factors. Transactions in foreign stocks can lead to additional losses stemming from foreign exchange fluctuations. "Moody’s", "Standard & Poor’s", and "Fitch" mentioned in this report are not registered credit rating agencies in Japan unless "Japan" is specifically designated in the name of the entity.

Russia: This information, interpretation and opinions submitted herein are not in the context of, and do not constitute, any appraisal or evaluation activity requiring a license in the Russian Federation.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>60 Wall Street</td>
<td>One International Place</td>
<td>222 South Riverside Plaza</td>
<td>101 California Street</td>
</tr>
<tr>
<td>New York, NY 10005</td>
<td>12th Floor</td>
<td>30th Floor</td>
<td>46th Floor</td>
</tr>
<tr>
<td>Tel: (212) 250 2500</td>
<td>Boston, MA 02110</td>
<td>Chicago, IL 60606</td>
<td>Tel: (415) 617 2800</td>
</tr>
<tr>
<td>101 California Street</td>
<td>700 Louisiana Street</td>
<td>700 Louisiana Street</td>
<td>1 Great Winchester Street</td>
</tr>
<tr>
<td>46th Floor</td>
<td>Houston, TX 77002</td>
<td>Houston, TX 77002</td>
<td>London EC2N 2E0</td>
</tr>
<tr>
<td>San Francisco, CA 94111</td>
<td>Tel: (832) 229-4500</td>
<td>Tel: (832) 229-4500</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>Tel: (415) 617 2800</td>
<td></td>
<td></td>
<td>Tel: (44) 20 7545 8000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>International Locations</th>
<th>Deutsche Bank AG London</th>
<th>Deutsche Bank AG</th>
<th>Deutsche Bank AG</th>
</tr>
</thead>
<tbody>
<tr>
<td>New York, NY 10005</td>
<td>1 Great Winchester Street</td>
<td>2-11-1 Nagatacho</td>
<td>Deutsche Bank Securities Inc.</td>
</tr>
<tr>
<td>United States of America</td>
<td>London EC2N 2E0</td>
<td>Sanno Park Tower</td>
<td>101 California Street</td>
</tr>
<tr>
<td>Tel: (212) 250 2500</td>
<td>United Kingdom</td>
<td>Chiyoda-ku, Tokyo 100-6171</td>
<td>46th Floor</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Japan</td>
<td>Tel: (41) 3 5156 6770</td>
</tr>
<tr>
<td>Deutsche Bank AG</td>
<td>Deutsche Bank AG</td>
<td>Deutsche Securities Inc.</td>
<td>Deutsche Bank AG</td>
</tr>
<tr>
<td>Filiale Hongkong</td>
<td>Deutsche Bank AG</td>
<td>2-11-1 Nagatacho</td>
<td>Deutsche Bank Place</td>
</tr>
<tr>
<td>International Commerce Centre,</td>
<td>Deutsche Bank AG</td>
<td>Sanno Park Tower</td>
<td>Level 16</td>
</tr>
<tr>
<td>1 Austin Road West,Kowloon,</td>
<td>Deutsche Bank Securities Inc.</td>
<td>Chiyoda-ku, Tokyo 100-6171</td>
<td>Corner of Hunter &amp; Phillip Streets</td>
</tr>
<tr>
<td>Hong Kong</td>
<td>2-11-1 Nagatacho</td>
<td>Japan</td>
<td>Sydney, NSW 2000</td>
</tr>
<tr>
<td>Tel: (852) 2203 8888</td>
<td>Sanno Park Tower</td>
<td>Tel: (81) 3 5156 6770</td>
<td>Australia</td>
</tr>
<tr>
<td></td>
<td>Chiyoda-ku, Tokyo 100-6171</td>
<td></td>
<td>Tel: (61) 2 8258 1234</td>
</tr>
<tr>
<td></td>
<td>Japan</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Disclaimer

The information and opinions in this report were prepared by Deutsche Bank AG or one of its affiliates (collectively "Deutsche Bank"). The information and opinions are believed to be reliable and has been obtained from public sources believed to be reliable. Deutsche Bank makes no representation as to the accuracy or completeness of such information.

Deutsche Bank may engage in securities transactions, on a proprietary basis or otherwise, in a manner inconsistent with the view taken in this research report. In addition, others within Deutsche Bank, including strategists and sales staff, may take a view that is inconsistent with that taken in this research report.

Deutsche Bank may be an issuer, advisor, manager, distributor or administrator of, or provide other services to, an ETF included in this report, for which it receives compensation.

Opinions, estimates and projections in this report constitute the current judgement of the author as of the date of this report. They do not necessarily reflect the opinions of Deutsche Bank and are subject to change without notice. Deutsche Bank has no obligation to update, modify or amend this report or to otherwise notify a recipient thereof in the event that an opinion, forecast or estimate set forth therein, changes or subsequently becomes inaccurate. Prices and availability of financial instruments are subject to change without notice. This report is provided for informational purposes only. It is not an offer or a solicitation of an offer to buy or sell any financial instruments or to participate in any particular trading strategy. Target prices are inherently imprecise and a product of the analyst judgement.

As a result of Deutsche Bank’s March 2010 acquisition of BHF-Bank AG, a security may be covered by more than one analyst within the Deutsche Bank group. Each of these analysts may use differing methodologies to value the security; as a result, the recommendations may differ and the price targets and estimates of each may vary widely.

In August 2009, Deutsche Bank instituted a new policy whereby analysts may choose not to set or maintain a target price of certain issuers under coverage with a Hold rating. In particular, this will typically occur for "Hold" rated stocks having a market cap smaller than most other companies in its sector or region. We believe that such policy will allow us to make best use of our resources. Please visit our website at http://gm.db.com to determine the target price of any stock.

The financial instruments discussed in this report may not be suitable for all investors and investors must make their own informed investment decisions. Stock transactions can lead to losses as a result of price fluctuations and other factors. If a financial instrument is denominated in a currency other than an investor’s currency, a change in exchange rates may adversely affect the investment.

All prices are those current at the end of the previous trading session unless otherwise indicated. Prices are sourced from local exchanges via Reuters, Bloomberg and other vendors. Data is sourced from Deutsche Bank and subject companies.

Past performance is not necessarily indicative of future results. Deutsche Bank may with respect to securities covered by this report, sell to or buy from customers on a principal basis, and consider this report in deciding to trade on a proprietary basis.

Derivative transactions involve numerous risks including, among others, market, counterparty default and illiquidity risk. The appropriateness or otherwise of these products for use by investors is dependent on the investors' own circumstances including their tax position, their regulatory environment and the nature of their other assets and liabilities and as such investors should take expert legal and financial advice before entering into any transaction similar to or inspired by the contents of this publication. Trading in options involves risk and is not suitable for all investors. Prior to buying or selling an option investors must review the "Characteristics and Risks of Standardized Options," at http://www.theocc.com/components/docs/riskstoc.pdf. If you are unable to access the website please contact Deutsche Bank AG at +1 (212) 250-7994, for a copy of this important document.

The risk of loss in futures trading, foreign or domestic, can be substantial. As a result of the high degree of leverage obtainable in futures trading, losses may be incurred that are greater than the amount of funds initially deposited.

Unless governing law provides otherwise, all transactions should be executed through the Deutsche Bank entity in the investor's home jurisdiction. In the U.S. this report is approved and/or distributed by Deutsche Bank Securities Inc., a member of the NYSE, the NASD, NFA and SIPC. In Germany this report is approved and/or communicated by Deutsche Bank AG Frankfurt authorized by the BaFin. In the United Kingdom this report is approved and/or communicated by Deutsche Bank AG London, a member of the London Stock Exchange and regulated by the Financial Services Authority for the conduct of investment business in the UK and authorized by the BaFin. This report is distributed in Hong Kong by Deutsche Bank AG, Hong Kong Branch, in Korea by Deutsche Securities Korea Co. This report is distributed in Singapore by Deutsche Bank AG, Singapore Branch, and recipients in Singapore of this report are to contact Deutsche Bank AG, Singapore Branch in respect of any matters arising from, or in connection with, this report. Where this report is issued or promulgated in Singapore to a person who is not an accredited investor, expert investor or institutional investor (as defined in the applicable Singapore laws and regulations), Deutsche Bank AG, Singapore Branch accepts legal responsibility to such person for the contents of this report. In Japan this report is approved and/or distributed by Deutsche Securities Inc. The information contained in this report does not constitute the provision of investment advice. In Australia, retail clients should obtain a copy of a Product Disclosure Statement (PDS) relating to any financial product referred to in this report and consider the PDS before making any decision about whether to acquire the product. Deutsche Bank AG Johannesburg is incorporated in the Federal Republic of Germany (Branch Register Number in South Africa: 1998/003298/10). Additional information relating to securities, other financial products or issuers discussed in this report is available upon request. This report may not be reproduced, distributed or published by any person for any purpose without Deutsche Bank’s prior written consent. Please cite source when quoting. Copyright © 2012 Deutsche Bank AG