LINE OF SIGHT

UNDERSTANDING FACTOR TILTS

STYLE FACTOR TILTS CAN PLAY A VALUABLE ROLE IN PORTFOLIOS

June 2013

Although their efficacy conflicts with modern financial theory, investors have successfully employed style factor tilts for more than 40 years to improve on passive capitalization weighted equity portfolios. Empirical studies of popular style tilts like value, small size and momentum repeatedly have been shown to outperform benchmarks across most global markets. Because these results are inconsistent with the classic notion that return is a function of risk alone, they are considered anomalous.

To better understand the role factor tilts can play in a portfolio, it is helpful to understand their origin and historical track record. From here, we will examine the theoretical justification for their effectiveness and the consistency of their returns through time. In doing so, we have expanded the traditional scope of style factors to include volatility, quality and dividend yield, among others.

Our research has shown that many factor tilts have indeed outperformed their benchmarks in domestic, international and even emerging markets. However, the excess return generated by "pure" single-factor tilts has, with few exceptions, not been entirely consistent across time. In contrast, many multi-factor tilts (high quality and low volatility,

Northern Trust

for example) are shown to both outperform and have persistent excess return. This is why we recommend that investors seek out products with multi-factor exposure to minimize the possibility of extended periods of negative performance.

THE CENTRAL PARADIGM

The first equity factor model was introduced by William Sharpe in 1964.[1] His model included only a single factor and, as a result, was quite straightforward:

$$ER_i = \alpha_i + \beta_i(ER_M)$$

In this equation, ER_i is the excess return on stock i over the risk-free rate of interest, ER_M is the return on the market, β_i is the exposure of stock i to the market excess return and α_i is the return on stock i not explained by the market excess return. This is, of course, the hallmark capital asset pricing model (CAPM).

The single factor in this model is β_i (beta), or the sensitivity of the stock to the market excess return. According to CAPM, a stock with a beta greater than one should earn a higher excess return than the market, but also should have a higher volatility. Sharpe conducted empirical tests to show that roughly two-thirds of the excess return of any given stock can be explained by this beta. This is the genesis of the notion that return increases with risk or, more specifically, return increases with systematic market risk.

If we accept CAPM as the true market model, we must also immediately reject the possibility of other factor tilts. Nowhere in Sharpe's model does it say we should be compensated for anything other than systematic risk. If, for example, a tilt toward small caps or value stocks achieved a higher return than the benchmark it must, according to the model, be the result of higher systematic risk. While CAPM allows returns to vary across sub-segments of stocks, the return per unit of risk (or the Sharpe ratio) must be always constant. In this sense, a size or value factor tilt is meaningless because these tilts are tantamount to increasing risk. Under Sharpe's theory, the same results could have been achieved by simply increasing the portfolio's beta by tilting toward high-beta stocks.

This notion of a constant Sharpe ratio has been challenged repeatedly in financial literature. For example, it is very easy to demonstrate that return per unit of risk varies considerably with value. Exhibit 1 shows the average annual equally weighted

EXHIBIT 1: VALUATION QUINTILE RETURNS AND SHARPE RATIOS

Level	Value Quintile	Russell 3000 1979 – 2012	MSCI World ex US 1996 – 2012	MSCI EM IMI 2005 – 2012
Average Annualized Equally Weighted Returns				
Index		12.9%	7.9%	14.6%
Cheap	1	16.3%	13.0%	21.8%
	2	15.6%	10.0%	16.9%
	3	12.2%	6.0%	14.2%
	4	11.7%	5.9%	13.5%
Expensive	5	10.0%	4.6%	6.5%
Average Sharpe Ratios				
Index		0.77	0.43	0.55
Cheap	1	0.86	0.54	0.72
	2	0.68	0.53	0.61
	3	0.67	0.34	0.55
	4	0.59	0.33	0.55
Expensive	5	0.42	0.24	0.25

Source: Northern Trust Quantitative Research

returns and Sharpe ratios for three indexes: the Russell 3000 Index, the MSCI World ex US Index and the MSCI Emerging Market Investable Markets Index. Each month, stocks in these indexes are broken into value quintiles based on their exposure to the MSCI BARRA value factor, and average returns and Sharpe ratios are then calculated. Note that first quintile (cheap) stocks have both a higher average return and a higher Sharpe ratio than their respective indexes. So across all global markets, value stocks have earned significantly higher returns per unit of risk than their more expensive growth counterparts. In other words, tilting toward value stocks would have produced a higher average return at a lower average risk than investing in the equally weighted index – a result that directly conflicts with CAPM.

This is only one example of several well-documented CAPM inconsistencies. Others include the higher returns and Sharpe ratios associated with high momentum, small size and low volatility, to name a few. Since these findings are incongruent with the classical notion of return as a function of risk alone, they are termed *anomalies* in financial literature. The large, and growing, number of anomalies suggests that CAPM may not be the final word in equity factor models.

In essence, the concept of being paid strictly for taking risk systematically has been muted, and the notion that factor tilts toward small value stocks could increase returns was born.

ENTER FAMA, FRENCH AND CARHART

While evidence of anomalies began to pile up almost immediately after the publication of Sharpe's paper, it took almost 30 years for a serious competitor of CAPM to emerge. In 1992 Eugene Fama and Kenneth French introduced a three-factor model that had much better success in explaining historical stock returns than CAPM.[2] Although Fama and French's model included Sharpe's original beta factor, they address some of the more prevalent CAPM issues by including factors for size and value.

$$ER_i = \alpha_i + \beta_i(ER_M) + \lambda_i(\text{size}) + \phi_i(\text{value})$$

where λ_i and ϕ_i are sensitivities to the size and value factors, respectively

Note that since the expected excess return of a security is a function of variables other than just market excess return, we no longer have a direct relationship between risk and return.[3] In essence, the concept of being paid strictly for taking risk systematically has been muted, and the notion that factor tilts toward small value stocks could increase returns was born.

In 1997, Mark Carhart extended Fama and French's model to include a fourth factor – momentum.[4] Although the Fama-French three-factor model was able to explain more than 95% of the variation in any given stock's return, the addition of momentum was found to be statistically robust and to increase the predictive power even further. Perhaps more importantly, Carhart's model was able to explain the three most prevalent equity market anomalies – size, value and momentum – in one succinct package. The case for factor tilts had only improved.

Still, despite their success in explaining historical stock returns, the Fama, French and Carhart models face some theoretical difficulties. In particular, they do not address why small size, high value and high momentum stocks should warrant a premium. Unlike

CAPM, which provides a very intuitive justification for returns (high risk = high return), the connection between small size, high value and high returns is not quite so clear. Even Fama and French themselves question their model's theoretical justification. In a 1996 paper they write:

> *"[have] we simply found three [factors] that provide a parsimonious description of returns and average returns, and so can absorb most of the CAPM anomalies? In other words, without knowing why, have we stumbled on … the three factor model?"*[5]

Despite the importance of these questions, explanations for why factor tilts should offer premium returns are wanting. The most viable candidates may be:

- Fama and French – and many other authors – characterize the value premium as compensation for potential financial distress.[6] One theory suggests that investors avoid distressed stocks and, hence, earn a higher premium because events that cause losses in distressed stocks are negatively correlated with investor income.
- Brennan, Chordia, Subrahmanyam and Tong argue that the size premium is actually a liquidity effect in which small-cap stocks are less liquid than large-cap stocks and therefore provide correspondingly higher returns to offset higher transaction costs.[7]
- Lakonishok, Shleifer and Vishney explain momentum as the by-product of irrational market participants who under-react to news.[8] As a result, the directional movement of a stock's price persists for a period longer than otherwise anticipated.

The fact that many anomalies have persisted for decades suggests they are not evidence of market inefficiencies but rather that our existing factor models are less-than-complete descriptions of returns.

Given these somewhat lackluster explanations it is tempting to view the value, size and momentum effects as simply transitory market inefficiencies, the magnitude of which will dissipate as investors seek to exploit them and that CAPM is, in fact, the appropriate long-run model. However, the fact that these and many other anomalies have persisted for decades suggests they are not evidence of market inefficiencies but rather that our existing factor models are less-than-complete descriptions of returns. So, even if we lack full theoretical justification, we have no reason to denounce them in favor of the traditional CAPM model.

THE EMPIRICAL EVIDENCE

Value, size and momentum are only a few of dozens of anomalies described in the financial literature. More recently, Ang, Hodrick, Xing and Zhang as well as Baker and Haugen have examined the "low volatility phenomenon," which is simply a tilt towards low volatility stocks.[9] Dividend yield was first introduced as a factor tilt by Brennan and later confirmed by Litzenberger and Ramaswamy, while quality, in numerous forms, has been a factor since at least the early 1970s.[10]

To illustrate the historical performance of factor tilts we again use the MSCI BARRA factor definitions for value, size, momentum and volatility, as well as Northern Trust's proprietary quality factor across domestic, international and emerging markets. The exception is dividend yield, for which we use the MSCI BARRA definition in the United States and the Worldscope dividend yield for international and emerging markets, because MSCI BARRA does not publish a dividend yield factor outside of the United States.

Each month, we rank all stocks in their respective universes based on these individual factors and place them into quintiles. We then measure the subsequent one-month return for each quintile. Exhibit 2 details the average annualized returns for each quintile, the Sharpe ratio, and the average annualized return and Sharpe ratio for the respective equally weighted benchmark. Note that the number in the column under the index name represents the average for all quintiles.

The results are striking. In the United States, the average annual equally weighted return of the Russell 3000 from 1979 to 2012 was 12.9% with a Sharpe ratio of 0.77. However, during this period a tilt toward *any* of the factors we examined – value, size, momentum, volatility dividend yield or Northern Trust's quality factor – would have performed significantly better in terms of both absolute return and Sharpe ratio.

EXHIBIT 2: RETURNS AND SHARPE RATIOS BY FACTOR QUINTILE

Domestic: Russell 3000

		Returns 1979 to 2012							Sharpe Ratio 1979 to 2012						
		Russell 3000	Value	Size	Momentum	Volatility	Dividend Yield	NT Quality	Russell 3000	Value	Size	Momentum	Volatility	Dividend Yield	NT Quality
High	1		16.3%	11.4%	15.7%	6.0%	15.2%	19.0%		0.86	0.38	0.77	0.17	1.03	1.01
	2		15.6%	12.9%	14.5%	15.0%	13.1%	16.5%		0.68	0.52	0.84	0.61	0.82	0.91
	3	12.9%	12.2%	13.2%	13.0%	15.7%	13.9%	12.5%	0.77	0.67	0.57	0.68	0.83	0.65	0.65
	4		11.7%	14.2%	13.4%	14.8%	11.7%	11.6%		0.59	0.78	0.70	0.95	0.48	0.56
Low	5		10.0%	15.1%	9.2%	14.3%	10.7%	5.1%		0.42	0.88	0.31	1.19	0.26	0.21

International: MSCI World Ex US

		Returns 1996 to 2012							Sharpe Ratio 1996 to 2012						
		MSCI AW Ex US	Value	Size	Momentum	Volatility	Dividend Yield	NT Quality	MSCI AW Ex US	Value	Size	Momentum	Volatility	Dividend Yield	NT Quality
High	1		13.0%	6.6%	10.8%	5.1%	13.7%	11.9%		0.54	0.36	0.58	0.16	0.64	0.69
	2		10.0%	6.8%	9.4%	7.9%	11.6%	9.5%		0.53	0.37	0.56	0.37	0.64	0.54
	3	7.9%	6.0%	8.5%	8.4%	7.3%	8.5%	8.7%	0.43	0.34	0.47	0.49	0.40	0.50	0.50
	4		5.9%	7.3%	6.0%	9.3%	3.6%	6.5%		0.33	0.38	0.31	0.59	0.21	0.34
Low	5		4.6%	10.2%	4.7%	9.8%	3.1%	4.8%		0.24	0.45	0.16	0.79	0.15	0.24

Emerging Markets: MSCI EM IMI

		Returns 2005 to 2012							Sharpe Ratio 2005 to 2012						
		MSCI EM IMI	Value	Size	Momentum	Volatility	Dividend Yield	NT Quality	MSCI EM IMI	Value	Size	Momentum	Volatility	Dividend Yield	NT Quality
High	1		21.8%	14.5%	17.2%	14.9%	20.4%	17.4%		0.72	0.62	0.63	0.40	0.82	0.67
	2		16.9%	13.5%	15.9%	14.6%	17.1%	16.9%		0.61	0.52	0.64	0.49	0.67	0.64
	3	14.6%	14.2%	15.2%	15.0%	14.9%	18.2%	14.9%	0.55	0.55	0.57	0.59	0.56	0.71	0.55
	4		13.5%	15.9%	11.0%	16.1%	13.8%	15.1%		0.55	0.57	0.41	0.67	0.50	0.54
Low	5		6.5%	13.2%	14.3%	12.3%	8.3%	10.4%		0.25	0.44	0.45	0.69	0.27	0.33

Source: Northern Trust Quantitative Research

For example, consider a value tilt. The highest value, i.e. cheapest, quintile of stocks earned an average annual return of 16.3% with a Sharpe ratio of 0.86. While this seems like a modest improvement over the index, recall that this represents an average annual outperformance of 3.4% per year for 34 years. If you had invested $1 in the equally weighted Russell 3000 index on January 1, 1979, that investment would have been worth almost $62 at the end of December 2012. In contrast, if you instead invested exclusively in the highest value quintile your investment would have been worth approximately $170.

The most extreme difference in returns is the Northern Trust quality factor. Here investing $1 in the lowest quality quintile would generate a final value of just $5 versus the $62 earned by the index. At the opposite end of the spectrum, high quality stocks would have earned more than $370 for every $1 invested. The difference in cumulative return between high and low quality is more than 6,800%.

Similar results hold for both international developed and emerging markets. In general, tilts toward higher value (cheap), momentum, dividend yield, high quality, lower volatility and size (small caps) would have produced returns that exceed their benchmarks with a lower level of total risk. While this has not been proven for emerging market size and volatility, note that only seven years of data is available for our emerging markets analysis. Because emerging markets are now more correlated with international developed markets and as more time elapses, we expect the return patterns of emerging market factors to converge to those of developed markets.

NORTHERN TRUST QUALITY DEFINED

For more than 40 years, Northern Trust has applied an investment philosophy focused on identifying and developing high-quality investments across client portfolios.

Our research shows that top-quality companies – as defined by our core quality philosophy based on our assessment of their management efficiency and profitability – typically outperformed the market. Using an empirically tested approach that identifies quality companies within each sector, our proprietary methods seek to create investor portfolios better positioned to deliver outperformance with controlled volatility.[10]

Of course, there is no reason why we must limit ourselves to a tilt toward a single factor. As Exhibit 3 illustrates, a multi-factor approach can produce returns that exceed any factor individually. For example, the top 20% of stocks at the intersection of high quality and high dividend yield earn an average annual return of 20.5% (the calculated average of the highlighted quality and dividend yield intersections), which is considerably greater than the returns to the top quality quintile (19.0%) or the top dividend yield quintile (15.2%) on their own, as illustrated in Exhibit 2.

Similarly, the top 20% of stocks at the intersection of quality and value earned an annual average return of 23.8%, quality and low volatility earned 21.1%, and quality small-caps earned 21.7%, all well in excess of singular factors. Similar results hold for both international and emerging markets.

While Exhibits 2 and 3 support the case for factor tilts, they do not tell the whole story. Equally important are the pattern and consistency of factor returns. Specifically, we want to be sure that the favorable performance of factor tilts isn't due to a few brief periods in history and that we don't run the risk of protracted intervals of poor performance.

EXHIBIT 3: RUSSELL 3000, 1979 TO 2012

Quality	Dividend Yield				
	High	Q2	Q3	Q4	Low
High	24.7%	21.1%	18.0%	20.7%	19.4%
Q2	19.1%	16.9%	15.0%	17.0%	18.4%
Q3	16.8%	12.4%	11.8%	13.8%	15.7%
Q4	5.9%	11.5%	8.5%	12.4%	14.4%
Low	9.5%	5.7%	8.2%	7.7%	11.5%

Quality	Value				
	Expensive	Q2	Q3	Q4	Cheap
High	17.9%	18.1%	18.5%	22.5%	28.6%
Q2	13.9%	14.6%	16.3%	20.1%	24.1%
Q3	10.4%	12.5%	12.8%	15.5%	19.4%
Q4	9.8%	10.7%	11.5%	14.5%	16.5%
Low	3.9%	4.9%	5.2%	10.2%	11.1%

Quality	Volatility				
	High	Q2	Q3	Q4	Low
High	16.1%	19.5%	22.4%	23.3%	22.9%
Q2	15.3%	16.4%	19.8%	19.9%	18.2%
Q3	13.2%	14.1%	15.3%	16.4%	12.5%
Q4	12.2%	13.3%	14.2%	14.8%	10.1%
Low	11.3%	12.0%	10.4%	9.9%	0.8%

Quality	Size				
	Small	Q2	Q3	Q4	Large
High	23.5%	22.8%	20.4%	19.2%	15.4%
Q2	22.2%	18.1%	19.5%	16.5%	13.5%
Q3	16.1%	15.2%	14.8%	13.8%	11.8%
Q4	13.2%	14.7%	13.1%	12.2%	11.5%
Low	5.7%	6.1%	8.2%	8.8%	8.7%

Source: Northern Trust Quantitative Research

To address this issue we used the same datasets from above to create "factor mimicking portfolios" for the Russell 3000 in the traditional style. Each month we sorted stocks into quintiles based on their exposure to individual factors, buying the highest Sharpe ratio quintile and selling short the lowest quintile. In this way we attempted to capture the "pure" return of the factor net of aggregate market activity.

Performance results for these mimicking portfolios are shown in Exhibit 4. The highlighted figures indicate five-year periods in which the factor had negative performance and, hence, during which a tilt toward that factor would have likely underperformed its benchmark. As you can see, most factors have several such periods. Small-caps, for example, suffer a 15-year stretch from 1984 to 1998 where the small size factor underperforms. Likewise, dividend yield, value and low volatility each have two five-year periods (out of the seven analyzed) with negative factor performance. These factor return "cycles" are well documented in the financial literature and suggest that there is indeed some timing risk for many factor tilts.

On the other hand, the Northern Trust quality metric and all the multi-factor tilts have consistent positive returns across all periods. We also see that including quality in multi-factor tilts increases the Sharpe ratio of their single factor counterparts by at least a multiple of two: value went from 0.40 to 1.06, size from 0.20 to 1.27, low volatility from 0.20 to 0.90 and dividend yield from 0.35 to 0.94. Clearly, quality acts to stabilize returns and smooth out factor cycles, thereby eliminating much of the timing risk and volatility associated with singular factor tilts.

EXHIBIT 4: RUSSELL 3000 FACTOR MIMICKING PORTFOLIO RETURNS (Q1 – Q5) AND SHARPE RATIOS, 1979 TO 2012

Period	Value	Size	Low Volatility	Dividend Yield	Northern Trust Quality	Quality & Value	Quality & Size	Quality & Low Vol.	Quality & Dividend
Annualized Returns									
1979 to 1983	6.6%	9.9%	4.6%	–0.3%	14.9%	13.6%	13.0%	6.6%	12.9%
1984 to 1988	9.8%	–4.0%	17.4%	9.9%	17.1%	17.8%	7.1%	21.3%	22.0%
1989 to 1993	–0.1%	–0.2%	2.5%	1.0%	14.0%	6.3%	9.0%	13.2%	11.1%
1994 to 1998	1.3%	–4.5%	5.3%	18.6%	15.0%	7.1%	9.1%	14.0%	16.0%
1999 to 2003	15.7%	14.8%	–0.4%	–5.5%	20.3%	16.6%	22.7%	12.3%	11.8%
2004 to 2008	–1.7%	0.7%	3.8%	12.8%	10.3%	7.1%	12.5%	21.5%	10.4%
2009 to 2012	13.4%	3.6%	–19.5%	15.9%	2.4%	9.3%	6.6%	1.0%	6.1%
Average	6.4%	2.9%	2.0%	7.5%	13.4%	11.1%	11.4%	12.8%	12.9%
Sharpe Ratios									
1979 to 1983	0.47	1.02	0.26	–0.02	2.89	1.34	1.75	0.54	1.09
1984 to 1988	1.16	–0.48	1.14	0.68	4.60	2.96	1.41	2.13	2.77
1989 to 1993	–0.01	–0.02	0.13	0.06	2.97	0.71	1.20	1.03	0.99
1994 to 1998	0.11	–0.46	0.24	0.75	2.16	0.61	1.70	1.05	0.74
1999 to 2003	0.58	0.71	–0.01	–0.09	1.01	0.68	1.45	0.30	0.27
2004 to 2008	–0.14	0.11	0.23	0.50	1.79	0.49	1.00	1.24	0.48
2009 to 2012	0.63	0.50	–0.57	0.58	0.27	0.65	0.40	0.05	0.25
Average	0.40	0.20	0.20	0.35	2.24	1.06	1.27	0.90	0.94
Drawdowns									
Years with Positive Return	20	18	25	22	31	28	31	27	27
Years with Negative Return	14	16	9	12	3	6	3	7	7
% Positive	58.8%	52.9%	73.5%	64.7%	91.2%	82.4%	91.2%	79.4%	79.4%
% Negative	41.2%	47.1%	26.5%	35.3%	8.8%	17.6%	8.8%	20.6%	20.6%
Max 1 Month Drawdown	–26.3%	–13.5%	–58.2%	–51.8%	–19.0%	–21.1%	–7.8%	–35.2%	–28.8%

Source: Northern Trust Quantitative Research

Quality also reduced the frequency and severity of drawdowns. Of the 34 years we analyzed, singular factor tilts had positive performance in about 60% of those years while quality outperformed more than 90% of the time. By adding quality to multi-factor tilts we increase the percentage of years of positive performance to about 83%. In addition, the average maximum one-month drawdown was reduced from 37.5% for single factor tilts to 23.2% for multi-factor tilts.

By combining value, size, dividend and volatility factor tilts with Northern Trust's quality metric, investors would have earned higher, more consistent returns and would have suffered less severe drawdowns. Again, results for international and emerging markets are similar.

HOW MUCH TILT?

To what extent should you tilt your portfolio towards a set of factors? In other words, what degree of tilt should your factor-based investment manager target? At the extreme you could select a manager that invests exclusively in stocks that comprise, say, the best quintile of your factor ranking and achieve results like those highlighted in Exhibits 2 and 3. While this may seem like the optimal solution, we must recognize that most equity mandates are benchmarked to classic market-capitalization weighted indexes. Allocating exclusively to the best quintile of stocks may therefore create very large tracking error to these benchmarks, which may create undesirable optic risk.

A common lament among investors is that their manager's performance didn't meet their risk and return expectations. When a manager is hired, these expectations are, no doubt, largely based on the manager's historical track record. The frequency with which investors are disappointed with their managers is, as we will show, high. Why is this the case?

The traditional manager selection process may contribute to this challenge. At a simplified level, the typical approach involves the investor creating a "short list" selected from the top quartile or quintile of managers, based on excess returns or information ratio, and then choosing one that has the desired level of tracking error to meet investment objectives. The problem with this process is that it is optimized to a single scenario – maximizing the investment's total utility assuming current data is an accurate forecast of the future. The result is a selection that does very well if historical financial market conditions (e.g., macroeconomic, political, regulatory, etc.) prevail, but could be dramatically suboptimal if circumstances change.

What is the alternative? In our view, investors should focus on maximizing the utility of their investments recognizing that conditions will likely evolve over our holding period. Specifically, their active risk posture should reflect the potential for significant underperformance due to change.

The pervasive underperformance of managers with strong track records is not a surprise. It is, in fact, an example of the statistical phenomenon known as "reversion toward the mean," whereby if a variable is extreme on its first measurement it will tend to be close to the mean on its second measurement. By selecting from only the top quartile or quintile of managers investors are, by design, choosing among managers who have an extreme first measurement. Their second measurement, i.e., the performance after the manager is hired, is likely to be more in line with the average performance of all funds – rather than those that are top quartile or quintile, which is why managers are very likely to underperform expectations. Reversion toward the mean in fund managers is well documented by Berk and Green (2004).[11]

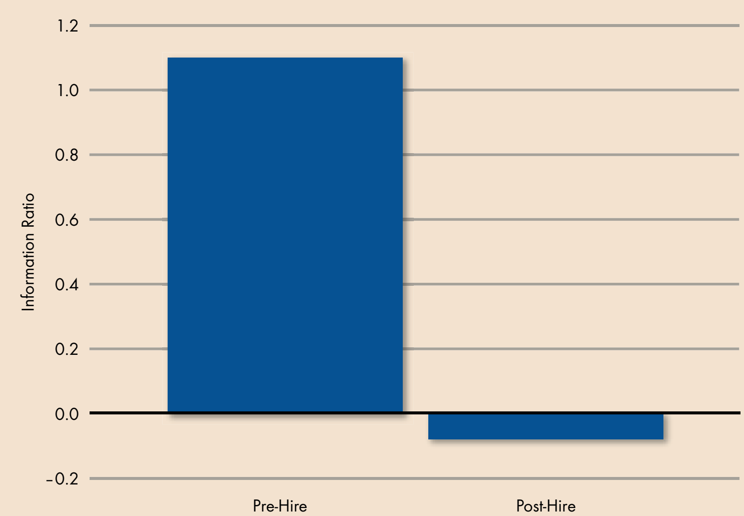
Investors should focus on maximizing the utility of their investments recognizing that conditions will likely evolve over our holding period. Specifically, their active risk posture should reflect the potential for significant underperformance due to change.

Considerable empirical evidence exists within finance literature to confirm this phenomenon. For example, Goyal and Wahal looked at the selection and termination of investment managers by 3,700 plan sponsors from 1994 to 2003.[12] They found that investment managers earned large positive active returns up to three years prior to being hired, but post-hire excess returns were close to zero.[13] Information ratios (IRs) for U.S. equity managers are shown in Exhibit 5. Note that both these IRs are calculated using alphas that are net of the three Fama-French factors. Clearly, investors are disappointed with the performance of their managers and we have probable cause to suspect reversion to the mean.

The only mathematical requirement for the existence of reversion to the mean is independence of returns from period to period. In other words, there can be, on average, no persistence in manager performance. If this is true, using the historical track record to select managers is pointless. However, Goyal and Wahal, like Carhart, show that some persistence in performance does, in fact, exist but that this persistence is due exclusively to factor tilts and not idiosyncratic "stock picking" bets.

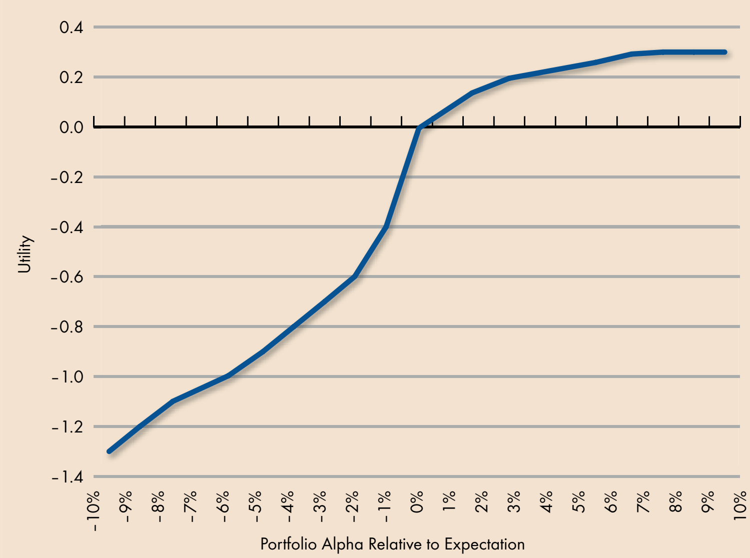
The findings of Goyal and Wahal and Berk and Green have further implications. If it is true that managers underperform their track record after they are hired, then investors should consider this when choosing an active risk posture. For example, say you are deciding between manager A and manager B, each of whom have two different levels of tracking error: 2% and 15%, respectively. Given a pre-hire IR of 1.10 (as shown in Exhibit 5) the return expectation is 2.20% for manager A and 16.50% for manager B. Post-hire, however, we expect actual returns of –0.16% and –1.20%, respectively. In other words, the magnitude of underperformance versus their track record is –2.36% and –17.70%.

EXHIBIT 5: THREE-YEAR PRE- AND POST-HIRE INFORMATION RATIOS



Source: Goyal, A. and Wahal, S., "The Selection and Termination of Investment Managers by Plan Sponsors," Journal of Finance 63(4) and Northern Trust Quantitative Research

EXHIBIT 6: UTILITY UNDER PROSPECT THEORY



Source: Northern Trust Quantitative Research

Which of these two results is most "painful"? Both managers have the same post-hire IR but manager B underperformed his benchmark by 15.34% (17.70% – 2.36%) more than manager A. We believe this difference creates significant "optic" risk for investors. All else being equal, the "pain" felt by underperforming a track record by 2.36% is much less than the pain of underperforming by 17.70%. The active risk posture should therefore shift to the smaller risk end of the spectrum.

The notion that larger losses are polynomially more painful than smaller losses is a key tenet of Prospect Theory of Kahneman and Tversky, which has seen wide application in the area of finance.[14] The authors define a utility function for returns:

$$U(r) = r^{\alpha}, \text{ for } x \geq 0$$
$$U(r) = -\lambda(-r)^{\beta} < 0$$
$$\text{where } 0 < \alpha \leq 1, 0 < \beta \leq 1, \lambda > 1$$

Here λ is a measure of relative risk aversion. A graph of the utility function for different levels of returns relative to expectation (i.e., relative to track record), where $\alpha = 0.5$, $\beta = 0.5$ and $\lambda = 2$ – typical values used in Prospect Theory studies – is shown in Exhibit 6. You can see that losses have relatively low disutility versus the same level of gains.

In our example, the utility of a 17.70% underperformance is –1.68 versus a utility of –0.61 for underperformance of 2.36%. Again, this suggests that utility-maximizing investors will appropriately move toward the more conservative end of the active risk spectrum.

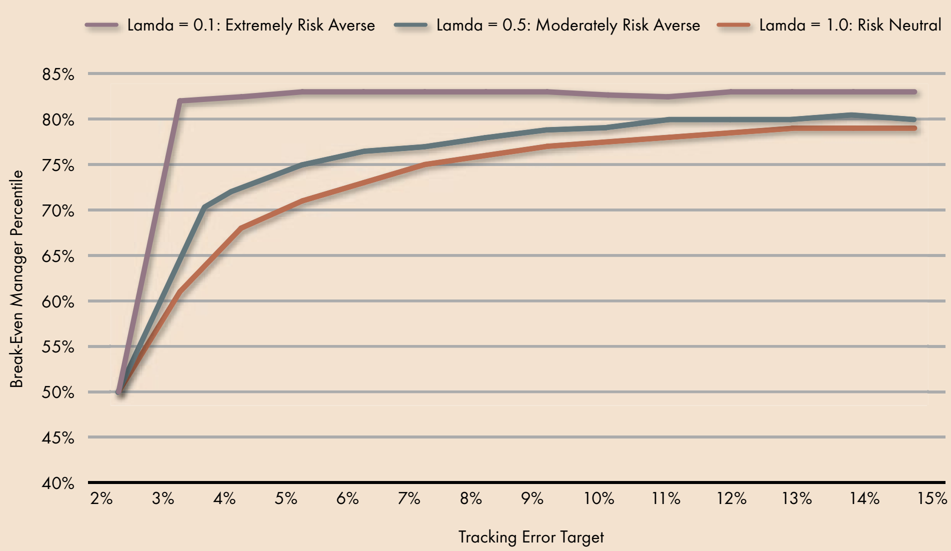
Using the concept of Prospect Theory, we can also define another version of "pain." Consider again our two managers: A and B. With the utility function of Prospect Theory, we could ask the question "if we are going to invest in manager B, what percentile of performance must we require from manager B to make the utility of B equal to the utility of manager A that performs at the 50TH percentile (average)?" In other words, if we believe in the findings of Goyal and Wahal, how good do we have to be at selecting active managers if we want to move out on the risk spectrum and keep our utility equal to a low risk – say, a 2% tracking error – manager?

This question can be answered using a relatively straightforward Monte Carlo study. At each iteration, we simulated an active return based on the IR from Goyal and Wahal and a specified tracking error target. We then computed the Prospect Theory utility for the return and stored these utility values. After the final iteration we examined the distribution of utilities and found the percentile of the distribution that was equal to the expected utility of a low risk manager (here 2% tracking error) computed using the same assumptions from Goyal and Wahal. This is the percentile performance you would need from a higher-risk manager to keep your utility equal to an average low-risk manager.

The results are shown in Exhibit 7 below. For all levels of risk aversion, the break-even percentile required of your higher-risk manager to match the utility of a low-risk manager increases as tracking error increases. For example, if you were to invest in a manager with a 6% tracking error and were risk-neutral, then your break-even manager percentile is about 72% – meaning that your manager would have to perform at or above the 72ND percentile to equalize your utilities. If you were extremely risk-averse, your manager would have to perform above the 82ND percentile to break even, and so on. Clearly, you have to be very good at selecting managers in order to justify an aggressive tracking error.

Of course, these results assume that there is absolutely no persistence in returns (per the Goyal and Wahal findings). While we have shown this to be true for fundamental managers, we've also concluded that some persistence exists for managers who make factor tilts. The actual degree of persistence varies by study but the range suggested by the papers cited above is from 30% to 60%.

EXHIBIT 7: BREAK-EVEN REQUIREMENTS BY LEVEL OF RISK AVERSION

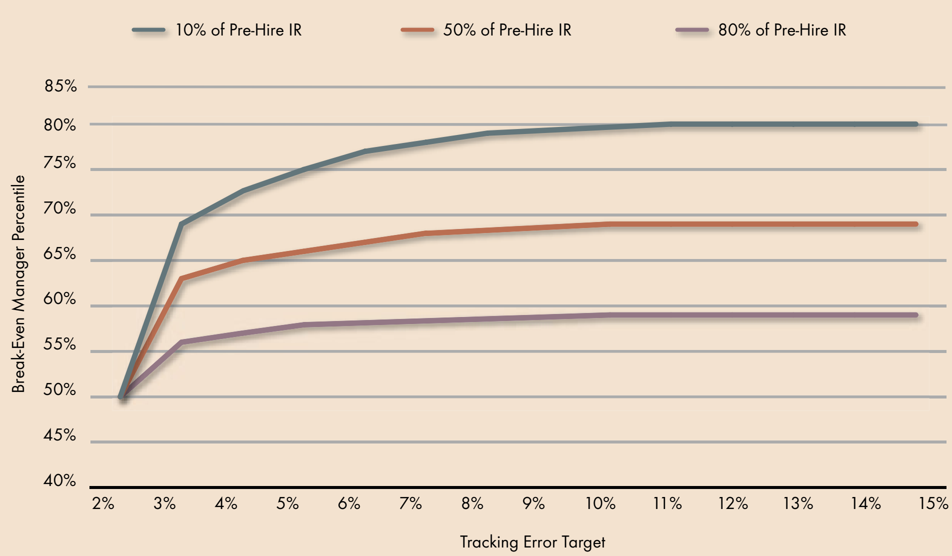


Source: Northern Trust Quantitative Research

To the extent they have less-than-perfect persistence, factor-oriented managers will still be exposed to some "pain." Exhibit 8 shows the percentile of performance needed to break even with a manager that has a 2% tracking error at different levels of IR persistence, assuming $\alpha = 0.5$, $\beta = 0.5$ and $\lambda = 2$. If, for example, you have 80% persistence and want to target a 7% tracking error, you need a manager who performs at or above the 55TH percentile. At the same tracking error level but a 50% persistence, you need a manager to perform at the 67TH percentile or above. With just 10% persistence, that number jumps to the 78TH percentile.

So even if your search is focused exclusively on factor-oriented managers, you still must consider the risk of underperformance. Given less-than-perfect persistence in IR, it makes sense for most investors to gravitate toward the conservative end of the active risk perspective, even though the "pain" associated with increasing active risk may be somewhat less extreme.

EXHIBIT 8: BREAK-EVEN REQUIREMENTS BY INFORMATION RATIO



Source: Northern Trust Quantitative Research

FACTOR TILTS CAN PROVIDE BENEFITS IF CHOSEN WISELY

Our research shows that style factor tilts can outperform traditional benchmarks and do so with lower risk. However, the specific tilts must be chosen wisely. Although value, size, volatility and dividend yield tilts all would have beaten their benchmarks over the long term, they are subject to significant timing risk and could potentially expose investors to extended periods of underperformance. However, we did see that multi-factor tilts that include quality can deliver both higher and more consistent returns, and reduce the frequency and severity of drawdowns.

Northern Trust's suite of quality products is specifically engineered to deliver the higher returns of factor tilts with more consistency and less severe drawdowns.

LEARN MORE

To learn more about our research into the effectiveness of factor tilts, or to see how using a multi-factor approach might work for your portfolio, please contact your relationship manager or visit northerntrust.com.

NORTHERN TRUST'S QUALITY SUITE

Northern Trust has developed proprietary asset management products that employ our core quality philosophy in conjunction with a specified factor tilt for investors seeking outperformance and risk-managed solutions. Our approach incorporates quality into the equity portfolio construction process and includes risk management techniques designed to minimize risks unlikely to be adequately compensated with excess returns. Our research suggests that over time investments in high-quality companies outperform a benchmark with lower or similar volatility. In addition, we apply thoughtful portfolio construction techniques to help eliminate unintended exposures and to create a portfolio well suited to various market environments. We offer solutions that combine quality with yield (Quality Dividend Focus), low volatility (Quality Low Volatility), value (Quality Value Portfolio), and size (Small Cap Core and Small Cap Value strategies) factors designed to generate higher, more consistent returns with less severe drawdowns.

FOOTNOTES

1 Sharpe, William F. "Capital asset prices: A theory of market equilibrium under conditions of risk," Journal of Finance XIX (3): 425–42, 1964.

2 Fama, E. F. and K. R. French, "The Cross-Section of Expected Stock Returns," Journal of Finance, Vol. 47, No. 2, 1992.

3 Although ignored by most subsequent authors, Fama and French also show that under their three-factor model return variance is not a sufficient statistic for a portfolio's risk. This is because the different factors in their model do not have the same weightings in the final return.

4 Carhart, M. M., "On Persistence in Mutual Fund Performance," Journal of Finance 52 (1): 57–82, 1997.

5 Fama, E. F. and K. R. French, "Multifactor Explanations of Asset Pricing Anomalies," Journal of Finance, Vol. 51, No. 1, 1996

6 Fama, E. F. and K. R. French, "Multifactor Explanations of Asset Pricing Anomalies," Journal of Finance, Vol. 51, No. 1, 1996

7 Brennan, M., T. Chordia, A. Subrahmanyam, and Q. Tong, Sell-side Liquidity and the Cross-Section of Expected Stock Returns, unpublished working paper

8 Vishny, Robert W. & J. Lakonishok, & A. Shleifer, Contrarian Investment, Extrapolation and Risk, Journal of Finance, 49(5), pp. 1541–78, 1994

9 Ang, Andrew, Robert J. Hodrick, Yunhang Xing, and Xiaoyan Zhang, "High Idiosyncratic Volatility and Low Returns: International and Further U.S. Evidence," Journal of Financial Economics, 9, 1–23, 2009 and Baker, Nardin and Haugen , Robert A., "Low Risk Stocks Outperform within All Observable Markets of the World," unpublished working paper, 2012.

10 Brennan, M. J., "Taxes, Market Valuation and Corporate Financial Policy," National Tax Journal XXIII (4): 417–428, 1970 and Litzenberger, R. H., "The Effects of Dividends on Common Stock Prices Tax Effects or Information Effects?," Journal of Finance XXXVI (2), 1982.

11 Berk, J., and R. Green, "Mutual Fund Flows and Performance in Rational Markets," Journal of Political Economy, 2004

12 Goyal, A. and Wahal, S., "The Selection and Termination of Investment Managers by Plan Sponsors," Journal of Finance 63(4) 1805–1847, 2008

13 The opposite is true for firing managers – pre-firing excess returns are negative but post-firing excess returns are significantly positive

14 Kahneman, D., and Tversky, A., "Prospect Theory: An Analysis of Decision under Risk," Econometrica, XLVII, 263–291, 1979.

We hope you enjoy the latest presentation from Northern Trust's *Line of Sight*. By providing research, findings, analysis and insight on the effects and implications of our changing financial landscape, *Line of Sight* offers the clarity you need to make better informed decisions.

Q 53766 (6/13)