WHAT IS QUALITY?

Understanding the quality phenomenon; what it is and why it exists.

Michael Hunstad
Head of Quantitative Research,
Active Equity Group,
mrh12@ntrs.com

Implementation of Northern Trust's quality philosophy, while integrated into the consistent framework that underlies our active equity and smart beta products, varies by strategy and market segment, based on our experience in managing these strategies.

Although the concept of quality has been deeply entrenched in other asset classes for decades, how this idea relates to stock selection has received considerable attention over the last few years. In this paper we explore the theme of equity quality and attempt to answer several important questions:

- How do we define quality? Are all definitions of quality equal? If not, what constitutes an "optimal" definition of quality?
- Is quality a reliable predictor of equity returns? In other words, is it a compensated risk factor? If so, why? Is there a rational explanation for why investors should be rewarded for holding high-quality stocks?
- Are returns to quality consistent or are they highly period specific? Is there a quality "cycle" and, if so, can we rotate in and out of quality to achieve better returns?
- Should index investors be concerned about quality?
- Is quality related to the low volatility phenomenon?

DEFINING QUALITY

Unlike other equity factors such as value or size, there is no generally agreed upon definition of quality. Although the pervasive, intuitive notion is that high-quality companies should have better overall performance than low-quality companies, their differentiating characteristics are hotly debated. Despite years of research, the finance community has made little headway in developing a core theory of equity quality.

Practitioners and academics alike often simply relate quality to some measure of profitability. Examples include Novy-Marx,[1] who defines profitability as the ratio of gross profits (revenue minus cost of goods sold) to assets; and Fama and French[2] who use the ratio of equity income to book value. The latter is roughly in line with Dimensional Fund Advisors'[3] direct profitability measure, defined as operating income (before depreciation and amortization) minus interest expense (scaled by book value). Numerous other authors and asset managers simply equate quality and profitability to return on equity (ROE).

Others extend beyond a single quality/profitability measure into a multi-metric definition. The MSCI Quality Indices[4] for example, identify quality as a combination of ROE, debt-to-equity ratios and earnings variability which is defined as the standard deviation of year over year earnings per share growth over the last five years. This is similar to the Russell U.S. Equity Indices[5] version of quality that includes return on assets, leverage and earnings variability. The S&P Quality Rankings[6] although proprietary, are known to be comprised of earnings and dividend metrics. Asset managers such as AQR Capital Management[7] have extended the work of Novy-Marx to include total profits over assets and gross margins plus free cash flow over assets. Finally, the F score developed by Piotroski[8] utilizes nine metrics including net income, operating cash flow, return on assets, stability of earnings, leverage, liquidity issuance, gross margins and asset turnover.

Unlike other equity factors such as value or size, there is no generally agreed upon definition of quality. Although the pervasive, intuitive notion is that high-quality companies should have better overall performance than low-quality companies, their differentiating characteristics are hotly debated.

Table 1: Measures of Profitability as Defined By Practitioners and Academics

Quality Definition	Identifier	Definition
AQR Capital Management	AQR	• Total profits/assets • Gross margins • Free cash flow/assets
Dimensional Fund Advisors Direct Profitability	DFA	• Operating income before depreciation and amortization minus interest expense scaled by book value
MSCI Quality Indices	MSCI	• Return on Equity (ROE) • Debt to equity • Earnings variability: Standard deviation of YOY earnings per share growth over last five fiscal years
S&P Quality Rankings	S&P	• Based on per-share earnings and dividend records for the most recent 10 years. Basic scores are computed for earnings and dividends and then adjusted by a set of predetermined modifiers for changes in the rate of growth, stability with long-term trends and cyclicality. Adjusted scores for earnings and dividends are then combined to yield a final ranking.
Piotroski F-Score	F-Score	• Nine metrics: - Return on assets - Operating - Cash flow - Quality of earnings - Net income - Leverage - Liquidity equity issuance - Gross margins - Asset turnover
Return on Equity (ROE)	ROE	• Trailing twelve month income/average equity

Sources: AQR: "A New Core Equity Paradigm, Using Value, Momentum, and Quality to Outperform Markets," AQR White Paper, March 2013.
MSCI: "MSCI Quality Indices Methodology," MSCI White Paper, December 2012.
Publicly available definition - S&P 500: "S&P 500 Quality Rankings Index, Index Methodology," S&P Indices White Paper, March 2011.
DFA: "Dimensional's Growth Portfolios," Dimensional Fund Advisors White Paper.
F-Score: Piotroski, J., "Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers," Journal of Accounting Research, Vol 38 Supplement, 2000

While these are just a handful of the available measures of quality, they reflect the diversity of views on how quality is defined. The litany of perspectives naturally leads us to question what exactly these varied definitions are trying to identify. We could, for example, characterize value stocks in a number of ways: price-to-earnings, price-to-sales, price-to-book value, etc., but all of them are easily recognized as measure of cheapness that are logically connected and should, at least in principle, convey the same information. This cannot necessarily be said of quality measures. How are measures of solvency directly comparable to profitability measures? The answer is that they are not, yet both may or may not end up in a definition of quality. From this result, we begin to suspect that, unlike size or growth, there may be multiple, unrelated facets of quality that must be gauged independently.

The same conclusion can be derived from a simple thought experiment. When asked the hallmark characteristic of a value stock, most would reply that it be 'cheap'. Similarly, we identify small-cap stocks by their small relative size. How would we characterize quality stocks? We could use profitability, management efficiency, prudent capital structure or cash flow and still be correct.

Unfortunately, this insight only complicates the matter. If quality is multidimensional, then isolating what it is trying to measure becomes even more difficult. We could generalize in a manner similar to Calvert[9] that quality gauges "sustainable competitive advantage," or be more vague and state that "quality companies generate consistent shareholder value over time and that value will be realized by the market[10]." Yet the ambiguity inherent in these definitions is unsatisfying and provides no guidance as to how to construct an optimal quality measure. At least for now, it seems that an acceptable definition of quality remains unarticulated.

QUALITY REDEFINED: NORTHERN TRUST'S QUALITY SCORE

Even if we have not yet settled on a clear definition of what quality is, there is ample evidence of what quality does.

Although quite general, we have evidence to state that "quality" includes those features of a company that particularly appeal to risk-averse investors. While lacking specifics, this definition gives us an intuitive framework with which to think about constructing a quality measure.

Risk-averse investors seek profitability, cash flow, solvency and ability to pay dividends. It is not enough simply to state that quality companies are profitable or have low earnings volatility – clearly, any quantification of quality must be truly multidimensional, and a high-quality company must have strong performance across a wide range of metrics.

To illustrate, Northern Trust developed a proprietary method, the Northern Trust Quality Score (NTQS)[11], which gauges multiple dimensions of quality grouped under the headings of Management Efficiency, Profitability and Cash Flow. These signals are based on our fundamental belief that a quality company should encompass the characteristics detailed in Table 2.

Table 2: The Dimensions of Quality as Defined by the Northern Trust Quality Score (NTQS)

Characteristic	Measurement
The ability to convert assets into sales	Measures management's ability to produce cash flow from booked assets. This is a measure of gross asset efficiency
The ability to convert assets into earnings	Quantifies the efficiency of total asset utilization (booked and intangible). Higher earnings per dollar of total assets means a higher level of efficiency
The ability to convert equity into returns	Gauges how well shareholder equity (paid in capital and retained earnings) is converted into profit
The ability to convert invested capital into returns	Assesses a company's efficiency at allocating the total capital under its control (equity + debt) to profitable investments
The ability to remain solvent	Quantifies the financial aggressiveness and overall leverage of the firm. This aggressiveness is seen as a measure of long-term liquidity risk
The ability to self-finance	Gauges a company's ability to generate internal liquidity and avoid costly alternative sources of funding
The ability to grow prudently without becoming overextended	Measures the trajectory and sustainability of management's growth targets

Is this multi-faceted definition of quality really a better proxy for risk aversion than, say, a simple ROE measure? Can it better distinguish risk-averse from risk-seeking behavior? How does it compare with the other definitions of quality cited earlier?

To answer these questions, a database was created of all stocks that received an S&P Quality Ranking from 1985 to 2012. For each month, we computed the NTQS and the six competing quality metrics detailed in Table 1 of each of these stocks. Stocks were then ranked and placed into quintiles for each component of the definition of quality, and one-month returns for each quintile were calculated. These returns were then regressed against the Fama-French-Carhart factors obtained from Ken French's website[12] to eliminate any returns to beta, value, size and momentum biases manifest in a quality definition. In this way, the resultant returns are "pure" in the sense that they represent returns only to quality and not other factors.

We then computed factor mimicking portfolios defined as the returns net of Fama-French-Carhart factors of the highest quality (first quintile) stocks minus the net returns of the lowest quality (fifth quintile) stocks. The objective of the factor mimicking portfolio is to proxy the true return of the quality factor, independent of aggregate market movements. In this way, we again hope to capture the "pure" exposure to the quality factor and minimize extraneous noise.

Note that all quality scores are computed in accordance with our interpretation of publicly available descriptions of the respective scoring methodology and should only be viewed as an approximation. Further, the method by which the scores are actually implemented into the portfolio may also differ significantly - returns to factor mimicking portfolios are not intended to parallel the performance of any specific product and are merely a convenient method to compare quality definitions.

The performance of the factor mimicking portfolios is shown in Table 3. The average annual return of the NTQS factor mimicking portfolios exceeded 9.4%, while that for the next highest alternative definition of quality was 8.4%. Perhaps even more importantly, the standard deviation of returns to the NTQS was just 5.2% yielding a return per unit of risk of 1.81 which is nearly double that of the next highest quality definition.

Table 3: Performance of Factor Mimicking Portfolios
S&P Quality Universe 1985 - 2012

	AVERAGE ANNUAL RETURNS						
	NTQS	**DFA**	**AQR**	**MSCI**	**F-Score**	**ROE**	**S&P**
Return	**9.4%**	6.3%	3.2%	4.9%	**8.4%**	5.5%	2.9%
Stdev	5.2%	8.4%	7.7%	8.3%	8.7%	8.5%	18.7%
Return/Stdev	**1.81**	0.75	0.41	0.59	**0.97**	0.64	0.15
	CORRELATIONS						
	AQR	**DFA**	**NTQS**	**MSCI**	**F-Score**	**ROE**	**S&P**
AQR	1.00						
DFA	0.55	1.00					
NTQS	0.62	0.49	1.00				
MSCI	0.54	0.54	0.54	1.00			
F-Score	0.27	0.50	0.37	0.45	1.00		
ROE	0.70	0.83	0.60	0.72	0.46	1.00	
S&P	0.33	0.26	0.25	0.18	0.09	0.38	1.00

Source: Northern Trust Quantitative Research
S&P Quality Rankings from Factset

DIFFERENCES BETWEEN HIGH- AND LOW-QUALITY STOCKS

Each month, stocks across the Russell 3000, MSCI World ex U.S. and MSCI Emerging Markets universes are ranked and formed into equally weighted quintiles based on their NTQS. The subsequent one month return is computed for each quintile and average returns and Sharpe ratios are shown in Table 4. Clearly, there is a significant difference in return performance between high- and low-quality stocks, regardless of the universe in question.

Even more impressively, these higher returns did not seem to come with higher risk. In fact, the opposite was true. The Sharpe Ratios, which measure the return per unit of risk, for domestic high-quality stocks has been nearly five times that of domestic low-quality stocks, while international developed and emerging markets show differentials of almost three and two times, respectively. This demonstrates that investors, at least historically, have been well-compensated for holding higher quality stocks.

Table 4: Returns and Sharpe Ratios by Northern Trust Quality Score Quintile

	NORTH AMERICA Russell 3000		INTERNATIONAL MSCI World Ex U.S.		EMERGING MARKETS MSCI EM IMI	
Quality	Returns 1979 to 2012	Sharpe Ratio 1979 to 2012	Returns 1996 to 2012	Sharpe Ratio 1996 to 2012	Returns 2005 to 2012	Sharpe Ratio 2005 to 2012
High	19.0%	1.01	11.9%	0.69	17.4%	0.67
Q2	16.5%	0.91	9.5%	0.54	16.9%	0.64
Q3	12.5%	0.65	8.7%	0.50	14.9%	0.55
Q4	11.6%	0.56	6.5%	0.34	15.1%	0.54
Low	5.1%	0.21	4.8%	0.24	10.4%	0.33

Source: Northern Trust Quantitative Research

While these results are exciting, they are also, unfortunately, counterintuitive for adherents to the Capital Asset Pricing Model (CAPM) of Sharpe[13] and Lintner[14] —perhaps the most widely used analytical model in all of finance—which suggests that return is related positively to risk, such that any higher returns associated with holding high-quality stocks should also come with elevated levels of risk. Empirically, however, this has not been the case. The higher returns of high-quality companies have, in fact, been associated with considerably lower levels of risk.

Perhaps even more interesting is the question of why should high-, as opposed to low-, quality stocks earn a risk premium? It is widely held that, for example, the value and size risk premia are compensation for the distressed and illiquidity risks inherent in cheaper and smaller names. What additional risks do we face with high-quality stocks that are avoided by holding lower quality? While the question seems almost absurd, we must find an answer if we can lend credence to quality as a sustainable source of alpha. Otherwise, we have no choice but to dismiss the results of Table 4 as spurious.

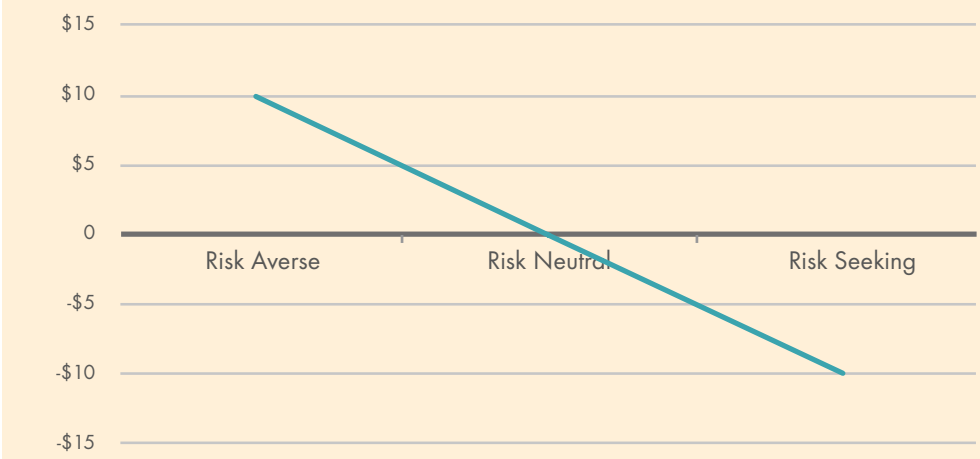
RISK PREFERENCES AND QUALITY

Although the "higher return equals higher risk" message of CAPM is deeply ingrained in our collective notion of capital markets, we must recall that CAPM requires certain assumptions about investor behavior. Specifically, CAPM suggests that investors are risk averse. If CAPM is the true model of equity returns then it is not sufficient that investors be risk averse on average, but that all investors must always share this same risk posture.

Risk aversion is an expression of an investor's affinity toward risk that is best described through example. Consider a simple coin toss whereby if the coin comes up heads you win $100 and if the coin comes up tails you pay $100. Since flipping heads and tails are equally likely, the expected value of the coin toss is zero ($100*0.5 + -$100*0.5 = $0). An investor who is risk averse would shy away from this type of game since on average there is nothing to be gained. To persuade a risk-averse investor to undertake the coin toss the payoffs must be adjusted such that the expected value of the game is something greater than zero. For example, a risk-averse investor is more likely to undertake the coin toss if heads means you win $120 and tails you lose $100. The expected value of this game is now positive at $10 ($120*0.5 + -$100*0.5 = $10) and reflects the compensation required by a risk-averse investor to bear the uncertainty of the coin toss. As a rule, risk-averse investors inherently try to avoid risk and require some additional premia to undertake investments with a high degree of uncertainty. This, of course, is the same central message of the CAPM.

While risk aversion seems like a reasonable model of investor behavior, there are other possibilities. By definition, risk-neutral investors are indifferent towards risk and, therefore, would be indifferent toward a coin toss with payoffs of $100 and -$100. To participate, a risk-neutral investor would require no additional compensation to bear the risk of the game, behavior that conflicts with CAPM. Likewise, risk-seeking investors actively pursue risk, such that they may be willing to undertake the coin toss even if it had a negative expected value, as illustrated by Chart 1. A risk-seeking investor may, in fact, be willing to undertake the coin toss if the payoffs were $100 and -$120 such that the expected value is -$10 ($100*0.5 + -$120*0.5 =-$10), again behavior in conflict with CAPM.

Chart 1: Hypothetical Required Risk Premium



Source: Northern Trust Quantitative Research

If we know the aggregate population exhibits a mixture of risk-seeking and risk-averse behavior, why shouldn't equity market investors? As stated, CAPM assumes universal risk aversion. Could a heterogeneous mix of investor behavior explain why high quality companies outperform low quality, in direct violation of CAPM?

Admittedly, the existence of risk-neutral and risk-seeking investors could be questioned. Is it logical that anyone would willingly participate in a game with a negative expected value? This is certainly the generally held belief in the finance community such that CAPM remains the most venerated model of equity returns despite its risk aversion assumption.

However, before we eliminate the possibility of non-risk-averse investors and write off the returns to quality in Table 3 as an artifact of data mining, let's change the coin toss example slightly. Consider a coin that instead of having a 50% likelihood of turning up heads, there is now just a 25% probability. By default then, the chance of tails is 75%. Is it now seemingly any more plausible that an investor would be willing to accept the toss if heads wins $200 and tails loses $100? No, probably not – the expected value is still negative ($200*0.25 + -$100*0.75 = -$25).

But what if we extend the example further? What if the probability of heads was just 1% but the payoff was increased to $900. The investor now wins $900 if heads comes up with 1% probability but loses $100 if tails comes up with 99% probability. What if the investor won $1 million with a 0.00009% likelihood but lost $1 with 99.99991% probability? While both of these examples still have a negative expected return, you can see that we are approaching probabilities and payoffs that are similar to state and national lotteries. Corbett and Geyer (2006) analyzed the multi-state Powerball Lottery and found that it did indeed have a negative expected value. Still, according to the American Gaming Association there were almost $25 billion in lottery revenues (amounts wagered minus winnings) in 2007, suggesting that the total amount wagered on lotteries could easily exceed $250 billion annually.

Negative expected value games extend beyond lotteries. It is well known that most, if not all, casino games have negative expected values as does pari-mutuel wagering. The same American Gaming Association study suggests that total legal gaming revenues in the United States (including lotteries) was more than $92 billion in 2007 with total amounts wagered easily approaching $1 trillion. More recent studies indicate even these estimates are conservative and that total amounts wagered could be in the $2 to $3 trillion range or about 7% to 13% of GDP. In any case, these figures represent about 10% to 20% of the market capitalization of the S&P 500 and are roughly equal to the total domestic small cap market. Clearly, the existence of risk-seeking behavior cannot be denied.

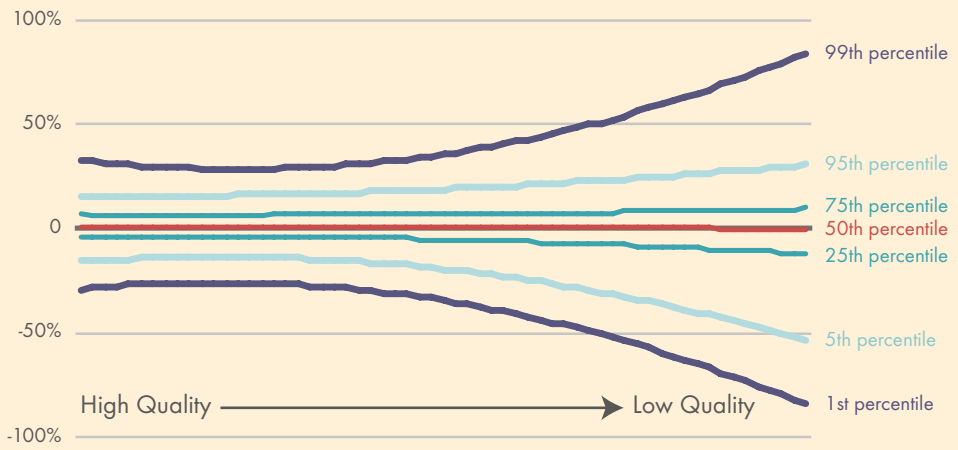
On the other hand, there is also significant evidence of risk aversion. Insurance products such as life and property and casualty coverage are known to appeal to individuals that are risk averse. According to Compustat, the total revenue of insurance companies in the Russell 3000 was approximately $900 billion in 2012 which also represents about 7% of GDP. Similarly, low risk products like certificates of deposit, short-term treasury bills, interest-bearing demand deposit accounts and money market funds have a large share of the overall investment market.

If we know the aggregate population exhibits a mixture of risk-seeking and risk-averse behavior, why shouldn't equity market investors? As stated, CAPM assumes universal risk aversion. Could a heterogeneous mix of investor behavior explain why high-quality companies outperform low-quality companies, in direct violation of CAPM?

HETEROGENEOUS INVESTING

To study this assertion, we begin by examining the distribution of risk across the spectrum of quality. Again we will use Northern Trust's Quality Score as our quality definition and perform quantile regressions on one month forward returns using the approach of Koenker and Hallock. Results are shown in Chart 2[15].

Chart 2: Quantile Regression of Quality on 1 Month Forward Returns Russell 3000 1979 - 2012



Source: Northern Trust Quantitative Research

It is clear that the distribution of quality returns expands as you move from high-quality to low-quality stocks. This not only confirms our previous contention that low-quality stocks indeed have higher volatilities than high-quality stocks, but it also suggests that the really big payoffs, e.g., 40% or more, occur more frequently in lesser quality names. Purely from a relative risk perspective, we might expect the lower quality end of the spectrum to be inhabited by risk-seeking investors while those who are risk averse will gravitate toward the higher quality end.

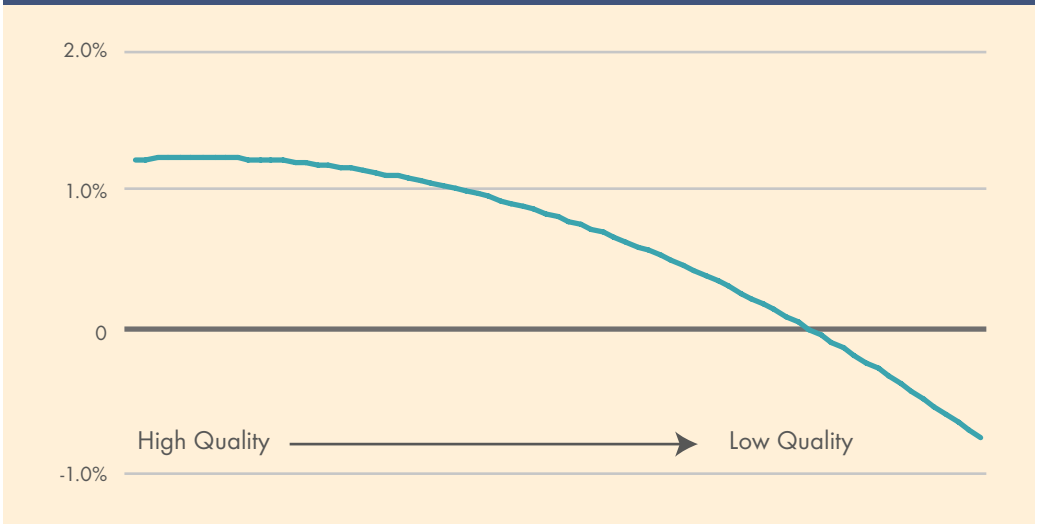
We can, however, compute the historical average return across the quality spectrum. To do so, we construct a polynomial regression of the NTQS on one month forward returns, the results of which are shown in Chart 3. Although the general shape of the curve simply confirms what we already presented in Table 4, here we see the more granular detail of how quality impacts expected return. Note that at lower levels of quality, the returns are close to zero or negative. How can this be?

If we apply the notion of heterogeneous investors, the shape of this graph makes sense. Stocks with low levels of volatility are the domain of risk-averse investors that demand, and receive, more certainty in their investments. As quality erodes, the risk of the stock increases. Risk-seeking investors bid up the price of lower quality names, thus lowering their expected returns. Extremely volatile stocks are most attractive to risk seekers, who bid the stock prices up to the point that expected returns are reduced.

We might conclude the primary driver of the quality phenomenon is the heterogeneity of investor risk preferences. High-quality stocks outperform low quality because low-quality/high-volatility names are relatively expensive versus high-quality/low-volatility names.

From this analysis, we might conclude the primary driver of the quality phenomenon is the heterogeneity of investor risk preferences. High-quality stocks outperform low quality because low-quality/high-volatility names are relatively expensive versus high-quality/low-volatility names. Risk-seeking investors bid up the price of low-quality stocks to the point they are bound to underperform. The risk-seeking objective is to capture the large potential upside even though it means introducing more uncertainty.

Chart 3: Regression of Quality on 1 Month Foward Returns
Russell 3000 1979 - 2012



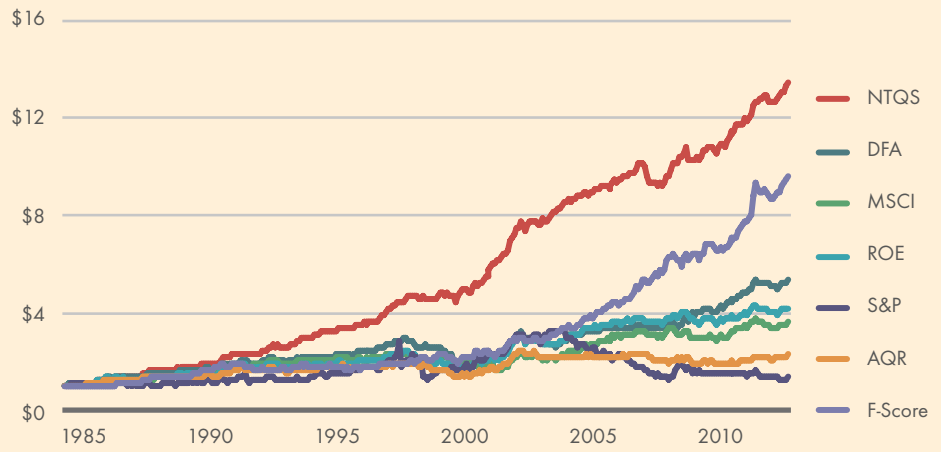
Source: Northern Trust Quantitative Research

While the analysis in Chart 3 was performed for the Russell 3000, virtually identical results are achieved with other developed and emerging market universes. In all cases, lower-quality names had both higher volatility and lower expected returns. It appears that heterogeneity of risk preferences can explain the quality anomaly in any equity market.

COMPARING QUALITY FACTOR PERFORMANCE

Chart 4 shows the cumulative performance of $1 invested in each of the seven alternate definitions of quality. The NTQS achieved a total return of nearly $13.50 with the Piotroski F-score next at $9.80. The rest of the definitions trailed significantly behind second place with two of the quality definitions actually achieving a negative overall return.

Chart 4: Value of $1 Invested in Factor Mimicking Portfolios
S&P Ranked Universe 1985 to 2012



Source: Northern Trust Quantitative Research

Note, however, that this analysis was restricted to only those stocks that received an S&P Quality Ranking. Because this ranking requires at least ten years of historical data, the number of stocks receiving an S&P Quality score is fairly small at roughly 1000 per average month. A better comparison might be made amongst the remaining definitions of quality by looking at a larger universe.

Table 5 details the performance of factor mimicking portfolios constructed with the full Russell 3000 universe from 1979 to 2012. Here again, the NTQS has the highest performance among all definitions of quality, beating the second highest performer by an average of 550 bps per year. The standard deviation of returns was also lower than any other quality definition resulting in a return per unit of risk that was more than 2.5 times higher than the second place Piotroski F-score.

Table 5: Performance of Factor Mimicking Portfolios
Russell 3000 1979 - 2012

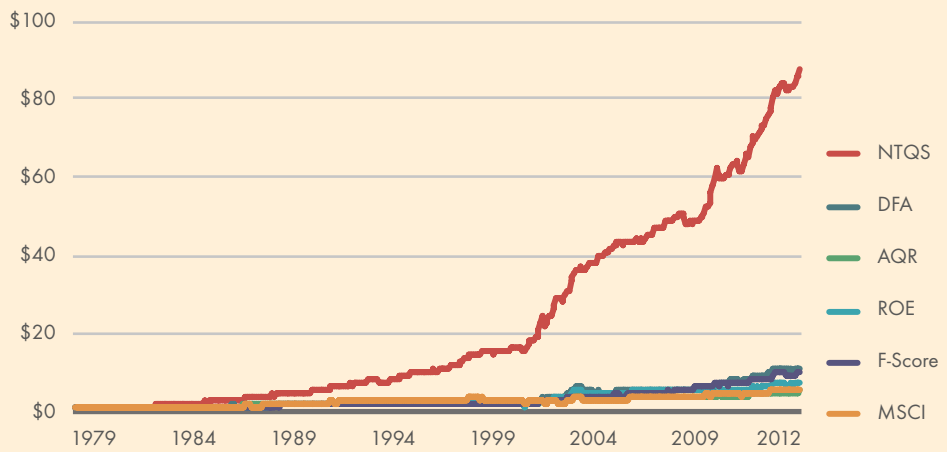
	AVERAGE ANNUAL RETURNS					
	NTQS	DFA	AQR	MSCI	F-Score	ROE
Return	13.4%	7.9%	5.4%	5.7%	7.4%	6.5%
Stdev	6.6%	12.4%	10.2%	10.3%	9.4%	11.2%
Return/Stdev	2.03	0.64	0.53	0.55	0.79	0.59
	CORRELATIONS					
	NTQS	DFA	AQR	MSCI	F-Score	ROE
NTQS	1.00					
DFA	0.73	1.00				
AQR	0.76	0.76	1.00			
MSCI	0.54	0.61	0.62	1.00		
F-Score	0.69	0.73	0.64	0.53	1.00	
ROE	0.77	0.90	0.86	0.74	0.73	1.00

Source: Northern Trust Quantitative Research

By better isolating and avoiding those companies with negative expected returns, the NTQS outperformed simpler and less effective measures of quality.

Chart 5 details the cumulative performance of quality factor mimicking portfolios for the Russell 3000 universe. One dollar invested in January of 1979 would have grown to almost $90 in the NTQS factor portfolio. Cumulative returns to other definitions of quality were less than $12.

CHART 5: Value of $1Invested in Factor Mimicking Portfolios Russell 3000 1979 to 2012



Source: Northern Trust Quantitative Research

These findings suggest the broad NTQS definition of quality does succeed at identifying those characteristics that most appeal to risk-averse investors. By better isolating and avoiding those companies with negative expected returns, the NTQS outperformed simpler and less effective measures of quality.

IMPLICATIONS FOR INVESTORS

These results have very important implications for investors employing index strategies. Because these strategies buy stocks across the quality spectrum investors should be concerned that a sizeable allocation of an index investment will be made to low-quality names with low or negative expected returns. Since investors are, on average, risk averse as evidenced by the positive expected value of index returns, they may be earning risk premia that are lower than what otherwise could be achieved in a high-quality portfolio.

CRISIS ASYMMETRY AND FACTOR TIMING

We have shown that the NTQS definition of quality has generated excess returns by better identifying companies possessing characteristics that are attractive to risk-averse investors. By avoiding names with low or negative expected returns, the performance of a portfolio of high-quality stocks will, on average, exceed those of lower quality. Although this is certainly a critical component of the total return to quality, it is only part of the story.

While we have made the case that heterogeneity of investor risk preferences produces the quality phenomenon, it would be an error to think risk preference is necessarily static through time. For example, it seems reasonable that there would be relatively fewer risk-seeking investors during times of market stress and relatively more during periods of market recovery. If investors are indeed changing their risk preferences as market conditions change this may also have important implications for how quality performs.

Consider an economy about to experience a recession. As evidence of the recession grows and the potential upside to high-volatility/low-quality stocks is diminished, fewer investors will be identified as risk seeking. We would, in fact, expect many formerly risk-seeking investors to become relatively risk averse as macroeconomic conditions deteriorate causing prices, and hence returns, of low-quality stocks to fall dramatically and high-quality stocks to rise on a relative basis. This, of course, is the classic flight to quality.

Is the mirror image necessarily true as macroeconomic conditions begin to improve? Do low-quality stocks outperform high-quality to the same degree as high previously outperformed low? Although this seems logical the empirical evidence suggests the answer is a resounding no.

Table 6 analyzes the returns to the NTQS factor mimicking portfolios during domestic and global recessions. Official recession dates are taken from the National Bureau of Economic Research (NBER) and the International Monetary Fund (IMF). High-quality stocks outperformed low-quality during the period six months prior to a recession to mid recession. In the U.S. for the Russell 3000 universe the average annualized factor mimicking portfolio return was almost 20%. Internationally for the MSCI All World ex-U.S. the factor mimicking portfolio return was almost 14%.

From mid-recession to six months after the official end of the recession, the mirror image of returns was not seen. Although low-quality stocks did tend to outperform high quality, the magnitude was relatively small. In the U.S. the excess return of low-quality over high averaged just 4% and was approximately the same internationally.

Table 6: Returns to NTQS Factor Mimicking Portfolio During Recessions

RUSSELL 3000: 1979 TO 2012			MSCI ALL WORLD EX-U.S. 1996 TO 2012		
NBER Recession	6M Prior to Mid Recession	Mid Recession to 6M After	IMF Recession	6M Prior to Mid Recession	Mid Recession to 6M After
U.S. Recession 12/07 - 6/09	22.5%	-16.0%	Global Recession 1998	10.8%	-6.8%
U.S. Recession 12/00 - 7/01	36.0%	-0.5%	Global Recession 2001-2002	18.2%	-3.6%
U.S. Recession 7/90 - 3/91	15.4%	1.4%	Global Recession 2008-2009	12.3%	-2.2%
U.S. Recession 7/81 - 11/82	11.5%	-4.1%			
U.S. Recesion 1/80 - 7/80	12.6%	0.1%			
Average	**19.6%**	**-3.8%**	**Average**	**13.8%**	**-4.2%**

Source: Northern Trust Quantitative Research

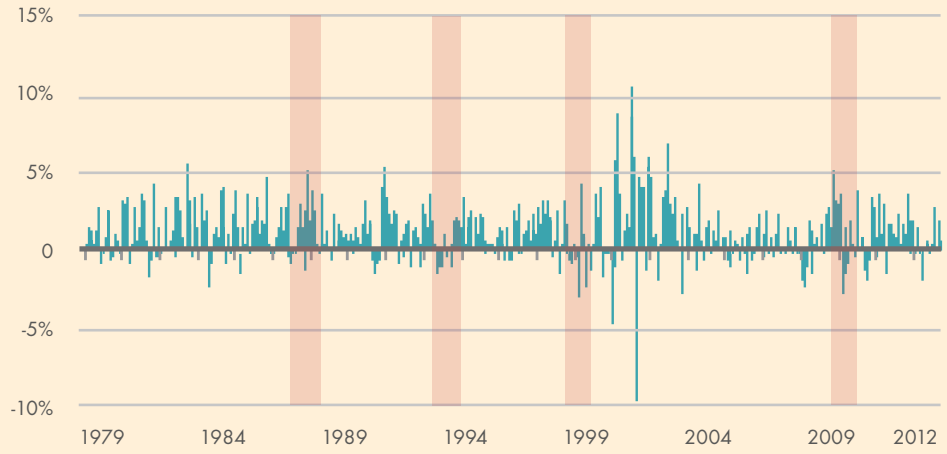
There are two reasons for why this is the case. First, it is not unusual for low-quality stocks, especially smaller low-quality stocks in broad indices such as the Russell 3000, to have such extreme price movements during a flight to quality that their total capitalization no longer qualifies them for inclusion in the index. Once they are removed, the loss is essentially locked in as they are likely replaced by other names with less potential for rebounding.

Second, and perhaps more importantly, it takes considerable time for investor risk preferences to revert back to their pre-recession posture. So while flights to quality are largely defined by macroeconomic conditions, "junk rallies" are not. Although it is true these rallies tend to occur in relatively benign economic climates, the triggering mechanism is not exclusively economic.

For the Russell 3000 since 1979 there have been only four junk rallies, defined as a negative return to the NTQS factor mimicking portfolio, that exceeded three consecutive months in length. Even amidst these protracted rallies, on average more than 80% of the total return in the rally was concentrated in a single month. As a rule, junk rallies tend to occur in very short bursts that do not follow the much slower macroeconomic cycle but, instead, tend to be driven more by sentiment. This makes junk rallies notoriously difficult to predict and, hence, timing or tactically rotating in and out of quality an unrealistic prospect.

Chart 6 shows the monthly performance of the NTQS factor mimicking portfolio. As you can see, the returns to low quality, i.e., negative factor mimicking portfolio returns, occur in small isolated patches as opposed to broad cycles.

Chart 6: Factor Mimicking Portfolio Returns
Russell 3000 1979 - 2012



Source: Northern Trust Quantitative Research

While low quality and high volatility are closely related concepts, the reverse is not true.

QUALITY AND LOW VOLATILITY

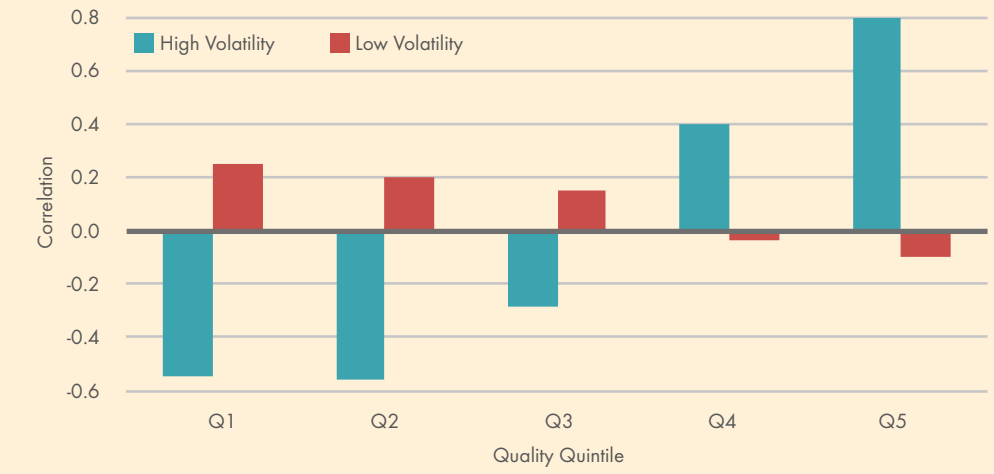
Based on the preceding analysis, a superficial case can be made that the quality phenomenon is simply a repackaging of the low volatility phenomenon documented by Blitz and van Vilet[16] and Ang, Hodrick, Xing and Zhang[17], among others. After all, we have attributed the excess returns of high-quality portfolios to simply avoiding high-volatility, low-quality names that tend to be overpriced. This is in some ways similar to the explanation given by Baker, Bradley and Wurgler[18] when describing the returns to low volatility. However there are very important differences.

While it is quite reasonable to expect a strong overlap between low-quality and high-volatility companies, it is not necessarily the case that high-quality companies have the lowest volatility. Recall our new definition of quality as the set of company characteristics that appeals to the risk-averse investor. While low historical volatility may be among that definition, it is only one of a multitude of factors. Thus, the relationship between low volatility and risk aversion is less well defined.

To illustrate this concept, we collected monthly data on all stocks in the Russell 3000 universe from 1979 to 2012. Each month we separately ranked stocks based on volatility and NTQS and, based on these results, placed each stock into a quality and a volatility quintile. The subsequent one month equally weighted returns for each of these quintiles was calculated. If our assertions are correct, we would expect to see a very high correlation between the returns of low-quality and high volatility but much more muted correlations between high-quality and low-volatility. This is indeed the case.

The results of this analysis are shown in Chart 7. The blue bar in the fifth quintile of quality suggests the return correlation between the low-quality quintile and the high volatility was a very strong 0.8. In contrast, the red bar in the first quintile of quality indicates the correlation between high quality and low volatility was a very modest 0.23. From this analysis it is clear that while low quality and high volatility are closely related concepts, the reverse is not true. Thus, there are significant differences between the aggregate low volatility and quality phenomenon.

Chart 7: Correlation of Volatility Quintiles to Quality Quintiles



Source: Northern Trust Quantitative Research

QUALITY AND REALIZED TRACKING ERROR

Finally, we note that the addition of quality to an equity portfolio will often cause the realized tracking error of that portfolio to be less than the initial target. In other words, a manager may construct a high-quality portfolio with an ex ante tracking error of, say, 2.5%, only to find that going forward the realized ex post tracking error is considerably lower at, say, 1.5%. In this section we will show this is the direct result of pursuing high-quality stocks and provide both a logical explanation and empirical evidence for this phenomenon.

Recall that computing ex ante tracking error requires two components: the historical volatilities of individual stocks and the historical correlations among stocks in the portfolio. In a simple portfolio of just two securities A and B the ex ante tracking error is:

$$TE_{ExAnte} = \left(d_A^2\sigma_A^2 + d_B^2\sigma_B^2 + 2d_A d_B \sigma_A \sigma_B \rho_{AB}\right)^{0.5}$$

Where d_A and d_B are the stock weight deviations from the benchmark, σ_A and σ_B are historical volatilities and ρ_{AB} is the historical return correlation between stock A and B. In contrast, the ex post tracking error is an average realized deviation from the benchmark return:

$$TE_{ExPost} = \left(\frac{1}{n}\sum_{i=1}^{n}\left(r_p - r_b\right)^2\right)^{0.5}$$

Where r_p is the return on the portfolio and r_b is the benchmark return. If stock volatilities and correlations are stable then ex ante and ex post tracking error should be roughly equal. However, for high-quality stocks both historical volatilities and correlations tend to be overstated causing ex ante tracking error to also be overstated and, hence, realized ex post tracking error to be less than predicted by the ex ante model.

Table 7 shows the average exponentially weighted stock volatility with a 2% daily decay, and subsequent ex ante stock volatility and subsequent one year ex post volatility for the five NTQS quintiles from 1985 to 2012. As you can see, the historical ex ante volatility of high-quality stocks tends to be higher than the realized ex post. Conversely, low-quality stocks tend to have somewhat higher ex post volatility. This is the result of stocks changing quality scores through their history.

Table 7: Average Ex Ante and Ex Post Volatility
Russell 3000 - 1985 to 2012

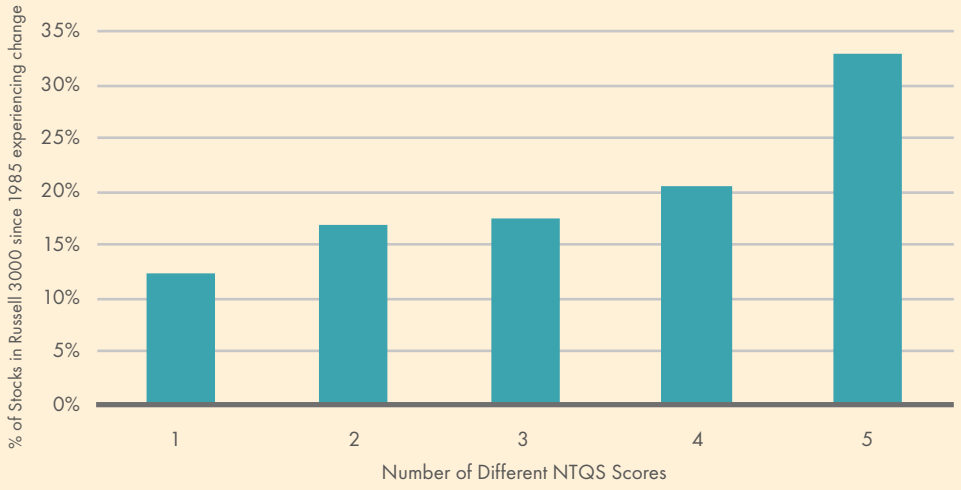
NTQS	Quality Quintile	Average Ex Ante Volatility	Average Ex Post Volatility	Difference
High Quality	Q1	24.8%	20.8%	4.0%
	Q2	23.8%	22.6%	1.2%
	Q3	23.0%	23.1%	-0.1%
	Q4	24.5%	24.9%	-0.3%
Low Quality	Q5	26.0%	30.1%	-4.2%

Source: Northern Trust Quantitative Research

While high-quality stocks clearly have lower ex post volatility, these stocks were not always of high quality. Because ex ante volatility is measured historically, this measurement could include periods where stocks were of lower quality and, hence, higher volatility. Thus, taking a historical view of volatility could overstate the expected ex post volatility. The converse is also true for low quality; low-quality names were not always low-quality so their history includes periods of relatively low volatility. Going forward, we would expect the volatility of low-quality stocks to exceed their historical average.

All this is, of course, predicated on stocks changing quality scores. Chart 8 shows the relative frequency of NTQS score changes since 1985. Note that more than one third of all stocks in the Russell 3000 have held all five quality scores (1 to 5) at some point in their history and less than 13% of stocks have held only a single score. Clearly, quality scores can be very dynamic.

Chart 8: Frequency of Distinct NTQS Scores Since 1985 (Russell 3000)



Source: Northern Trust Quantitative Research

While this suggests that stock volatilities are not stable due to changing quality, a similar analysis shows that stock correlations are also unstable for the same reason. Although the full description of our empirical tests and results is beyond the scope of this paper, we note that high-quality stocks tend to have more stable correlations than low-quality and that, due to the historical averaging effect, ex ante portfolio volatility which relies on these correlations as an input will again overstate ex post volatility. The total impact on ex post volatility is a combination of the instability of individual stock volatilities and correlations.

CONCLUSIONS

- Higher quality stocks—as identified by the Northern Trust Quality Score (NTQS) definition—tend to outperform lower quality stocks and do so with considerably less risk. This phenomenon is seen across domestic, international developed and emerging markets.
- Unlike value or size, there is no single generally accepted definition of an equity quality factor. Although a wide range of attempts have been made to define quality, their ability to capture the quality phenomenon has been extremely varied.
- Much of the difficulty in defining quality stems from the lack of a theoretical justification. Classic models such as CAPM suggest the quality phenomenon should not exist so we lack a practical framework with which to form a definition. However, these models make restrictive assumptions such as uniformity of investor risk posture.
- On the other hand, our model of heterogeneity in investors' views toward risk does a better job of explaining the quality phenomenon as well as the asymmetry of returns to quality during crises and recessions.
- Heterogeneity of risk postures suggests risk-seeking investors drive up the price of low-quality/high-risk stocks until their expected values are negative. Risk-averse investors gravitate toward low-risk stocks where positive expected values, i.e., positive risk premia, are an equilibrium condition.
- With this guidance we can define quality as those features of a company that appeal to risk-averse investors – a definition that is inherently multidimensional and will vary by market segment.
- With metrics encompassing multiple dimensions of quality, the NTQS has performed exceptionally well relative to other alternative definitions. Implementation of Northern Trust's quality philosophy, while integrated into the consistent framework that underlies our active equity products, varies by strategy and market segment.
- Tactically rotating in and out of quality is impractical since "junk rallies" are tied directly to macroeconomic cycles and are, hence, even more difficult to predict. Further, most of the return for holding low-quality names is concentrated into a few, relatively brief periods which are easily missed with factor rotation.
- The quality and low-volatility phenomenon are only partially related. While there is strong overlap between low quality and high-volatility stocks, there is only a weak relationship between high quality and low-volatility stocks. Thus, the two phenomena are fundamentally different.

If you would like to learn more about using the Northern Trust Quality Score within your equity portfolio, please contact your local Northern Trust relationship manager, or visit northerntrust.com.

ENDNOTES

1 Robert Novy-Marx, "The Other Side of Value: The Gross Profitability Premium," The Journal of Financial Economics,108, no. 1 (2013).

2 Eugene Fama and Kenneth French., "Dissecting Anomalies," The Journal of Finance 63, no. 4 (2008), 1653-1678.

3 Dimensional Fund Advisors, "Dimensional's Growth Portfolios," Dimensional Fund Advisors White Paper.

4 MSCI, MSCI Quality Indices Methodology (2012).

5 Russell Investments, Russell U.S. Equity Indices (2013).

6 Standard and Poors, S&P 500 Quality Rankings Index, Index Methodology (2011).

7 AQR Capital Management, A New Core Equity Paradigm, Using Value, Momentum, and Quality to Outperform Markets, (2013).

8 Joseph Piotroski, "Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers," Journal of Accounting Research 38 (2000).

9 Calvert Investments, Quality: The Third Dimension (2012).

10 Regions Investment Management, "Equity strategy of Regions Investment Management," http://www.regions.com/commercial_banking/rim_equity_strategy.rf (accessed Aug. 2013).

11 The model referenced in this paper is one of several quality implementations that Northern Trust employ. For the purposes of this paper, we are using a proprietary quality score that is typically implemented for broad-cap mandates.

12 Dartmouth College, "Fama/French Portfolios," http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html (accessed Aug. 2013).

13 William F. Sharpe, "Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk," The Journal of Finance 19, no. 3 (1964).

14 John Lintner, "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets," Review of Economics and Statistics 47, no. 1 (1965).

15 A quantile regression of order 2 was utilized and all polynomial parameters, including those at the 99th and first percentiles were statistically significant at the 99% confidence level. Standard errors were bootstrapped using 1000 iterations.

16 David Blitz and Pim Van Vliet, "The Volatility Effect: Lower Risk Without Lower Return," Journal of Portfolio Management (2007).

17 Andrew Ang and Robert J. Hodrick, Yunhang Xing, and Xiaoyan Zhang, "High Idiosyncratic Volatility and Low Returns: International and Further U.S. Evidence," Journal of Financial Economics 91, no. 1 (2009).

18 Malcolm Baker, Brendan Bradley and Jeffrey Wurgler, "Benchmarks as Limits to Arbitrage: Understanding the Low-Volatility Anomaly," Financial Analysts Journal 67, no. 1 (2011).

NOTES

IMPORTANT INFORMATION
This material is directed to eligible counterparties and professional clients only and should not be distributed to or relied upon by retail investors.

LEGAL, INVESTMENT AND TAX NOTICE.
Information is not intended to be and should not be construed as an offer, solicitation or recommendation with respect to any transaction and should not be treated as legal advice, investment advice or tax advice. Clients should under no circumstances rely upon this information as a substitute for obtaining specific legal or tax advice from their own professional legal or tax advisors. All material has been obtained from sources believed to be reliable, but the accuracy, completeness and interpretation cannot be guaranteed. The views expressed are those of the authors as of the date noted and not necessarily of the Corporation and are subject to change based on market or other conditions without notice.

There is no guarantee that the investment objectives of any fund or strategy will be met. Risk controls and models do not promise any level of performance or guarantee against loss of principal. Past performance is no guarantee of future results. Returns of the indices also do not typically reflect the deduction of investment management fees, trading costs or other expenses. It is not possible to invest directly in an index. Indices are the property of their respective owners, all rights reserved.

Northern Trust Corporation, Head Office: 50 South La Salle Street, Chicago, Illinois 60603 U.S.A. Incorporated with limited liability in the U.S. Products and services provided by subsidiaries of Northern Trust Corporation may vary in different markets and are offered in accordance with local regulation. This material is directed to professional clients only and is not intended for retail clients. For Asia-Pacific markets, it is directed to institutional investors, expert investors and professional investors only and should not be relied upon by retail investors. **For legal and regulatory information about our offices and legal entities, visit northerntrust.com/disclosures.** The following information is provided to comply with local disclosure requirements: The Northern Trust Company, London Branch; Northern Trust Global Services Limited; Northern Trust Global Investments Limited. The following information is provided to comply with Article 9(a) of The Central Bank of the UAE's Board of Directors Resolution No 57/3/1996 Regarding the Regulation for Representative Offices: Northern Trust Global Services Limited, Abu Dhabi Representative Office. Northern Trust Global Services Limited Luxembourg Branch, 2 rue Albert Borschette, L-1246, Luxembourg, *Succursale d'une société de droit étranger* RCS B129936. Northern Trust Luxembourg Management Company S.A., 2 rue Albert Borschette, L-1246, Luxembourg, *Société anonyme* RCS B99167. Northern Trust (Guernsey) Limited (2651)/Northern Trust Fiduciary Services (Guernsey) Limited (29806)/Northern Trust International Fund Administration Services (Guernsey) Limited (15532) Registered Office: Trafalgar Court Les Banques, St Peter Port, Guernsey GY1 3DA.

IRS CIRCULAR 230 NOTICE: To the extent that this communication or any attachment concerns tax matters, it is not intended to be used, and cannot be used by a taxpayer, for the purpose of avoiding any penalties that may be imposed by law. For more information about this notice, see http://www.northerntrust.com/circular230.

NOT FDIC INSURED	May lose value	No bank guarantee

Northern Trust