The Current State of Quantitative Equity Investing

Ying L. Becker
Ying L. Becker is professor of finance at Suffolk University and was formerly managing director at State Street Global Advisors.

Marc R. Reinganum
Marc R. Reinganum serves on the advisory board at Driehaus Capital Management and was formerly senior managing director at State Street Global Advisors and the Mary Jo Vaughn Rauscher Chair in Financial Investments at Southern Methodist University.

Introduction: Risk and Return

The current approaches and products of quantitative equity investing stand on the shoulders of major theoretical and empirical contributions in financial economics. At the root of disciplined, modern investment processes are two intuitive concepts: risk and return. The notion of total return is obvious—price appreciation plus any dividend payments. Risk is not so straightforward. Indeed, in *Risk, Uncertainty, and Profit*, Knight (1921) distinguished between risk and uncertainty. In essence, uncertainty involves environments in which investors cannot articulate potential outcomes or the likelihood of those outcomes. In contrast, risk is much more precise, like a roulette wheel. The possible outcomes are well specified and the likelihood of each outcome is known, but in advance, an investor does not know which outcome will be realized. Quantitative methods rely on this view of risk, although in the literature, risk and uncertainty are often used interchangeably. In the early 20th century, however, links between risk and asset pricing were not established. Economic analysis dealt with the concept of expected utility. For example, in "The Utility Analysis of Choices Involving Risk," Friedman and Savage (1948) offered an explanation of why an individual may both gamble and buy insurance based on the shape of her utility curve.

Models explicitly linking risk and return began to blossom in the 1950s. Markowitz (1952) famously wrote about portfolio selection in a

mean–variance framework. He defined an efficient portfolio as one that minimizes variance for a given level of expected return and maximizes expected return for a given level of variance. Investors should consider only portfolios along the efficient boundary, although the precise portfolio of choice depended on an investor's utility preferences. Markowitz established relationships between risk and return at a portfolio level but not at the level of individual assets. In a somewhat different context, Tobin (1958) used similar thinking to analyze liquidity preferences. Another major theoretical development occurred shortly afterward when Arrow and Debreu (1954) proved the existence of a unique general equilibrium. Their work allowed for state pricing, in which case a security would pay one unit in that state and zero otherwise, a paradigm that led to much theoretical work in financial economics. Viewed from the vantage point of a company, Modigliani and Miller (1958) modeled a partial equilibrium world in which the cost of capital of a company with risky projects was independent of capital structure, although the equity component of return may contain an additional premium related to financial risk arising from leverage. In more aggregate macro models, risk and return were being explicitly linked. The importance of these contributions can be gleaned by looking at the list of Nobel Prize in Economic Sciences winners from this period.

Modern Portfolio Theory and Asset Pricing

After Markowitz (1952) developed portfolio theory, it took more than a decade for his insights to be used in creating asset pricing models at the level of individual risky securities. Sharpe (1964), Lintner (1965), Mossin (1966), and Treynor (1962) developed what came to be called the "capital asset pricing model" (CAPM). This model had alluring simplicity and powerful insights. For individual securities, it demonstrated that there is a cross-sectional relationship between expected returns and security risk as measured by beta, the covariance between the security return and the market return scaled by the variance of the market return. More formally,

$$\mathrm{E}(R_i) = R_f + \beta_i\left[\mathrm{E}(R_M) - R_f\right], \tag{1}$$

where $\mathrm{E}(R_i)$ is the expected return of risky asset i; $\mathrm{E}(R_M)$ is the return of the market portfolio, the cap-weighted portfolio of all risky assets; R_f is the risk-free rate of interest; β_i is the relative risk of security i in the market portfolio (beta), which is equal to $[\mathrm{cov}(R_i, R_M)/\mathrm{var}(R_M)]$; and $\mathrm{E}(R_M) - R_f$ is often called the "market risk premium."

Although the modeling required much heavy lifting, the end result linked risk and return in a straightforward way: The expected return of a risky asset equals the risk-free rate of interest plus a risk premium, where the risk premium is proportional to the asset's beta. In this framework, there may be many risky events that could affect the realized returns of securities, but only beta risk is systematically priced. Stated differently, the cross section of security expected returns should depend on only betas and nothing else. No risks except for beta, a covariance measure of risk, are meaningful in the pricing of risky securities because other risks can be "diversified away."

The market portfolio also plays a pivotal role in the CAPM. Indeed, according to the CAPM theory, the risky portfolio each and every investor holds is the market portfolio. The market portfolio is efficient in terms of how Markowitz described efficiency. Furthermore, it can be shown to be the portfolio that maximizes the Sharpe ratio, where

$$\text{Sharpe ratio} = \frac{\mathrm{E}\left(R_p\right) - R_f}{\sigma\left(R_p\right)}. \tag{2}$$

That is, the Sharpe ratio is a portfolio's expected return in excess of the risk-free rate, scaled by the standard deviation of the portfolio return. The market portfolio in the CAPM is the unique portfolio that attains the maximum value of the Sharpe ratio and offers investors the best possible risk–return trade-off. It is the only portfolio of risky assets that all investors need and want.

The theoretical advances of the 1960s in asset pricing were accompanied by similar advances in computer technology and "big data." For example, the IBM System/360 mainframe computer was introduced in 1964 and began delivery in 1965. Concurrently, the first big dataset in finance was created at the University of Chicago. As Fisher and Lorie (1964) explained, the Center for Research in Security Prices (CRSP) was created to transcribe into computer-readable form monthly data on individual NYSE companies dating back to 1926. These data included security prices, capital changes, and dividends—all the data required to compute total return. With large-scale data, computing power, and a theoretical model in hand, academic research empirically tackled the question of whether the CAPM is a good quantitative approximation to the observed behavior of stock prices. This early period of testing the CAPM perhaps culminated with the seminal study of Fama and MacBeth (1973) that both set an enduring standard for methods to test cross-sectional relationships in equity markets with time-series data and concluded that the null hypothesis of the CAPM could not be rejected.

Although most empirical interest in the CAPM focused on the cross-sectional relationship described in Equation 1, the model was in fact tested with time-series data. Apart from potential econometric estimation complications, understanding the time-series properties of security returns became important too. Earlier statistical work by Working (1934), Kendall (1953), and Roberts (1959) suggested that security and commodity prices behaved like cumulative series of random numbers, a concept popularized by Malkiel (1973) in his book *A Random Walk down Wall Street*. The serial independence of price changes over time led to another key quantitative insight: the efficient market hypothesis (EMH). According to the EMH, price changes over time are uncorrelated because current prices "fully" reflect all relevant information. Fama (1970) clearly articulated this hypothesis and refined the notion of "fully" by classifying tests as weak form, semi-strong form, or strong form based on the type of information used to test the hypothesis.

The necessary introduction of time-series data into cross-sectional tests of relationships in asset pricing models leads to an unavoidable ambiguity in interpreting and using the empirical results. The reason is that empirical research actually tests a joint hypothesis—that the cross-sectional model is correct and that markets are informationally efficient. For example, if empirical tests reject the cross-sectional relationship between expected returns and betas in Equation 1, is it because the CAPM is misspecified? Or is it because markets are informationally inefficient?[1] Similar questions can be and are asked of even the most contemporary research. Trying to discern between the two is often a matter of examining the reaction of stock prices to certain events, such as earnings releases (for an early study in this vein, see Ball and Brown 1968).

Anomalies: The Precursors to Factors

By the late 1970s and early 1980s, cracks in the CAPM and the efficient market hypothesis were beginning to surface in academic research. Basu (1977) reported that portfolios of low-P/E stocks earned on average about 7% more annually than portfolios of high-P/E stocks even after adjusting for estimates of CAPM betas, at least during the 14-year period from 1957 to 1971. Basu interpreted his results as a rejection of the semi-strong form of the efficient market hypothesis, insofar as the information contained in P/Es did not appear to be "fully reflected" in prices. Interestingly, Basu chose not to attribute his results to any shortcoming in the CAPM, a reflection perhaps of how deeply ingrained the CAPM had become as perceived truth in the academic community.

[1]Roll (1977) questioned whether tests of asset pricing models were ever empirically possible.

Banz (1981) and Reinganum (1981a) changed the tenor of the discussions regarding the CAPM. Both researchers reported that differences in the average returns of portfolios grouped by company size (as measured by stock market capitalization) could not be explained by differences in estimated betas. In particular, portfolios of small-cap stocks outperformed portfolios of large-cap stocks on a beta risk-adjusted basis—the so-called size effect. Reinganum further reported that P/Es could not explain the size effect, because the size effect was present even after controlling for P/Es. Perhaps the enduring contribution of Banz and Reinganum was that both attributed their research results to a misspecification of the CAPM rather than a market inefficiency. A misspecified CAPM meant that size proxied for some missing risk factor(s) and that the cross-sectional relationship between expected returns and market betas was incomplete; other risk factors seemed to be needed. At the time of publication, Banz's and Reinganum's results were viewed as anomalous, tentative, and not necessarily correct. In retrospect, the course of modern quantitative equity investments had been altered.

Other anomalous evidence began appearing in the literature. For example, Keim (1983) found that the size effect had a strong seasonal component concentrated in the month of January. This result certainly challenged the belief that stock returns were independently and identically distributed and suggested that a simple seasonal dummy variable could help predict stock returns. Latané and Jones (1977) reported that quarterly standardized unexpected earnings (based on reported earnings) were reflected in stock prices only with a lag. Unlike Banz's and Reinganum's results, this result was centered on a specific event—the reporting of quarterly earnings—and eventually dissipated, a pattern more likely to be market inefficiency. In a somewhat similar vein, Givoly and Lakonishok (1979) concluded that revisions in analysts' forecasts of earnings could be used to earn abnormal returns in the two months following the release—a violation of the semi-strong form of the efficient market hypothesis. Rosenberg, Reid, and Lanstein (1985) reported that book-to-price ratios could help investors exploit pricing errors after controlling for several "risk indexes." Although the authors considered this finding a market inefficiency, the book-to-price ratios are not events like earnings releases or revisions in analysts' forecasts. The initial foundations of the CAPM and the efficient market hypothesis were showing cracks that might be exploited by investors.

The Age of Factor Investing

The era of factor investing might be traced back to Ross (1976), who developed a theory of security pricing with multiple factors. Ross's arbitrage pricing

theory (APT) was purely theoretical and did not specify how many factors were appropriate, nor did it label what the factors might be.[2] The effective era of equity factor investing, however, might more appropriately be attributed to the empirically based three-factor model explicated by Fama and French (1993). Fama and French essentially combined previous research on the size effect, the value effect, and the overall market factor into one cross-sectional equation. Unlike the CAPM, this model was not explicitly anchored in a theory but did have much empirical evidence to support it. Fama and French suggested that, at least empirically, a three-factor model better fits the cross section of average equity returns than does the CAPM. The three factors are MKT (the difference in return between the cap-weighted market portfolio and the risk-free rate of interest); SML (the difference in returns between a portfolio of small-cap stocks and a portfolio of large-cap stocks); and HML (the difference in returns between a portfolio with high book values of equity relative to market values of equity [BE/ME] and a portfolio with low BE/ME). This paradigm became the baseline standard for much quantitative equity work that followed it.[3]

Following Fama and French (1993), the floodgates for factor investing seemed to open. For example, Jegadeesh and Titman (1993) reported that over 3- to 12-month time horizons, an investor who purchased stocks that had performed well in the past and sold past poor performers would earn significant abnormal returns. Stated differently, over short and intermediate time horizons, stock returns exhibit momentum. More recently, Geczy and Samonov (2016) documented momentum in more than two centuries of price data. Carhart (1997) suggested that momentum should be considered a common factor, and indeed, later studies often spoke of a four-factor model—one with the three Fama–French factors plus momentum. De Bondt and Thaler (1985) reported the opposite of short- to intermediate-term momentum at longer (three- to five-year) investment horizons; specifically, prior stock market losers tend to become subsequent winners relative to prior winners De Bondt and Thaler strongly suggested that the stock market overreacts to unexpected and dramatic news events and that this overreaction can account for a return differential between winners and losers of 25% over three years.

Momentum and reversal are often characterized as expressions of investor sentiment. De Bondt and Thaler (1985) may very well be considered the seminal study in "behavioral finance," because the authors rooted their work in

[2]Connor and Korajczyk (1988) suggested that an asymptotic principal components technique for estimating and testing the APT yields better results than the CAPM does.

[3]Jacobs and Levy (1988) looked at a variety of anomalies within a multiple regression analysis to disentangle the various effects but did not explicitly link their results to asset pricing.

the psychology of investor decision making.[4] Understanding investor sentiment became a key aspect to model empirical asset pricing, regardless of whether sentiment represented behavioral biases.[5] For example, Lakonishok, Shleifer, and Vishny (1994) argued that the value effect arose because these strategies exploited suboptimal choices of the typical investor; in essence, investors overpay for "glamour" stocks, perhaps overestimating their future growth prospects. In subsequent research, this expectational "error" was attributed in part to the finding that value stocks tended to have earnings surprises that were systematically more positive than those of glamour stocks (see La Porta, Lakonishok, Shleifer, and Vishny 1997). But this explanation could not be a complete one, because the return differences between value and glamour stocks persisted much longer than did the magnitude of differences in the earnings surprises.

Apart from the market, size, value, and momentum factors, other credible factors muscled their way into the empirical equity asset pricing world. Amihud and Mendelson (1991) argued that liquidity—as measured by, say, the bid–ask spread—is not only an issue for market microstructure. Rather, they saw illiquidity premiums as a significant influence on expected returns because even seemingly small differences in transaction costs could make a meaningful difference in asset values. Pástor and Stambaugh (2003) reported that expected returns are cross-sectionally related to sensitivities of stock returns to fluctuations in aggregate liquidity; the difference in annual returns between high-sensitivity stocks and low-sensitivity stocks was about 7.5% after controlling for the market, size, value, and momentum factors. Taking a different tack, Sloan (1996) made a convincing case that investors pay too much attention to bottom-line earnings numbers and not enough attention to the quality of those earnings, as reflected in accruals, a noncash component of earnings. Sloan found that accruals and subsequent stock returns were negatively related: High-accrual companies underperformed low-accrual companies over the course of at least a year, and about half of the annual outperformance was concentrated around the four quarterly earnings announcements.

The search for better factors continues in quantitative equity investing. Fama and French (2015) revisited their three-factor pricing model from the early 1990s.[6] In this more recent work, they proposed a five-factor model that includes the market premium (MKT), size (SML), and value (HML), plus

[4]Kahneman and Tversky wrote extensively about potential biases and heuristics used in the psychology of decision making under risk (for example, see Tversky and Kahneman 1974; Kahneman and Tversky 1979, 1984).

[5]Barberis, Schleifer, and Vishny (1998) developed a model that attempted to explain both investor overreaction and investor underreaction.

[6]Grover and Kizer (2016) estimated the cost of these factor exposures in well-known mutual funds and exchange-traded funds.

RMW (the difference in returns between portfolios with robust versus weak operating profitability) and CMA (the difference in returns between portfolios with low [conservative] and high [aggressive] investment, where investment is measured by the change in total assets). In part, Fama and French (2015) was a reaction to new research findings. For example, Novy-Marx (2013) found a gross profitability premium in the cross section of average returns, where gross profitability is simply defined as revenues minus cost of goods sold, all scaled by total assets. Novy-Marx reported that the gross profitability premium is equally strong as the book-to-market premium; further, gross profitability is complementary to value and diversifies the risk of value.[7] In addition, research has documented that substantial increases in capital investments tend to be associated with subsequent negative abnormal returns (see Titman, Wei, and Xie 2004); empire building might be hazardous to shareholder wealth! This general sentiment is echoed in work by Cooper, Gulen, and Schill (2008), who reported an inverse relationship between asset growth and subsequent stock returns. Thus, as a result of research since the early 1990s, the venerable three-factor Fama–French model has now been modulated to a five-factor version.

Perhaps the most puzzling cross-sectional anomaly is risk itself. Intuition and theory certainly suggest that average returns and systematic risk should be positively correlated: Higher systematic risk should be associated with higher average returns, but all other risks should not be priced. Ang, Hodrick, Xing, and Zhang (2006) reported that companies with high idiosyncratic risk relative to the Fama–French (1993) model had abysmally low average returns. This finding survived after the authors controlled for size, book-to-market ratio, leverage, liquidity, volume, turnover, bid–ask spreads, co-skewness, and dispersion in analysts' forecasts. In theory, this risk should not be priced at all. An equally startling result was published by Clarke, de Silva, and Thorley (2006), who found that, on the basis of optimizations from 1968, the minimum-variance portfolios had about three-fourths the risk of the cap-weighted market portfolio but higher average returns. In other words, the low-volatility portfolio had a better Sharpe ratio than the market portfolio had. Li, Sullivan, and Garcia-Feijóo (2016) and Baker (2016) concluded that this empirical regularity is a systematic mispricing of risk.

In CAPM theory, no portfolio has a better Sharpe ratio than the market portfolio. The anticipated link between risk and return seemed to be broken: Low-volatility portfolios were earning higher returns relative to high-volatility portfolios. Baker, Bradley, and Wurgler (2011) analyzed the low-risk

[7]Ball, Gerakos, Linnainmaa, and Nikolaev (2016) found that cash-based operating profitability, which excludes accruals, outperforms gross profitability that includes accruals.

anomaly, where risk is defined as both total volatility and beta. That is, unexpectedly large returns are seen in the data with respect to low-volatility stocks (versus high-volatility ones) and with respect to low-beta stocks (versus high-beta ones). Sometimes the inverted relationship between betas and returns is called "betting against beta" (BAB; see Frazzini and Pedersen 2014). This reversal in the relationship between risk and return is puzzling. Indeed, Baker et al. (2011) believe that this may be the greatest anomaly in finance.[8]

Investor interest in low volatility and betting against beta seems to have increased since the dot-com bubble crash in the early 2000s and the global financial crisis of 2008. Investors seem to be acutely attuned to "downside protection," which low-volatility and low-beta strategies tend to offer, especially in long-only portfolios. Yet the supposedly unusual relationship between risk and return was noted many years earlier. Black, Jensen, and Scholes (1972) reported that the cross-sectional relationship between CAPM betas and average returns was flatter than the theory would predict; contrary to the CAPM, high-beta securities had significantly more-negative returns and low-beta securities had significantly more-positive returns than the theory would predict. Haugen and Heins (1975) found little support for the notion that risk premiums had manifested themselves in realized rates of return over long periods; the authors also reported that portfolios with low-variance stocks had earned greater average returns than portfolios with higher-variance stocks. Reinganum (1981b) reported that estimated betas were not systematically related to average returns across securities and that high-beta securities and portfolios with widely different estimated betas possessed statistically indistinguishable average returns. He cautioned that the CAPM may lack significant empirical content. Even a close look at the evidence in Fama and MacBeth (1973) suggests that the cross-sectional relationship between average returns and security variance may not have been completely extinguished by beta. Whatever the reason, after a nearly 30-year hiatus, low risk seems to be back on the table as a bona fide investment tool and strategy, supported by more empirical work.

If your head is spinning with the proliferation of factors beyond market, size, and value, you are not alone. In his American Finance Association Presidential Address, Cochrane (2011) described the situation as a "zoo of new factors." Indeed, Harvey, Liu, and Zhu (2016) analyzed 313 published papers that studied cross-sectional return patterns. The backlash against the proliferation of potential factors emerged well before their exponential growth in the literature. Lo and MacKinlay (1990) raised the issue of data snooping in tests of asset pricing models. Lo and MacKinlay argued that

[8]Auer and Schuhmacher (2015) confirmed the BAB phenomenon even among the 30 large, liquid stocks in the Dow Jones Industrial Average from 1926 through 2013.

newly discovered relations must be weighed in view of past inferences; the conditions underlying standard statistical tests may be violated. This concern about overfitting and potential solutions to it has only grown over time. Hsu, Kalesnik, and Viswanathan (2015) proposed a three-step heuristic framework to assess factors: (1) Does a factor persist across time and geographies? (2) Does a factor persist with respect to sensible perturbations in its definition and construction? (3) Is a factor validated and vetted numerous times in top-tier journals? More formally, Harvey and Liu (2015) developed a measure called the "haircut Sharpe ratio" that takes into account multiple testing and data mining. Harvey et al. (2016) provided a nice discussion and analysis of the issue of multiple testing and suggested that a t-statistic of 3.0 is a good cutoff value for testing factors; they argued that a t-statistic of 2.0 is no longer appropriate. Suhonen, Lennkh, and Perez (2017), using a sample of 213 "alternative beta" index funds, found that performance deteriorated substantially after a strategy went live, compared with performance during its backtest period.

In short, there is a potpourri of potential and reasonable risk factors that investors may consider. The frustration arises, perhaps, because there is currently no one right answer as with the CAPM, in which the cap-weighted market portfolio is the optimal and right choice for each and every investor. But the lack of a single right answer may be an insight in itself. Perhaps the answer is that there is no one right choice for each and every investor, and perhaps the future of factor investing lies in illuminating different investor-appropriate paths.

The Prime Factor and Smart Factors

In the beginning, there was one factor and it was the market factor. A cap-weighted return of all risky securities, the market factor was reasonably easy to calculate, and it was buttressed by an elegant theory that asserted no other factor need be worshipped. The final 25 years of the 20th century was a golden era for this factor, in terms of both performance and industry adoption (see Reinganum 2014). Indeed, the market factor anchored quantitative equity in terms of performance attribution, compensation schemes, and terminology. The cap-weighted market portfolio became synonymous with "passive" investing; "active" portfolios were defined as portfolios whose security weights differed from those of cap weights. In many institutions, the view of risk also shifted subtly from the total risk of the market portfolio to tracking error risk—the variability in differential returns between an investor's portfolio and the market portfolio. Such metrics as the information ratio were developed to assess whether deviations from the benchmark return were

worth the tracking error risk (for example, see the classic book *Active Portfolio Management* by Grinold and Kahn 2000).

Investor appetite for factor investing seems to have increased so far in the 21st century. Invesco PowerShares Capital Management LLC (2015) reported that factor-based investing is gaining traction in the institutional community and that institutions plan to increase their use of smart-beta exchange-traded funds (ETFs) more than any other category. Morningstar (2014) also reported that factor investing is a fast-growing segment in the marketplace and offered its analysis of what it termed "strategic beta." Hill (2016) argued that these new products are the next evolutionary step in "the triumph of indexing."

In part, the increasing acceptance of factor-based quantitative equity investing may be driven by the relatively low economic returns that the prime factor (the market return) has delivered relative to investor and actuarial expectations. Part of the allure of factor investing is its potential to improve returns at lower cost. In a well-publicized study for the Norwegian Government Pension Fund, Ang, Goetzmann, and Schaefer (2009) recommended that this very large sovereign fund adopt a factor risk premium approach to gain exposures that were otherwise attained by its group of active managers. The report suggested that appropriate factor exposures could be achieved by moving assets out of many actively managed strategies without detriment and with lower costs (for example, see Kidd 2014). In a similar vein, Bender, Briand, Nielsen, and Stefek (2010) suggested that strategy and style risk premiums could be used as the building blocks of a diversified portfolio. Kahn and Lemmon (2015) suggested that asset owners may be disappointed by their aggregate active performance because they may be overpaying by 43% on average for their active risk, given that this part of risk could be obtained through low-cost factor solutions.

The discussion about factors in the previous section was a bit vague in some respects. Any portfolio manager knows that, in the final analysis, implementation of a strategy and its performance are about portfolio weights. The same group of securities can be combined into two portfolios with very different risk and return characteristics by weighting the same securities differently. For example, consider a portfolio formed on the basis of, say, sensitivities to the Fama–French HML factor. What does this exactly mean? One can rank securities from high to low using this metric, but this ranking does not create a portfolio. The portfolio is a set of weights, and so one needs to devise an approach to construct the portfolio weights. Thus, factor investing, whether for one factor or multiple factors, involves portfolio construction, and in the end, portfolio construction ends up with one set of security weights.

With the prime factor—the market—portfolio construction was straightforward: Just cap-weight the securities. We term the other factors "smart

factors" to correspond to the current industry practice of referring to "smart beta." The smart factors can be implemented in a variety of weighting schemes. But the smart factors and the strategies based on them all scream out at least one common chorus: "I am not cap-weighted!"

If not cap-weighted, then what? Arnott, Hsu, and Moore (2005) developed indexes that explicitly avoid price and market-cap metrics in the weighting scheme. Instead, the authors used such items as gross revenue, equity book value, and total employment to calculate security weights. Arnott et al. (2005) called these "Fundamental Indexes" and argued that they deliver superior mean–variance performance relative to cap-weighted indexes; that is, they have higher Sharpe ratios.[9] Perold (2007) argued that cap-weighting is not an intrinsic drag on performance and that fundamental indexing is actually a form of value investing. In a similar vein, Chow, Hsu, Kalesnik, and Little (2011) argued that the outperformance (relative to a cap-weighted index) of most alternative equity strategy indexes can be attributed to their value and size factor exposures, a view basically corroborated by Dubil (2015). Van Gelderen and Huij (2014) found that mutual funds that tilted toward low-volatility, value, and small-cap factors outperformed mutual funds that did not implement the factor approach.[10]

Qian (2005) suggested another approach for portfolio construction and coined it "risk parity." Qian noted that market-cap allocations are not equivalent to risk allocations. For example, a stock/bond portfolio weighted 60/40 in terms of capital allocation had a risk allocation of 93/7. Qian suggested allocating weights such that each risky asset has the same risk allocation. He proposed that doing so would lead to better diversification and more portfolio efficiency, a conclusion later buttressed by his colleagues Sorensen and Alonso (2015). Like Arnott et al.'s Fundamental Indexes, Qian's risk parity approach yields weights that differ from market-cap weights. Fisher, Maymin, and Maymin (2015) argued that although risk parity is a fast and frugal heuristic, it tends to outperform both knowledge-intensive mean–variance approaches and knowledge-independent equal-weighted approaches.[11]

[9]In a couple of clever articles, Arnott and coauthors suggested that almost any weighting scheme of securities that differs from weighting by current market cap will outperform cap-weighted benchmarks (see Arnott, Hsu, Kalesnik, and Tindall 2013; Arnott, Beck, and Kalesnik 2015).

[10]Simon, Omar, Lazam, and Amin (2015) studied Shariah-compliant factors based on the Musharakah principle and reported superior results relative to cap-weighted indexes for securities in Malaysia from 2009 through 2013.

[11]Another line of research explored whether genetic programming algorithms might better articulate security weights (for example, see Becker, Fei, and Lester 2006).

Amenc, Goltz, Lodh, and Martellini (2014) nicely articulated a more nuanced approach to smart factor construction. They suggested that smart factor construction needs to balance exposures to desired and rewarded risks while mitigating and diversifying away as much unrewarded risk as possible. The result will, of course, be a non-cap-weighted portfolio. Not surprisingly, Amenc, Ducoulombier, Goltz, Lodh, and Sivasubramanian (2016) suggested that smart factor tilts be well diversified rather than highly concentrated among a limited number of securities. Of course, as Clarke, de Silva, and Thorley (2002) demonstrated much earlier, the impact of factors in realized returns will diminish as more constraints are placed on the portfolio construction process. Sorensen, Hua, Qian, and Schoen (2004) suggested some simple criteria that might assist factor construction in a sensible way. Bender and Wang (2016) as well as Clarke, de Silva, and Thorley (2016) pointed out that combining factor subportfolios is not an efficient way to capture the information content of multiple factors; bottom-up approaches using individual securities capture factor exposure more efficiently than top-down approaches do because bottom-up approaches better capture nonlinear cross-sectional interaction effects between factors.

The term "smart beta" was probably introduced into the quantitative equity management lexicon by Towers Watson, a global consulting firm, in 2013. In its 2013 report, Towers Watson suggested that exposures to different return drivers could be achieved without hedge-fund-like fees. This conclusion was echoed by Mladina (2015), who concluded that the risk and return profile of hedge funds could be explained by a mix of systematic risk factors. These factor exposures could be achieved simply and with transparency. The term "smart beta" itself has engendered some spicy comments. For example, Malkiel (2014) argued that smart beta is more about marketing than investing. He cautioned that the realizations of these smart factor premiums are not always positive. In "Beta as an Oxymoron," Anson (2015) argued smart beta isn't smart; it's dumb. If it were smart, providers would charge much higher fees. Qian, Alonso, and Barnes (2015) echoed the view that smart beta is surely a misnomer and can't be that smart *now* because quantitative and fundamental managers have been using these factors for years. Bogle (2016) remained unconvinced that smart beta has slain traditional cap-weighted indexing.

With all due respect to criticisms of the term "smart beta," smart-beta and smart factor investing are probably here to stay, as evidenced by the attention focused on them by large asset managers and advisory firms.[12] For example, Morningstar published its "A Global Guide to Strategic-Beta

[12]Vadlamudi and Bouchey (2014) questioned whether smart-beta solutions are actually smart for taxable investors on an after-tax basis.

Exchange-Traded Products" in 2014; the firm noted that "strategic beta" is its term for smart beta. Kahn and Lemmon (2016), both with BlackRock, discussed how smart-beta products are disrupting the investment management industry by providing an important component of active strategies in a low-cost, transparent, and rule-based way. Philips, Bennyhoff, Kinniry, Schlanger, and Chin (2015) of Vanguard published "An Evaluation of Smart Beta and Other Rules-Based Active Strategies," in which they concluded that an index should be constructed using market-capitalization weights. State Street Global Advisors (2016) published a piece titled "The Factor Revolution: Moving beyond Traditional Investment Models" to address the issue of investors' realizations that their desired outcomes are likely to be difficult to achieve with current approaches. Melas (2016) of MSCI asserted that factor investing will have a profound effect on long-term portfolio management. In short, from its humble and outcast beginning nearly 40 years ago, factor investing has now gone mainstream.

The Second Coming of Big Data and Technology in Quantitative Equity

The first era of big data in finance occurred in the 1960s with the creation of CRSP and Standard & Poor's Compustat tapes and mainframe computers. The data in this era were predominantly well structured, numeric, standardized, and curated, primarily consisting of company-specific information on stock prices, dividends, and capital changes, as well as officially filed financial and accounting information conforming to proscribed standards. Time-stamping information release, a necessary input for understanding the incorporation of information into stock prices, was not an easy or trivial matter and often required hand-collected datasets. Eventually, time stamping and event identification became easier as newswire services began to put press releases in computer-searchable databases.

Perhaps with the widespread acceptance of momentum as a factor, researchers and investors tried to better understand the drivers of sentiment. An early indicator of sentiment came from earnings forecasts made by financial analysts, although the early focus of this research was whether financial analysts made better earnings forecasts than did econometric models of actual earnings (for example, see Brown, Richardson, and Schwager 1987). Often this research used data supplied by I/B/E/S that aggregated the forecasts of individual financial analysts. The data were generally believed to be trustworthy because they came from bona fide financial firms, although they were not standardized because the earnings that analysts were forecasting were not

necessarily consistent with GAAP. Since these early measures of sentiment, the range of investor sentiment metrics has expanded greatly.

Big data is often discussed in terms of the four Vs: volume, variety, velocity, and (more recently) veracity. With advances in technology, the different types of information that quantitative equity might find useful has exploded. In practice, most "big data" analyses to date in quantitative equity have focused on unstructured data emanating from text-based sources, with various degrees of credibility and curation. Some of the early research in this area dates to the so-called dot-com era. For example, Tumarkin and Whitelaw (2001) studied the relationship between postings on a specific internet bulletin board (RagingBull.com) and stock prices during the period from April 1999 to February 2000 for a group of stocks classified in the internet service sector. The authors concluded that message board activity could not predict stock returns, consistent with the efficient market hypothesis. But they also noted that strong positive returns preceded days with unusual message board activity and strong positive opinions. Dewally (2003) examined the stock recommendations on two newsgroup sites in April 1999 and February 2001 and concluded that newsgroups provided no value to their readers in terms of predicting subsequent returns. A slightly more positive spin on the information content of internet stock message boards was reported by Antweiler and Frank (2004). These authors analyzed about 1.5 million messages posted on Yahoo! Finance and RagingBull.com during 2000 for 45 firms using computational linguistic algorithms (naive Bayes and support vector machine) to assign buy, hold, and sell tags for each message. The authors concluded that internet chatter is pertinent for predicting trading volume and volatility. Clarkson, Joyce, and Tutticci (2006) studied the market reaction to takeover rumors posted on the Australian internet discussion site HotCopper between May 1999 and March 2000. From this very narrowly defined event in a sample of 189 firms, the authors found intraday abnormal returns and trading volume in the 10-minute intervals around the posting of the rumors, driven mostly by firms not identified in the press in the preceding year. Das and Chen (2007) proposed a method using statistical and natural language processing techniques to classify opinions from internet stock message boards. The authors used their method to extract views on 24 high-tech stocks from messages in July and August 2001 and found some contemporaneous (but not predictive) relationships between message board activity and market variables.

The internet stock bulletin boards of the late 1990s and early 2000s certainly represented some elements of big data in terms of volume, variety, and velocity, but the sources were not typically verified or vetted in terms of expertise. Tetlock (2007) took a different approach by analyzing the relationship

between content in the *Wall Street Journal*'s "Abreast of the Market" column and stock market returns. Unlike stock market bulletin boards, the source of this content is easily identifiable and presumably a bit higher in veracity. Tetlock used the General Inquirer linguistic content program to quantify daily changes in the column over the 1984–1999 period. He reported that pessimism in the column's content did predict short-run downward price movements in the market, especially for smaller-cap companies. It was unclear, however, whether trading on this information would yield economic profits after accounting for transaction costs.

Tetlock, Saar-Tsechansky, and Macskassy (2008) extended the work of Tetlock (2007) by assessing the impact of negative words in all *Wall Street Journal* and Dow Jones News Service stories from 1980 to 2004 for S&P 500 Index companies.[13] This study reported the following: (1) The fraction of negative words can forecast low company earnings; (2) statistically, negative words predict negative abnormal returns on the next day (although transaction costs may wipe out the economic profit of a trading strategy); and (3) the predictability of earnings and returns from negative words is greatest for stories that focus on fundamentals (as measured by stories containing the word stem "earn"). In a study of 2.2 million news articles between 1989 and 2010, Hillert, Jacobs, and Muller (2014) suggested that the momentum effect is exacerbated by news coverage. That is, prior stock market winners with excessively high media coverage experience returns substantially greater than the returns of prior losers with excessive media coverage over about nine months. Most importantly for this study, the return differential between winners and losers is much smaller for companies with excessively low media coverage compared with high-media-coverage companies. The interest in news stories, sentiment, and stock returns has remained. For example, using a proprietary Thomson Reuters neural network measure of sentiment, Heston and Sinha (2017) found that daily news predicts stock returns for perhaps only one or two days, confirming previous research; news aggregated over a week has longer-lasting effects.[14] Unlike much of the previous research cited, however, this research offers no transparency into how tone or sentiment were actually constructed.

Moving from media to corporate-released information, Li (2008) analyzed the relationship between the readability of a company's annual report and its subsequent performance and earnings persistence. On the basis of the

[13]Tetlock (2010) studied an even larger dataset of news stories from 1979 to 2007 and focused on how news stories might resolve asymmetric information. Tetlock (2011) also analyzed investor reactions to stale information in news stories, where "stale" was defined as textual similarity in the previous 10 stories for a company.

[14]Using sentiment data from Thomson Reuters' News Analytics, Uhl, Pedersen, and Malitius (2015) claimed that filtered sentiment data could be used to perform tactical asset allocation.

Fog Index from computational linguistics and the length of reports for the period from 1994 to 2004, Li concluded that companies with easier-to-read annual reports tend to have more-persistent profits and that poor-performing companies have more-difficult-to-read annual reports. The *tone* of how companies voluntarily communicate with investors through press releases was studied by Henry (2008), who reported that abnormal returns increase as the tone of the earnings press release becomes more positive. Feldman, Govindaraj, Livnat, and Segal (2010) focused specifically on tone changes in the management discussion and analysis (MD&A) section of Forms 10-Q and 10-K. The authors concluded that tone changes do convey information not embedded in the regular financial statements. In particular, they concluded that tone changes are significantly correlated with short-window contemporaneous returns around SEC filing dates and with drift returns. Li (2010) isolated the forward-looking statements (FLSs) in the MD&A sections of 10-Q and 10-K reports to measure tone. This research found that FLS tone is correlated with several variables, including current performance, accruals, company size, return volatility, and company age. The author concluded that the Bayesian tone measure of FLSs is positively associated with future earnings.

In an analysis of 10-K filings, Loughran and McDonald (2011) developed a list of negative words that they considered more appropriate for financial analysis. According to the authors, the revised list, along with term weighting, should yield more informative measures of tone in financial documents. Jegadeesh and Wu (2013) used term weighting based on market reactions to 10-K filings from 1995 through 2010 downloaded from the SEC's EDGAR database. The authors reported that their measure of tone is significantly related to companies' market returns around their SEC filing dates; there is some initial underreaction to tone, but it quickly corrects within two weeks. The authors extended their method to IPO prospectuses and found a negative relation between tone and IPO underpricing.[15]

An interesting offshoot of corporate-released information is the management conference call. On these calls, company managers present prepared remarks, which are typically followed by a question and answer session with financial analysts. Brockman, Li, and Price (2015) extracted the linguistic tones of both managers and analysts during earnings conference calls. Perhaps not surprisingly, the authors found that manager tones are more optimistic than analyst tones; in addition, the market reacts much more attentively to analyst tones than to manager tones.

[15]Hanley and Hoberg (2010) examined the implications of the information content of IPO prospectuses using textual analysis on various aspects of the underwriting process.

This section began with a review of big data as gleaned from internet stock bulletin boards around 2000 and will come full circle by concluding with opinions expressed on social media. Bollen, Mao, and Zeng (2011) analyzed daily Twitter feeds using two mood-tracking tools (OpinionFinder and Google Profile of Mood States) and concluded that at least one metric of mood—calmness—improves the predictability of changes in the Dow Jones Industrial Average level. Using textual analysis, Chen, De, Hu, and Hwang (2014) extracted the tone of all opinion pieces between 2005 and 2012 from Seeking Alpha (SA), one of the biggest investment-related social media websites. Opinion content on SA is curated by a panel. The authors observed that the fraction of negative words in SA opinions and commentaries seems to have predictive power for individual stock returns over the ensuing three months. Azar and Lo (2016) measured Twitter sentiment around Federal Open Market Committee (FOMC) meetings between 2007 and 2014. Their evidence suggests that the information content of Twitter sentiment is predictive of market returns around FOMC meetings. In another Twitter-based study, Liew and Wang (2016) studied the cross-sectional relationship between sentiment in tweets and first-day IPO performance, from opening price to closing price. The authors did not calculate the sentiment measure themselves in a transparent way but, rather, used a proprietary metric provided by iSENTIUM, LLC. Based on this source, the study reports that prior-day sentiment can predict an IPO's first-day return and that there is a contemporaneous relationship between sentiment and IPO first-day returns. Liew and Budavari (2017) used a bullish–bearish sentiment bar indicator, filled in by users, from StockTwits. The authors found that for 15 stocks with a high volume of sentiment data, the "percent bullish" measure can add explanatory power to time series of daily returns beyond the five factors of Fama–French (2015). Karagozoglu and Fabozzi (2017) analyzed minute-by-minute, proprietarily calculated sentiment data from the commercial company PsychSignal in its Trader Mood product. The authors reported that this product and associated algorithms contain useful information about future stock market volatility.

Firms can employ, but also must contend with reactions on, social media.[16] For example, Lee, Hutton, and Shu (2015) explored crises caused by product recalls. Based on a sample of 405 recalls between 2000 and 2012, the evidence suggests that companies with social media messaging can attenuate the negative price consequences of the recall relative to companies with no social media presence. The attenuation benefits of social media declined as social media became more interactive. Indeed, the frequency of tweets by

[16]In April 2013, the SEC decided to allow companies to use social media to disclose key information.

disgruntled individuals exacerbates the negative price reaction, which can be offset somewhat by more-frequent tweeting by companies. Companies might also use Twitter to improve market liquidity for their publicly traded equity. Blankespoor, Miller, and White (2014), using a final sample of 85 technology companies, found that corporate tweets with links to their press releases lowered bid–ask spreads and increased market depth, presumably because the tweets reduced informational asymmetry.

The growing interest in big data in finance is perhaps illustrated by the fact that large financial firms have found it necessary to offer clients their views on this issue. These firms include Citi ("Big Data & Investment Management: The Potential to Quantify Traditionally Qualitative Factors," 2015), BNY Mellon ("Big Data and Investment Management: The Application of Data to Product Management and Client Satisfaction," 2015), Deutsche Bank ("Big Data Investment Management," 2016), and Goldman Sachs Asset Management ("The Role of Big Data in Investing," 2016).

To date, most published work on big data in equity management has focused on investor sentiment extracted from natural language processing algorithms applied to social media, official 10-Q and 10-K documents, press releases, and company conference calls.[17] The preponderance of published evidence indicates that to the extent that big data does contain useful sentiment information, it is for the most part short-lived in terms of profitable stock trading. Although these data may be quite relevant for market makers and trade desks, they do not seem to contain hidden, easy-to-exploit gems of information. Indeed, for long-term investors, it is not yet clear that big data per se is a big deal for their investment processes. As of this writing, we have yet to see refereed journal articles suggesting that big data and software can create long-term, persistent insights about quantitative equity management. Nonetheless, with so much digital data untagged and unexplored, this question remains open.[18]

Getting Dynamic

Although unstructured big data is one focus of current quantitative equity research, another focus may very well be dynamic factor models. In some ways, even with the explosion in potential factors, standard cross-sectional research

[17]Other suggested uses of big data include satellite imagery and microtransaction data. Conceptually, these potential uses might be considered as similar extensions to the work on economic links (Cohen and Frazzini 2008) and supply chains (Shahrur, Becker, and Rosenfeld 2010).

[18]Deutsche Bank (2016) reported that Google has indexed only about 0.01% of the accessible data on the internet.

between security returns and factors may be approaching a point with limited additional insights. Indeed, McLean and Pontiff (2016) even questioned whether academic research destroys stock return predictability and found evidence that it does, particularly after publication. But future quantitative equity insights may turn out to be centered much more on what factors are rewarded at given points in time than on what factors are rewarded on average over time. Some might label this approach "factor timing." As the search to outperform standard, passive benchmarks intensifies, some might view dynamic factor models as a reasonable investment approach. Of course, dynamic factor models and factor timing are not without critics and skeptics. For example, Asness (2016) opined that factor timing just might be a siren song.

The academic underpinnings for dynamic factor models can arguably be traced back to at least the 1980s. Ferson, Kandel, and Stambaugh (1987) reported evidence of a time-varying risk premium over the period 1963 through 1982 in common stock portfolios formed on the basis of market capitalization. The notion of time-varying risk premiums is also supported by Fama and French (1989), who found a risk premium in expected returns that varies with business conditions. Keim and Stambaugh (1986) also detected evidence of changing expected risk premiums. Ferson and Harvey (1991) concluded that time variation in risk premiums, as contrasted with time variation in betas, accounted for most of the return predictability. This list of early articles is not exhaustive but merely indicative that time-varying risk premiums were reported in the literature about 30 years ago. That they should be considered controversial now is somewhat puzzling. Perhaps the controversy is attributable partially to the observation that investors tend to implement timing decisions poorly, which causes realized returns to fall well short of proven strategy returns (see Hsu, Myers, and Whitby 2016).

Connor (1995) suggested classifying factors into three types: macroeconomic, fundamental, and statistical. For example, Boguth and Kuehn (2013) claimed that macroeconomic uncertainty affects asset pricing and that exposure to consumption volatility predicts future returns. Feldman, Jung, and Klein (2015) claimed that the Conference Board's Leading Economic Indicators could be used to create a time strategy that beats a simple buy-and-hold strategy.[19] The cyclically adjusted price-to-earnings ratio (CAPE; see Campbell and Shiller 1998), a fundamental-type factor with long-horizon predictability, continues to receive attention. For example, Siegel (2016) suggested ways to improve the predictive power of this approach by substituting NIPA (national income and product account) after-tax corporate profit data

[19]Bali, Brown, and Tang (2017) found that exposure to an economic uncertainty index is reflected in cross-sectional returns and labeled this risk an "uncertainty premium."

for GAAP earnings. Philips and Ural (2016) also investigated the CAPE and developed a list of recommendations to improve its efficacy. Hull and Qiao (2017) argued that the CAPE can be used to reduce sequencing risk in the decumulation phase of investing. Momentum, a statistical factor, continues to be a bedrock of factor investing, yet it doesn't always work and sometimes crashes dramatically, such as in 2009. Daniel and Moskowitz (2016) demonstrated that a dynamic momentum strategy can double the alpha and the Sharpe ratio of a static momentum strategy. Garcia-Feijóo, Kochard, Sullivan, and Wang (2015) concluded that there are cycles in low-volatility investing and that the performance of low volatility is time varying and influenced by the economic environment. Miller, Li, Zhou, and Giamouridis (2015) suggested a dynamic factor-weighting framework to respond to changes in factor predictability. Incorporating classification tree analysis, these authors concluded that their multifactor dynamic approach generated reward-to-risk ratios nearly four times greater than those generated by static approaches. Using more-conventional regression analysis, Reinganum, Becker, and He (2011) presented a dynamic multifactor model, conditioned on macroeconomic, fundamental, and statistical variables, that also significantly outperformed its fixed-weight, static counterpart.

Dynamic modeling is a current, promising area of quantitative research, with roots dating back to the 1980s and models of time-varying expected returns. Evidence of predictable returns can be found in very recent research as well (for example, see Hull and Qiao 2017). Closely related to dynamic models in quantitative equity investing are regime-shifting models. In an analysis of both developed and emerging markets, Pereiro and González-Rozada (2015) reported that regime-shifting models outperformed single-regime models. Using a different methodological toolbox, Nystrup, Hansen, Madsen, and Lindström (2015) also suggested that a simple regime-shifting approach will outperform a static allocation approach. Mulvey and Liu (2016) used a machine learning algorithm, trend filtering, to categorize regimes. They found these approaches most useful for long-term planning, and they suggested that such methods might help reduce the downside risks for university endowments and foundations. Xiong, Idzorek, and Ibbotson (2016) demonstrated the value of forecasting left-tail risk—though not the same as a regime—because it provided better downside protection while maintaining Sharpe ratios.

Ang, Madhavan, and Sobczyk (2017) developed a methodology to separate the effects of static factor exposures from dynamic timing from security selection. They reported that in a sample of mutual funds, each component tends to be distinct. But perhaps more important than this specific empirical result itself is that mainstream investment practitioners are recognizing

the importance of being able to attribute performance to dynamic factor-timing skills. In part, this trend is undoubtedly a reflection of the explosion in smart-beta products and ETFs that tend to deliver static factor exposures in a very low-cost approach. Perhaps this is why Jacobs (2015) questioned whether smart beta is really state of the art. Indeed, the author argued that dynamic, multifactor approaches can lead to better outcomes than those of static, smart-beta approaches.

The surge in interest on dynamic factor timing may very well alter what is meant by the term "factor." For decades, factors have been associated with exposures that are priced on average over time. But as conditional models become better understood, this definition may evolve. Whereas forecasting factors that are priced on average will always remain of keen interest, factors that are not priced on average (at least in a statistical sense) may garner interest. For example, oil prices may not be a priced factor in the standard long-term sense but may very well speak loudly in certain economic environments and conditions. One can imagine that the search for dynamic factors that condition expected returns will continue and perhaps expand. Separating the wheat from the chaff will be a continuing challenge for quantitative dynamic factor investing. But dynamic factor-timing models appear to be here to stay and are likely to grow in importance.

Conclusion

Quantitative equity management is alive and well—and intellectually active—as investors seek to better manage risk and return. Factor investing has taken off commercially in the form of smart-beta products and strategies, vetted by decades of prior and current research. Dynamic factor-timing approaches are probably still in the early stages, especially from a commercial perspective. However, one might reasonably forecast this to be a growth area for the quantitative equity field. A new generation of big data approaches is developing in the field and will likely grow as technology becomes more capable and more data are digitally available. Quantitative equity management techniques are helping investors achieve more-efficient and appropriate investment outcomes.

If you'd like to contact the authors, please send an email to marc@reinganum.com.