Factor Modeling: The Benefits of Disentangling Cross-Sectionally for Explaining Stock Returns

Bruce I. Jacobs and Kenneth N. Levy

Bruce I. Jacobs
is principal and co-chief investment officer of Jacobs Levy Equity Management in Florham Park, NJ.
bruce.jacobs@jlem.com

Kenneth N. Levy
is principal and co-chief investment officer of Jacobs Levy Equity Management in Florham Park, NJ.
ken.levy@jlem.com

KEY FINDINGS

- More than three decades ago, the authors pioneered a cross-sectional approach to factor modeling that disentangled the unique contributions of numerous factors to the pricing of individual stocks.
- A time-series approach using portfolio sorts has dominated the asset pricing literature, but cross-sectional analysis using firm characteristics has greater explanatory power for stock returns and helps practitioners address one of the most fundamental issues in investment management: understanding and predicting the returns of individual stocks.
- The authors revisit disentangling returns using a cross-sectional model, compare factor models using cross-sectional factors with those using time-series factors, and discuss the benefits and challenges of cross-sectional models.

ABSTRACT

More than three decades ago, Jacobs and Levy introduced the idea of disentangling stock returns across numerous factors. They identified the relationships between individual stock returns and firm characteristics using a cross-sectional analysis and examined the benefits of using the resulting time series of returns to the disentangled factors for return forecasting. Some years later, an alternative factor model proposed by Fama and French made use of time-series factors based on portfolio sorts (examples of these time-series factors include the return differences between small- and big-capitalization stocks and between high- and low-book-to-price stocks). Recently, Fama and French found that the cross-sectional approach using firm characteristics is better able to explain stock returns than the time-series approach based on portfolio sorts. This article compares models that use cross-sectional factors across firm characteristics with models that use time-series factors based on portfolio sorts and discusses the benefits and challenges of the cross-sectional approach for investment management.

TOPICS

Factor-based models, statistical methods, security analysis and valuation, equity portfolio management*

Factor models have dominated the asset pricing literature for decades now. Jacobs and Levy (1988, 1990) were among the early innovators in the study of multifactor asset pricing models. Building upon work by Fama and MacBeth (1973), they adopted a cross-sectional approach to factor modeling that disentangled the unique contributions of numerous factors to the pricing of individual stocks. Fama and French (1993) proposed an alternative approach, time-series factor modeling, for the pricing of portfolios rather than individual stocks. It became the dominant factor methodology in empirical asset pricing. However, recent work by Fama and French (2020a) suggests that cross-sectional approaches, similar to those proposed by Jacobs and Levy (1988), may be more effective at explaining returns. We agree. Cross-sectional analysis helps practitioners address one of the most fundamental issues in investment management: understanding and predicting the returns of individual stocks.

The literature has for some decades favored the time-series approach. Many practitioners are not aware of the distinct advantages of the cross-sectional approach, including the ability to use numerous factors, some of which may be conditioned on the market environment, to capture the complexity of security pricing. Even practitioners familiar with the basic principles of the cross-sectional approach often struggle with the appropriate contexts for its application—including when its use, rather than Fama and French's time-series approach, is advisable.

This article will help practitioners navigate those difficult decisions by providing a careful analysis of the cross-sectional approach, including its appropriate applications. We will revisit disentangling returns using a cross-sectional model and discuss the usefulness of the resulting time series of returns to the disentangled factors for return forecasting. We also will compare factor models using cross-sectional factors with those using time-series factors and discuss the benefits and challenges of cross-sectional models.

FROM FACTORS TO FACTOR MODELS

We begin by defining the term *factor*, which has several meanings. Factor may refer to an attribute proxying for a common source of risk, such as the size factor or the book-to-market (BM) factor in the three-factor model of Fama and French (1993); in that sense, it is also used to describe the risk premium for each factor (i.e., factor premium or factor return). For example, the Fama–French size factor SMB (small minus big) is technically the risk premium for the size-related common factor. An asset's sensitivity to each risk factor is called a factor *loading*. Factor can also refer to a firm characteristic that has a relationship with stock return, regardless of whether the relationship results from risk or from mispricing (e.g., Jacobs and Levy 1988; Harvey, Liu, and Zhu 2016; Green, Hand, and Zhang 2017). Here, a factor represents an attribute, such as earnings accruals or idiosyncratic volatility, or the return associated with the attribute.

There is still considerable debate over whether factor premiums reflect risk or mispricing. Are the premiums due to the risk loadings on the common factors associated with firm characteristics or to the mispricing of the firm characteristics? Consider the value premium. One strand of literature (e.g., Fama and French 1993; Davis, Fama, and French 2000) suggests that, in a rational risk-based framework, characteristics should proxy for sensitivity to common risk factors. That is, value stocks, defined as stocks with high BM, tend to earn relatively higher returns because these firms have high loadings on (or covariance with) the common distress factor, which has a positive risk premium. In contrast, evidence in favor of mispricing has been provided by another strand of literature (e.g., Daniel and Titman 1997; Chordia, Goyal, and Shanken 2017), which found that high-BM stocks have higher expected returns regardless of their actual distress factor loadings; this implies that high-BM stocks are undervalued—that is, mispriced. However, as Cornell (2020) noted, this debate is less critical for practitioners than for academics. The practical question is whether certain characteristics explain the return differences across stocks.

Factor models can help answer that question. Factor models can be built in a myriad of ways. Some emphasize parsimony with a few factors; others are broad with numerous factors. Some use portfolio sorts based on firm characteristics; others use many firm characteristics of individual stocks. Some periodically update for evolving firm characteristics; others update routinely. Some seek to explain the behavior of portfolios, others the behavior of individual stocks. Some have the objective of testing the validity of the capital asset pricing model (CAPM), others of measuring portfolio risk, and still others of explaining or forecasting returns. Some project historical returns forward; others condition return forecasts on the market environment. Various factor models have differing abilities to explain portfolio or individual stock returns and can be used in different ways in portfolio management.

Factor models are generally estimated and tested with either time-series or cross-sectional regression. A time-series regression uses data from a single entity that has been observed at fixed intervals over time. In time-series applications, the entity is typically a portfolio. For example, the factor loadings of a portfolio for the Fama and French (2015) five-factor model can be estimated by running a time-series regression of the portfolio's monthly excess returns (over the risk-free rate) on the five monthly time-series factors: the market excess return; the size factor (SMB), measured as the return spread between small and big stocks; the value factor (HML), measured as the return spread between stocks with high-BM ratios and those with low-BM ratios; the profitability factor (RMW), measured as the return spread between stocks with robust operating profitability and those with weak profitability; and the investment factor (CMA), measured as the return spread between stocks with conservative investment and those with aggressive investment (as measured by asset growth rates). Each of these time-series factors is constructed by using a portfolio sort based essentially on a single corresponding characteristic.[1] The time-series approach has, for more than two decades, been the primary approach used by academics for asset pricing and measuring abnormal returns in event studies and by practitioners for selecting factor portfolios and evaluating portfolio performance.

A cross-sectional regression uses data observations that come from different entities as of a point in time. In cross-sectional applications, the entities can be either portfolios or individual stocks. For example, one can run a cross-sectional regression of stock returns for a given month on several firm characteristics observed at the end of the prior month to estimate the relationship of return with each characteristic. Cross-sectional factors are estimated from cross-sectional analysis in which asset returns for each time period are regressed on multiple firm characteristics. The cross-sectional approach has not been widely used by academics but has been used by some practitioners to address one of the most fundamental issues in investment management: explaining and predicting returns for individual stocks.

[1] Fama and French control for firm size when constructing factors via a two-way sort (by market capitalization and by the characteristic of interest); this controls for only one characteristic. Put precisely, SMB is constructed based on independent double sorts of stocks on size and BM ratio, and HML, RMW, and CMA are constructed based on independent double sorts on size and the corresponding characteristic (BM, operating profitability, or investment). That is, SMB is controlled for BM. HML, RMW, and CMA are controlled for size. Specifically, Fama and French (2020a) first produced six size/BM portfolios from the intersection of the 2×3 size/BM independent sorts. They constructed SMB as the average return of the three small portfolios (with high, medium, and low BM) minus the average return of the three big portfolios. Similarly, they obtained HML as the average of the difference between the returns on the high- and low-BM portfolios of big stocks and the return difference for high- and low-BM portfolios of small stocks. RMW and CMA are constructed in the same way as HML except the second sort is on either operating profitability or investment. Note that Fama and French (2015) also constructed the alternative version of SMB, HML, RMW, and CMA based on the intersection of four independent sorts into two groups on size, BM, operating profitability, and investment ($2 \times 2 \times 2 \times 2$). It has become standard practice to construct the four factors from the 2×3 sorts on size and one other variable. However, portfolio sorts do not optimally isolate each factor from the others.

The adequacy of a factor model can be tested based on how well it explains assets' average returns. For example, Fama and French (1993) constructed 25 test portfolios from the intersections of the size and value quintiles to determine how well their three-factor time-series model explained the average returns of the portfolios. Since then, it has become standard practice to use the 25 Fama–French size and value portfolios as initial test assets when evaluating asset pricing models.

Although many factor models do a good job of fitting the average returns of the 25 Fama–French size and value portfolios, their ability to explain the average returns of other portfolios is limited. There are two plausible and non-competing explanations for this. First, Fama–French time-series factor models favor parsimony and typically contain a limited number of factors, potentially ignoring other factors that may further contribute to explaining stock returns. Second, the time-series factor models do not take into account the interrelationships among factors. Yet the return spreads SMB, HML, RMW, and CMA are correlated because size, BM ratio, profitability, and investment are related characteristics; Fama and French (2015), for instance, found that large growth stocks with low BM ratios tend to be more profitable and invest more.

Prior to the advent of the earliest Fama–French model, Jacobs and Levy (1988) developed a cross-sectional model that uses numerous factors to explain stock returns, taking into account their interrelationships. Jacobs and Levy (1988) used cross-sectional regressions at the individual stock level to disentangle multiple firm characteristics, or factors, to estimate the pure returns to each factor. Disentangling can reveal which factors really matter; it provides the pure return to each factor, uncontaminated by the effects of other factors. By contrast, when a single firm characteristic is used in a portfolio sort or a simple regression (i.e., a regression with a single independent variable), there is no disentangling across related characteristics; returns estimated from simple regressions are naïve factor returns. Jacobs and Levy (1988, 1989b) pioneered these insights and introduced the terms "disentangling" as well as "pure" and "naïve" returns.

Fama and French have recently (2020a) found that models using cross-sectional factors do a better job of explaining the average returns of various anomaly portfolios than do models using time-series factors. This finding is supportive of Jacobs and Levy's (1988) much earlier insight of disentangling returns cross-sectionally across factors.

THE DEVELOPMENT OF CROSS-SECTIONAL MODELING

Jacobs and Levy (1988) used cross-sectional regressions to measure anomalies—as factors were called then, being anomalous to the efficient market hypothesis (Fama 1970) and the single-beta-factor CAPM (Sharpe 1964). The paper analyzed the returns to multiple firm characteristics simultaneously to determine which anomalies mattered and which were mere proxies for others. It identified which had statistically significant average returns by examining the time series of cross-sectional factor returns in the same way as the recent Fama and French (2020a) model using cross-sectional factors. It also examined autocorrelation and seasonality of the time series of the disentangled returns to firm characteristics. The paper was followed by another that proposed a forecasting model for the disentangled returns to the size factor conditioned on the market environment (Jacobs and Levy 1989b).[2]

Prior to Jacobs and Levy (1988), Fama and MacBeth (1973) regressed portfolio returns on three potential descriptors of risk: market beta, the square of market beta,

[2] Editor's note: Professor Charles A. D'Ambrosio, former editor of the *Financial Analysts Journal*, which published both papers during his tenure as editor, said in the *Wall Street Journal* that "Jacobs and Levy were the first to bring so much of this anomaly material together" (White 1991).

and idiosyncratic volatility. Fama and MacBeth (1973) focused on testing the CAPM. To minimize statistical concerns associated with formal tests of the CAPM, the article used portfolios as test assets. Fama and MacBeth (1973) implemented the two-step regression method: First, portfolio betas were estimated using time-series regression; then the risk premium for the beta estimate was estimated using cross-sectional regression. The beta estimates obtained from the first-step time-series regression are likely to be more precise for portfolios than for individual stocks (Blume 1970).

Because all the determinants of each portfolio's return are not knowable, the error terms of the regression model (which are not explained by the factors used) may be correlated across portfolios because of the effect(s) of one or more omitted factors. In that case, the ordinary least squares (OLS) factor return estimates are unbiased, but the standard errors of those estimates are no longer valid for quantifying their precision. Fortunately, as Fama and MacBeth (1973) noted, when each factor return estimate is stacked across time, the cross-sectional correlation issue may be circumvented by calculating the standard error based on the time-series variation of the estimates.[3]

Jacobs and Levy (1988) had a different focus than Fama and MacBeth (1973)—identifying anomalies that mattered and their corresponding firm attributes that were robustly associated with the cross section of individual stock returns. As such, Jacobs and Levy employed individual stocks instead of portfolios. The cross-sectional regression methodology used by Jacobs and Levy (1988) differed from that of Fama–MacBeth in other ways as well. Beyond descriptors of risk, it included numerous fundamental factors based on accounting data and information from security analysts, as well as momentum and reversal factors. The cross-sectional regression incorporated 38 industry variables, as well as the 25 firm characteristics, so that industry affiliation did not contaminate the return effects associated with the firm characteristics. For example, a change in oil prices would affect oil stock prices, and to the extent that oil companies are large, a regression that did not control for industry affiliation would partially attribute such return effects to the size factor.

Jacobs and Levy (1988) also standardized each firm characteristic for consistent scaling over time to have a cross-sectional standard deviation of one and truncated outliers. The regression coefficient for each characteristic is then interpreted as the payoff to a long–short portfolio that has unit value of that characteristic and zero value of all other characteristics. Furthermore, weighted least squares (WLS), a special case of generalized least squares (GLS), was used rather than OLS to make valid inferences in the presence of heteroskedasticity by placing less weight in the regression on stocks with higher volatility.[4] Jacobs and Levy (1988) measured volatility via the variance of the residuals from market model time-series regressions of the individual stocks. The article noted the relationship between the variance of the residuals and market capitalization: "Because higher residual risk is correlated with small size, GLS weights generally lie between capitalization and equal weights."

[3] Petersen (2009) demonstrated that the Fama–MacBeth procedure produces unbiased standard errors in simulated data in which residuals are cross-sectionally correlated. He further reported that, in cross-sectional dependence, the Fama–MacBeth standard errors are better than the clustered standard errors, another popular alternative for cross-correlation adjustment. Petersen (2009) also identified situations in which the Fama–MacBeth standard errors are likely to be downward biased. When residuals are correlated over time, he found a substantial downward bias in the Fama–MacBeth standard errors. This is good news from an asset pricing perspective because it means that the Fama–MacBeth standard errors are unlikely to exhibit significant biases, because returns for any given firm are likely uncorrelated across time. As Cochrane (2005) noted, "[t]he assumption that the errors are not correlated over time is probably not so bad for asset pricing applications, since returns are close to independent."

[4] Although a GLS procedure with a full covariance matrix can be applied to models in which cross-sectional correlation of the error terms is present, the covariance matrix of the error terms is not known.

As Cochrane (2005) put it, "[a] GLS regression can be understood as a transformation of the space of returns to focus attention on the statistically most informative portfolios…. The statistically most informative portfolios are those with the lowest residual variance." Although Cochrane's remarks refer to portfolios, they also apply to individual stocks, as explored by Jacobs and Levy (1988).

In practice, it is quite common to use OLS estimates with Fama–MacBeth standard errors in the anomaly literature. For example, Harvey, Liu, and Zhu (2016) relied on the reported *t*-statistics of historical factors from 313 papers, most of which used OLS and the Fama–MacBeth procedure. However, it would be more appropriate to use WLS, rather than OLS, for anomaly research using individual stocks for at least two reasons. First, given that research using individual stocks is likely to have a wider range of error variance than that using portfolios, WLS, with weights inversely proportional to the residual variance (as done by Jacobs and Levy 1988), would be more desirable than OLS. Second, although microcap stocks are rarely incorporated into institutional portfolios, their impact on anomaly returns can be influential (Fama and French 2008; Hou, Xue, and Zhang 2015). When the WLS weight is set to the market capitalization of each stock, the WLS approach helps mitigate the influence of microcap stocks (Green, Hand, and Zhang 2017).

Several earlier studies also considered multiple factors in different contexts. King (1966) recognized the possibility of multiple factors in asset pricing by documenting that industry factors can capture cross-sectional differences in stock return. Rosenberg and Marathe (1976) applied cross-sectional factor models to risk analysis and portfolio optimization. Their work attributed the residual risk of individual stocks to six risk indexes and 39 industry classifications and led to the Barra multifactor risk models. Rosenberg, Reid, and Lanstein (1985) tested the performance of two strategies—one based on the BM ratio and the other based on specific return reversal—constructed to be orthogonal to one another, to 11 risk indexes, and to 55 industry classifications. Sharpe (1982) examined five factors and eight broad industry classifications to test the implication of the single-factor CAPM and provided presumptive evidence that beta was not the only priced factor. Jacobs and Levy (1988) were the first to simultaneously examine a comprehensive set of virtually all the then-known anomalies (25 in number), as well as 38 industry classifications, and to provide a multifactor framework for security return forecasting.[5]

Cross-sectional regressions measure all anomaly effects jointly, purifying each effect by controlling for all the other effects. For Jacobs and Levy (1988), each pure return represents the return to a portfolio that has one standard deviation of exposure to one factor but no exposures to the other factors. The concept of a pure return has since become widely accepted, and the term has entered industry parlance.[6,7]

As mentioned earlier, Jacobs and Levy (1988) described returns that derive from simple regressions as naïve. Simple regressions, much like portfolio sorts, naively

---

[5] In the Rosenberg and Marathe (1976) paper, the six risk indexes, each composed of various descriptors that characterize the firm, were market variability, earning variability, unsuccess and low valuation, immaturity and smallness, growth orientation, and financial risk. In the paper by Rosenberg, Reid, and Lanstein (1985), the 11 risk indexes were variability in markets, success, size, trading activity, growth, earnings/price, earnings variation, financial leverage, foreign income, labor intensity, and yield. For Sharpe (1982), the five factors were beta, yield, size, bond beta, and alpha. For Jacobs and Levy (1988), see the Appendix for the list of the anomalies examined.

[6] Earlier, Rosenberg, Reid, and Lanstein (1985) introduced the phrase "pure hedge portfolio." The article used the adjective "pure" to describe a self-financed ("zero investment") hedge portfolio in which the shorts provide the financing for the longs. In contrast, Jacobs and Levy (1988) used the adjective "pure" to describe the return associated with a firm characteristic disentangled from the returns associated with other firm characteristics.

[7] Clarke, de Silva, and Thorley (2017) spoke of "pure factor portfolios," each of which has one unit (or one standard deviation) of exposure to the factor of interest but no exposure to the other factors.

measure only one anomaly at a time, with no control for other related effects. The returns to a single anomaly will often proxy for several related return effects. For example, a naïve return to low price-to-earnings ratio (P/E) may proxy for returns to small capitalization, high dividend yield, certain industries such as utilities, and so on. In contrast, a multivariate framework properly attributes return to its underlying sources. A pure return to low P/E can be thought of as the return to a portfolio that has a low P/E but is market-like in other respects. Such a portfolio would be immunized against other effects that could contaminate the measurement of the pure return to low P/E.

Jacobs and Levy (1988) found that the average returns to many of the anomalies analyzed were significant at standard levels of significance. Even at the higher levels of significance recommended today to control for data snooping (Harvey, Liu, and Zhu 2016), many of the anomalies would still have been deemed significant. The evidence suggested that returns were predictable and the market not totally efficient. The finding that beta was insignificant was not supportive of the one-factor CAPM. Although the factor returns analyzed in 1988 have evolved over time and some have become less significant (Green, Hand, and Zhang 2017), the novel approach of disentangling numerous anomalies or firm characteristics simultaneously proved its usefulness over the years and is even more relevant in the current investment environment with its "zoo of new factors" (Cochrane 2011).[8]

Several studies have followed the approach of Jacobs and Levy (1988). Fama and French (1992) used monthly cross-sectional regressions of stock returns on beta, size, value (BM), and other variables and found that size and value helped explain average stock returns, but beta did not. In the spirit of Jacobs and Levy (1988), Haugen and Baker (1996) ran monthly cross-sectional regressions of returns on more than 50 characteristics to estimate the returns associated with each characteristic.[9] Ziemba and Schwartz (1992) applied the Jacobs and Levy methodology in the Japanese market.[10] Fama and French (2008) used cross-sectional regression to examine the relationships between average returns and size, value, profitability, growth, accruals, net stock issues, and momentum. That article estimated cross-sectional regressions separately for different size groups and found that, controlling for other anomalies,

---

Their methodology, terminology, and insights are virtually identical to those of Jacobs and Levy (1988). Wall Street research reports commonly adopt the term *pure factors* in various contexts, including factor construction (Nomura 2012; Société Générale 2017; J.P. Morgan 2019), factor risk modeling (MSCI 2011), factor exchange-traded funds (Axioma 2012), and factor volatility (Wolfe Research 2020). They often acknowledge the disentangling benefit of pure factors. For example, the Wolfe Research report states that "Looking at the pure factors allows us to disentangle the impact of other factors from the target factor." Bloomberg's Factors to Watch (FTW <GO>) function offers a "Pure Factor Returns" mode that allows users to obtain pure returns to various style factors.

[8] Editor's note: Nobel laureate Harry M. Markowitz later referred to Jacobs and Levy (1988) as "seminal" in his foreword to their book *Equity Management* (Markowitz 2000, 2017). He also acknowledged the methodology, terminology, and insights of the article in other work (Markowitz and van Dijk 2006, p. 188):

> Before Jacobs and Levy (1988) anomaly studies considered small numbers of variables, usually one to three at a time. Observing that some apparent anomalies may be surrogates for others, Jacobs and Levy fit a series of monthly cross-sectional regressions of security excess returns against 25 anomaly and 38 industry variables. This allowed them to "disentangle" what they called the "pure" (i.e., underlying) anomalous effects from what they called the "naïve" effects observed from simple regressions against anomalous variables one at a time. The Jacobs and Levy methodology may be used for expected return estimation as well as for explaining observed anomalies.

[9] Professor William Ziemba noted in Ziemba and Ziemba (2007) that Haugen's "real crime" was not citing Jacobs and Levy (1988), which "is the original reference for the U.S.," and his own subsequent work on anomalies in Japan (Ziemba and Schwartz 1992). Ziemba said his work was done subsequent to seeing Jacobs and Levy (1988), after which he "explained these things to him [Haugen] repeatedly." Ziemba continued, "This is typical sadly of modern academic finance. If it's not published in one of about five or perhaps 10 top journals, it does not exist!"

[10] See also Schwartz and Ziemba (2000).

returns associated with net stock issues, accruals, and momentum were strong for all size groups (microcap, small, and big). Their finding suggested a market with greater dimensionality.

USING TIME-SERIES FACTORS OR CROSS-SECTIONAL FACTORS

It is useful to examine in more detail how factor models that use time-series or cross-sectional factors work and the differences between them. For expositional purposes, although any multifactor model can be used, we follow Fama and French (2020a), taking the Fama and French (2015) five-factor model as an example.

Use of Time-Series Factors

The original Fama–French five-factor model, which uses time-series factors, is as follows:

$$R_{it} - R_{ft} = a_i + b_i(R_{mt} - R_{ft}) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + e_{it} \tag{1}$$

where R_{it} is the return on security or portfolio i for period t, R_{ft} is the risk-free return, R_{mt} is the return on the market portfolio, and e_{it} is an error term. As introduced earlier, SMB_t, HML_t, RMW_t, and CMA_t are differences between returns on diversified portfolios of small and big stocks, stocks with high- and low-BM ratios, stocks with robust and weak operating profitability, and stocks with conservative and aggressive investments, respectively.

To test the Fama–French five-factor model, a time-series regression (Equation 1) of excess return is run on the five time-series factors to estimate the intercept (a_i) and the factor loadings (b_i, s_i, h_i, r_i, and c_i) for each asset i. The resulting OLS factor loading estimates are not time varying (i.e., they are fixed over the estimation period). The pricing error for asset i is given by the estimate of the intercept ($\hat{a}_i$); a perfect asset pricing model would result in an intercept of zero.

Use of Cross-Sectional Factors

The cross-sectional counterpart of the Fama–French five-factor model starts from the following cross-sectional regression:

$$R_{it} = R_{zt} + R_{MCt} MC_{it-1} + R_{BMt} BM_{it-1} + R_{OPt} OP_{it-1} + R_{INVt} INV_{it-1} + e_{it} \tag{2}$$

where R_{it} is the return on security or portfolio i for period t, R_{zt} is the common return to all assets not captured by the regression explanatory variables, MC_{it-1} is the size (market capitalization) observed in period $t-1$, BM_{it-1} is the value (BM ratio) observed in $t-1$, OP_{it-1} is the operating profitability observed in $t-1$, INV_{it-1} is the change in investments (asset growth rate) observed in $t-1$, and e_{it} is an error term.

Each explanatory variable in Equation 2 is a characteristic of asset i that motivates four of the five factors (the exception being the market factor $R_{mt} - R_{ft}$) in the Fama–French five-factor model (Equation 1). Fama and French (2020a) did not include beta with respect to the market factor as a characteristic; hence the cross-sectional counterpart has four, rather than the original five, factors.[11] To estimate these four factors, a cross-sectional regression (Equation 2) of asset returns is run on the four

[11] Note that, for Jacobs and Levy (1988), beta is one of the explanatory variables in the cross-sectional regressions of returns on firm characteristics.

characteristics (MC_{it-1}, BM_{it-1}, OP_{it-1}, and INV_{it-1}) for each period t. The resulting OLS estimates are the time series of the intercept ($\hat{R}_{zt}$) and the four cross-sectional factors ($\hat{R}_{MCt}$, $\hat{R}_{BMt}$, $\hat{R}_{OPt}$, and $\hat{R}_{INVt}$), which will be plugged into the following four-factor model:

$$R_{it} - R_{zt} = MC_{it-1}R_{MCt} + BM_{it-1}R_{BMt} + OP_{it-1}R_{OPt} + INV_{it-1}R_{INVt} + e_{it} \quad (3)$$

Note that Equation 3 is not a regression model. It simply interchanges characteristics and factors in the regression model (Equation 2) to mimic the traditional factor model (Equation 1).[12] Because the observed values of the explanatory variables (characteristics) and the coefficient estimates (factor estimates) from the regression model (Equation 2) are plugged into Equation 3, the pricing error for asset i in Equation 3 is the average across t of the residuals $\hat{e}_{it}$ in the regression model (Equation 2).

Fama and French (2020a, p. 1892) claimed that "Our insight is that when the cross-section regression … is stacked across t, it becomes an asset pricing model that can be used in time-series applications." This was the insight of Jacobs and Levy (1988), who examined the time series of factor returns derived from cross-sectional regressions. Furthermore, Jacobs and Levy (1989b) used the time series of size factor pure returns to produce forecasts conditioned on market and economic information.

In the cross-sectional regression of Jacobs and Levy (1988), the characteristics of each asset were prespecified and allowed to evolve over time. Equivalently, in the cross-sectional model (Equation 3), factor loadings (characteristics MC_{it-1}, BM_{it-1}, OP_{it-1}, and INV_{it-1}) are prespecified (i.e., not estimated) and time varying, unlike the factor loadings (b_i, s_i, h_i, r_i, and c_i) in the time-series model (Equation 1). In sum, Equation 1 estimates constant loadings on the factors in the time-series regression separately for each asset i. In contrast, in Equation 3, the factors are estimated from the cross-section regression (Equation 2), and the factor loadings are the time-varying size, value, profitability, and investment characteristics of each asset.

Comparing the Use of Time-Series Factors with Cross-Sectional Factors

An apparent advantage of time-series factors over cross-sectional factors is that in practice they are easier to replicate. For example, HML is essentially the return of a long–short portfolio based on the BM ratio. Note, however, that each cross-sectional factor is also the return of a long–short portfolio based on the standardized exposure to the factor of interest, with the added construction that the portfolio is neutralized to other factors. Another difference is that time-series factors such as HML can be measured over short periods, whereas many firm characteristics are typically only observable at certain intervals. Nonetheless, the factor returns to firm characteristics can be estimated at any frequency, given the most recent observed values of those firm characteristics.

Fama and French (2020a) compared models that use cross-sectional factors and models that use time-series factors in their ability to describe average equity returns. They examined various performance metrics based on pricing errors and found that, on every performance metric, models based on cross-sectional factors dominate the traditional models based on time-series factors used in their previous articles (e.g., Fama and French 1993, 2015).[13]

---

[12] Equation 3 appears redundant because it is mathematically equivalent to Equation 2. However, Equation 3 follows the traditional factor model structure of Equation 1, in which factor returns are independent variables for explaining asset returns and factor loadings are regression coefficients. In contrast, Equation 2 has characteristics as independent variables and factor returns as regression coefficients.

[13] The performance metrics used by Fama and French (2020a) include the average of the absolute values of the pricing errors, the average of the absolute values of the t-statistics for the pricing errors, and the average of the squared pricing errors (both raw and adjusted for the standard error of each

There are two potential explanations for this dominance of models that use cross-sectional factors over those that use time-series factors: (1) time-varying factor loadings and (2) factors constructed from cross-sectional regressions.[14] Fama and French (2020a) presented a simple yet intuitive way to test the impact of time-varying factor loadings on the pricing errors; they examined the pricing errors for the variant of Equation 3 that replaced the time-varying factor loadings with the (fixed) time-series averages of the factor loadings as follows:

$$R_{it} - R_{zt} = \overline{MC}_i R_{MCt} + \overline{BM}_i R_{BMt} + \overline{OP}_i R_{OPt} + \overline{INV}_i R_{INVt} + e_{it} \tag{4}$$

where $\overline{MC}_i$ is the time-series average of MC_{it-1}, $\overline{BM}_i$ is the time-series average of BM_{it-1}, $\overline{OP}_i$ is the time-series average of OP_{it-1}, and $\overline{INV}_i$ is the time-series average of INV_{it-1}. They showed that, although the pricing errors of Equation 4 are slightly higher than those of Equation 3, Equation 4 still performs better than Equation 1, which uses time-series factors. This finding suggests that, in terms of explaining equity returns, the outperformance of models based on cross-sectional factors over those based on time-series factors owes more to the cross-sectional factors than to the time-varying factor loadings.

Factors constructed from cross-sectional regressions have at least two advantages over time-series factors. First, the time-series factors *SMB*, *HML*, *RMW*, and *CMA* in Equation 1 are constructed from sorts of stocks into groups based on size, value, profitability, and investment, respectively (see footnote 1 for the definition of the factor sorts). As acknowledged by Fama and French (1993), the number of groups and the breakpoints chosen are arbitrary and ad hoc.[15] Because Fama–French regressions use ad hoc factors as explanatory variables, the resulting time-series estimates of factor loadings may not be optimal. In contrast, cross-sectional factors resulting from the regressions of returns on prespecified characteristics are optimal.

Second, each time-series factor sort unavoidably captures return effects from the other factor sorts, while the cross-sectional regression (Equation 2) disentangles the size, value, profitability, and investment effects. For example, Fama and French (1992) reported a negative correlation between the size and value characteristics; value stocks tend to be lower priced and thus smaller in size. To address that, Fama and French (1993) constructed SMB and HML based on double sorts on size and value. Unlike the regression method of least squares, however, portfolio sorts do not optimally disentangle the return effects from each factor. Jacobs and Levy (1988, 1989b) provided pure returns to each factor, avoiding the confounding influence of the other factors by disentangling in a 25-factor cross-sectional model.

Using the terminology of Jacobs and Levy (1988), the traditional Fama–French time-series factors are naïve returns from portfolio sorts, and their cross-sectional factors are pure returns. The dominance of models that use cross-sectional factors over those that use time-series factors in explaining equity returns is consistent with the cross-sectional approach used by Jacobs and Levy (1988). The findings there

pricing error) divided by the cross-sectional variance of average returns. For time-series factor models only, they also used the GRS statistic of Gibbons, Ross, and Shanken (1989) and the maximum squared Sharpe ratio for the performance metrics.

[14] Several studies have proposed factor models that allow the factor loadings of an individual asset to vary over time (see, e.g., Rosenberg 1974; Harvey 1989; Jagannathan and Wang 1996).

[15] Fama and French (1993, p. 9) stated: "Our decision to sort firms into three groups on *BE/ME* and only two on *ME* follows the evidence … that book-to-market equity has a stronger role in average stock returns than size. The splits are arbitrary, however, and we have not searched over alternatives. The hope is that the tests here … are not sensitive to these choices. We see no reason to argue that they are." They also used the 30th and 70th percentiles among NYSE stocks as breakpoints for low and high BM; the choice of these breakpoints is also arbitrary.

revealed a much greater dimensionality to the stock market than suggested by the one-factor CAPM or by previous studies that looked at only one or a few anomalies or factors. A model with greater dimensionality may be better able to explain the cross section of stock returns. Furthermore, pure returns to anomalies from cross-sectional regression can more fully explain stock returns than can the naïve returns from analyzing each anomaly individually, whether using simple regression or portfolio sorts. Jacobs and Levy (1989a) found the market to be permeated by a complex web of price behaviors, reflecting the interaction of numerous fundamental and behavioral factors as well as institutional features such as the regulatory environment.

Challenges for Cross-Sectional Modeling

Data mining, multicollinearity, out-of-sample predictability, and machine learning application present some challenges for statistical analysis. Cross-sectional approaches permit more factors, which can further heighten these challenges.

Data mining. In the last few decades, the anomaly literature has examined hundreds of factors. The ensuing concerns about data mining have called for a more rigorous examination of alleged factors. For example, McLean and Pontiff (2016) studied the return predictability of 97 anomalies and found that the average return was 26% lower out of sample and 58% lower once findings were published. Harvey, Liu, and Zhu (2016) studied more than 300 anomaly papers and proposed a higher hurdle for the *t*-statistic on new factors—a value that exceeds 3.0—to account for multiple testing. Green, Hand, and Zhang (2017) tested 94 firm characteristics and found that only 12 characteristics were independent determinants of average stock returns. Taken together, the evidence suggests that data mining presents a challenge to investment practitioners searching for reliable return-predictive signals.

Multicollinearity. Adding more factors can lead to multicollinearity. It is, however, possible to detect multicollinearity using diagnostics such as the variance inflation factor (VIF), tolerance, and condition index.[16] Jacobs and Levy (1988) addressed multicollinearity with a simple diagnostic test, comparing the time-series standard deviation of naïve returns and pure returns. That article showed that the time-series standard deviation of all 25 anomalies was smaller with pure returns and concluded that multicollinearity was not a serious problem.[17]

Out-of-sample predictability. A cross-sectional model with a large number of factors can potentially overfit the data at the expense of out-of-sample predictability. It turns out that many widely available statistical procedures can be used to mitigate this problem. For example, one can use lasso or ridge regressions to penalize those firm characteristics that make only a minor contribution to the expected return (e.g., Chinco, Clark-Joseph and Ye 2019; Feng, Giglio, and Xiu 2020; Gu, Kelly, and Xiu 2020). An alternative approach is to use methods such as principal components analysis (PCA) or partial least squares (PLS) to transform a large set of the original variables into a lower-dimensional set of features (e.g., Light, Maslov, and Rytchkov 2017; Gu, Kelly, and Xiu 2020). A drawback of PCA, however, is that it does not explain

[16] The VIF for independent variable *j* is obtained from the following formula: $VIF_j = 1/(1 - R_j^2)$, where R_j^2 is the R^2 from regressing independent variable *j* on all other independent variables. Tolerance is the reciprocal of VIF. The condition index is computed as the square root of the ratio of the largest eigenvalue to each individual eigenvalue of $X'X$, where X is a matrix of values of explanatory variables.

[17] If severe multicollinearity is detected among variables (e.g., among various value metrics such as book-to-market ratio, cash-flow-to-price ratio, and earnings-to-price ratio), several approaches can address the issue. The most common approaches fall into three categories: removing some of the variables, combining those variables to create a composite measure, and using more advanced dimensionality reduction techniques such as PCA or PLS, which will be discussed in more detail shortly.

how each independent variable (i.e., firm characteristic) is related to the dependent variable (i.e., expected stock return).

Machine learning application. Recent years have witnessed a surge of interest in machine learning methods as a potential replacement for traditional regression methods in both academia and industry (see Bartram, Branke, and Motahari 2020 for an overview of machine learning applications in asset management). Gu, Kelly, and Xiu (2020) documented that machine learning methods can be beneficial in return prediction, particularly when the relationships are nonlinear among predictors and between predictors and expected return.[18]

Recent research, however, suggests that it is important to use caution when applying machine learning to investment management practice. Israel, Kelly, and Moskowitz (2020) argued that, although machine learning thrives in a big data and high-signal-to-noise-ratio environment, stock return prediction has relatively small data and low-signal-to-noise ratios, a less-than-ideal environment for machine learning. Bartram, Branke, and Motahari (2020) also noted that the opacity and complexity of machine learning models may have adverse consequences for asset managers in three ways. First, because the inferences made by machine learning models may be difficult to interpret, users may not be able to properly model and monitor the potential risk of systematic crashes. Second, machine learning may detect spurious and irrelevant patterns, which can lead to bad decisions. Third, performance attribution can be challenging because of the lack of interpretability. For example, some performance attribution systems are based on intuitive linear factor models, but the success of machine learning models mainly comes from less intuitive nonlinearity.

The Practical Benefits of a Cross-Sectional Approach

Despite the various challenges discussed in the preceding section, a cross-sectional approach to factor models has considerable merit for investment management. As noted previously, models based on cross-sectional factors have greater explanatory power than those based on time-series factors. Some of the other benefits of the cross-sectional approach are discussed next.

A focus on individual securities. Cross-sectional models in which stock returns are regressed on firm characteristics are readily applicable to individual stocks, and the values of most characteristics are directly observable rather than estimated (with the exception of parameters estimated from market model regression, such as beta). Investment practitioners are generally concerned with understanding and predicting the cross-sectional difference in expected returns of individual stocks.

Greater opportunity and better diversification. Cross-sectional models can be extended to encompass a large number of characteristics associated with stock returns. In contrast, the standard time-series factor models, such as those used by Fama and French (1993, 2015) and the *q*-factor model of Hou, Xue, and Zhang (2015), typically limit the number of factors to five or fewer. Parsimony is a virtue of a good model.[19] In practice, however, exploring more factors may allow one to take fuller advantage of the market's multidimensionality (Jacobs and Levy 1988, 2014a).

[18] Multiple regression can also accommodate nonlinearity with transformations of explanatory variables and interaction terms.

[19] Aside from parsimony, degrees of freedom may also limit the number of time-series factors. The degrees of freedom increase with the number of observations—that is, the number of time periods (T) in time-series regression or the number of stocks (N) in cross-sectional regression. For each additional factor added to a regression, there is a loss of one degree of freedom, which lowers the precision of the estimates. Because T is often far smaller than N, limiting the number of factors and thus saving the degrees of freedom is particularly important for time-series models.

A multidimensional portfolio with exposures to a large number of factors can exploit more opportunities than a portfolio based on only one or a few targeted factors, such as a smart beta portfolio. Furthermore, a multidimensional portfolio can benefit from diversification across numerous factors. It is less susceptible than a smart beta portfolio, for example, to the poor performance of any one factor. As some factors underperform, others may outperform, fostering greater consistency of performance.[20]

Fama and French (2018) documented that, although the monthly equity market premium and the monthly premiums of value and small stocks are on average positive over the period 1963–2016, their volatility is so high that the chances of having negative premiums are substantial for 3- to 5-year periods and nontrivial even for a 10-year period. In this regard, Fama and French (2020b) reported that the value premium was, on average, much lower in the second half of the 1963–2019 period. Blitz (2020) documented that the return on each of the Fama–French five factors (excluding the market factor) fell well short of its long-term average during the 2010–2019 period, delivering on average a negative return. Yet, many other factors not part of the five-factor model—including low risk, price momentum, earnings momentum, analyst revisions, seasonals, and short-term reversals—delivered positive returns over the same period.

Given the market's multidimensionality, including more factors can also provide valuable insights into portfolio performance evaluation. A performance attribution analysis using numerous factors, including industry factors, provides an appropriate framework for analyzing the sources of returns to a multidimensional portfolio.

Better forecasts. A cross-sectional approach based on many firm characteristics provides investors with a coherent framework to obtain a composite estimate of a stock's expected return. As Markowitz said in respect to Jacobs and Levy (1988), "such disentangling of multiple equity attributes improves estimates of expected return" (Markowitz 2000, 2017). Later, Lewellen (2015) reported that expected returns derived from cross-sectional regression have strong predictive power for actual returns. A cross-sectional approach also provides the flexibility to incorporate transient factor signals (i.e., temporary drivers of security returns). For example, the COVID-19 pandemic in 2020 has had a huge impact on individual stocks, but its impact is likely to diminish considerably as infections subside. Time-series factor models, which assume all factor premiums are persistent over time, are not well suited to accommodate transient factors.

Usefulness to anomaly and factor research. Cross-sectional regression of returns on firm characteristics, as implemented by Jacobs and Levy (1988) and Fama and French (1992), is useful for anomaly or factor research because it allows academics and practitioners to test a new factor while controlling for the effects of other factors. For example, Fama and French (2008) ran cross-sectional regressions to investigate several anomalies. Time-series factors included in asset pricing models, as Fama and French (2020a) noted, were often first discovered by cross-sectional regression.

Allowing for evolving firm characteristics. Cross-sectional models also allow for time-varying firm characteristics. Firm characteristics evolve as a firm moves through the various stages of its life cycle. For example, beta tends to converge to the market beta as a firm matures (Blume 1975); a small company grows into a large company by becoming more profitable and successful; a growth stock turns into a value stock by transitioning to the low-growth stage; a firm with weaker profitability emerges as a firm with robust profitability by entering a high-growth stage; and a firm with aggressive investment becomes a firm with conservative investment when it reaches the maturity stage. Fama and French's (2015) five factors were inspired by these characteristics;

[20] For a comparison of smart beta strategies and multidimensional strategies (which, in the spirit of smart beta, could be called "smart alpha"), see Jacobs and Levy (2014b).

however, the time-series regression of their five-factor model is not able to capture the evolution of these characteristics for each firm over time.[21]

Dynamic portfolios. The use of cross-sectional factor returns allows for the construction of dynamic portfolios (Jacobs and Levy 2014b). The time series of disentangled factor returns derived from cross-sectional regression can be used to generate factor forecasts that can be conditioned on market and economic information (Jacobs and Levy 1989b). These factor forecasts can be combined with the observed values of each firm's characteristics to obtain the predicted return for each stock. Portfolios constructed from these return forecasts can adapt to varying market and economic conditions. In contrast, static strategies that assume that the historical average of each factor return will persist tend to misestimate prospective returns because the factor returns can fluctuate significantly as market conditions change. The advantage of dynamic portfolios over static portfolios mirrors the skeptical view of Cornell (2020) that the historical relations between factors and stock returns are of limited use because the relationships are not stationary.

CONCLUSION

Jacobs and Levy (1988) disentangled stock returns across numerous anomalies (or firm characteristic factors) simultaneously, separating each potential source of return via cross-sectional analysis from the background noise created by other factors. A recent Fama and French article (2020a) found that models based on cross-sectional factors offer greater explanatory power for equity returns than do models based on time-series factors.

Disentangling controls for cross-contamination of factors and results in pure returns that are additive and can be more predictive than the estimates from naïve single-factor analyses or the Fama–French portfolio sorts because the market has greater dimensionality than naïve returns or sorts can accommodate. For factor returns that are influenced by market and economic conditions, cross-sectional analysis and the resulting time series of the disentangled factor returns facilitate such conditional forecasts. Furthermore, the application of cross-sectional analysis to individual security prediction is direct, unlike the Fama–French time-series modeling of portfolios.

Disentangling stock returns across numerous firm characteristics facilitates the construction of multidimensional portfolios that can achieve diversification by combining the independent insights from many factors (Jacobs and Levy 2014a). Multidimensional portfolios that take into account time-varying factor returns challenge smart beta portfolios, which maintain exposures to only one or a few chosen factors, typically regardless of underlying conditions.

Jacobs and Levy (1988) pioneered the use of cross-sectional regression to disentangle the relationships between numerous factors and stock returns and the use of the resulting time series of the disentangled factor returns to make conditional forecasts of security returns (Jacobs and Levy 1989b). These ideas are as relevant today as they were more than three decades ago.[22]

---

[21] In the time-series regression of the Fama–French five-factor model, portfolio factor exposures are updated over time; as firms evolve, they transition to different portfolios, which have different exposures to the five factors.

[22] Fama and French's articles have made no reference to the much earlier cross-sectional work of Jacobs and Levy (1988, 1989b).

APPENDIX

EXHIBIT A1

List of Anomalies Examined by Jacobs and Levy (1988)

Anomaly	Variable Definition for Anomaly
Low P/E	The trailing year's fully diluted earnings per share divided by price
Small Size	The negative of the natural log of market capitalization
Yield	Dividend divided by price
Zero Yield	A binary indicator of zero yield
Neglect	The negative of the natural log of one plus the number of analysts
Low Price	The negative of the natural log of price
Book/Price	Common equity per share divided by price
Sales/Price	The trailing year's sales per share divided by price; industry relative
Cash/Price	The trailing year's per-share earnings plus depreciation and deferred taxes, divided by price
Sigma	Dispersion of error terms from a rolling 60-month market model regression
Beta	Beta estimated from a rolling 60-month market model regression
Coskewness	$\Sigma (R_i - \bar{R}_i)(R_m - \bar{R}_m)^2 / \Sigma (R_m - \bar{R}_m)^3$, where R_i is the excess stock return, R_m is the S&P 500 excess return, and $\bar{R}_i$ and $\bar{R}_m$ are rolling 60-month arithmetic averages.
Controversy	The standard deviation of analysts' FY1 earnings estimates, normalized by price
Trend in Estimates (–1)	The change in consensus FY1 earnings estimate in month $m - 1$
Trend in Estimates (–2)	The change in consensus FY1 earnings estimate in month $m - 2$
Trend in Estimates (–3)	The change in consensus FY1 earnings estimate in month $m - 3$
Earnings Surprise (–1)	Actual earnings minus consensus estimate, normalized by price; announced in month $m - 1$
Earnings Surprise (–2)	Actual earnings minus consensus estimate, normalized by price; announced in month $m - 2$
Earnings Surprise (–3)	Actual earnings minus consensus estimate, normalized by price; announced in month $m - 3$
Earnings Torpedo	The change from earnings per share last reported to FY1 consensus estimate, normalized by price
Relative Strength	The alpha intercept from a rolling 60-month market model regression
Residual Reversal (–1)	The residual from the market model regression for month $m - 1$
Residual Reversal (–2)	The residual from the market model regression for month $m - 2$
Short-Term Tax	A measure of potential short-term tax-loss selling pressure
Long-Term Tax	A measure of potential long-term tax-loss selling pressure

ACKNOWLEDGMENTS

We are thankful for the encouragement of Kent Daniel of Columbia University, Frank Fabozzi of EDHEC Business School, Andy Lo of MIT, Marty Gruber of NYU, Bill Sharpe of Stanford University, Bill Ziemba of the University of British Columbia, Brad Cornell of UCLA, Allan Timmermann of UC San Diego, Elroy Dimson of the University of Cambridge, John Hand of the University of North Carolina, Sheridan Titman of the University of Texas at Austin, Chris Geczy, Mike Gibbons, Craig MacKinlay, and Jessica Wachter of The Wharton School at the University of Pennsylvania, Steve Horan of the CFA Institute, Ingrid Tierens of Goldman Sachs, and Edward O. Thorp. We are also grateful for the assistance of Sangwoo Lee, the comments of Phil Davies and Christian Xu, and the editorial assistance of Judith Kimball and David Landis, all from Jacobs Levy Equity Management.

REFERENCES

Axioma. “Neutralizing Unintended Bets in Factor ETFs: Purifying the Target Signal Improves Performance.” 2012.

Bartram, S. M., J. Branke, and M. Motahari. *Artificial Intelligence in Asset Management*. CFA Institute Research Foundation, 2020.

Blitz, D. 2020. “Factor Performance 2010–2019: A Lost Decade?” *The Journal of Index Investing* 11 (2): 57–65.

Blume, M. E. 1970. “Portfolio Theory: A Step Toward Its Practical Application.” *The Journal of Business* 43 (2): 152–173.

——. 1975. “Betas and Their Regression Tendencies.” *The Journal of Finance* 30 (3): 785–795.

Chinco, A., A. D. Clark-Joseph, and M. Ye. 2019. “Sparse Signals in the Cross-Section of Returns.” *The Journal of Finance* 74 (1): 449–492.

Chordia, T., A. Goyal, and J. Shanken. “Cross-Sectional Asset Pricing with Individual Stocks: Betas versus Characteristics.” Working paper, November 2017.

Clarke, R., H. de Silva, and S. Thorley. 2017. “Pure Factor Portfolios and Multivariate Regression Analysis.” *The Journal of Portfolio Management* 43 (3): 16–31.

Cochrane, J. H. *Asset Pricing*, revised ed. Princeton, NJ: Princeton University Press, 2005.

——. 2011. “Presidential Address: Discount Rates.” *The Journal of Finance* 66 (4): 1047–1108.

Cornell, B. 2020. “Stock Characteristics and Stock Returns: A Skeptic’s Look at the Cross Section of Expected Returns.” *The Journal of Portfolio Management* 46 (7): 131–142.

Daniel, K., and S. Titman. 1997. “Evidence on the Characteristics of Cross Sectional Variation in Stock Returns.” *The Journal of Finance* 52 (1): 1–33.

Davis, J. L., E. F. Fama, and K. R. French. 2000. “Characteristics, Covariances, and Average Returns: 1929–1997.” *The Journal of Finance* 55 (1): 389–406.

Fama, E. F. 1970. “Efficient Capital Markets: A Review of Theory and Empirical Work.” *The Journal of Finance* 25 (2): 383–417.

Fama, E. F., and K. R. French. 1992. “The Cross-Section of Expected Stock Returns.” *The Journal of Finance* 47 (2): 427–465.

——. 1993. “Common Risk Factors in the Returns of Stocks and Bonds.” *Journal of Financial Economics* 33 (1): 3–56.

——. 2008. “Dissecting Anomalies.” *The Journal of Finance* 63 (4): 1653–1678.

——. 2015. “A Five-Factor Asset Pricing Model.” *Journal of Financial Economics* 116 (1): 1–22.

——. 2018. “Volatility Lessons.” *Financial Analysts Journal* 74 (3): 42–53.

——. 2020a. “Comparing Cross-Section and Time-Series Factor Models.” *The Review of Financial Studies* 33 (5): 1891–1926.

——. 2020b. “The Value Premium.” Working Paper 20-01, Fama–Miller Center, January.

Fama, E. F., and J. D. MacBeth. 1973. “Risk, Return, and Equilibrium: Empirical Tests.” *Journal of Political Economy* 81 (3): 607–636.

Feng, G., S. Giglio, and D. Xiu. 2020. “Taming the Factor Zoo: A Test of New Factors.” *The Journal of Finance* 75 (3): 1327–1370.

Gibbons, M. R., S. A. Ross, and J. Shanken. 1989. “A Test of the Efficiency of a Given Portfolio.” *Econometrica* 57 (5): 1121–1152.

Green, J., J. R. M. Hand, and X. F. Zhang. 2017. “The Characteristics That Provide Independent Information about Average U.S. Monthly Stock Returns.” *The Review of Financial Studies* 30 (12): 4389–4436.

Gu, S., B. Kelly, and D. Xiu. 2020. "Empirical Asset Pricing via Machine Learning." *The Review of Financial Studies* 33 (5): 2223–2273.

Harvey, C. R. 1989. "Time-Varying Conditional Covariances in Tests of Asset Pricing Models." *Journal of Financial Economics* 24 (2): 289–317.

Harvey, C. R., Y. Liu, and H. Zhu. 2016. "… and the Cross-Section of Expected Returns." *The Review of Financial Studies* 29 (1): 5–68.

Haugen, R. A., and N. L. Baker. 1996. "Commonality in the Determinants of Expected Stock Returns." *Journal of Financial Economics* 41 (3): 401–439.

Hou, K., C. Xue, and L. Zhang. 2015. "Digesting Anomalies: An Investment Approach." *The Review of Financial Studies* 28 (3): 650–705.

Israel, R., B. T. Kelly, and T. J. Moskowitz. 2020. "Can Machines 'Learn' Finance?" *Journal of Investment Management* 18 (2): 23–36.

J.P. Morgan. "The Quest for Pure Equity Factor Exposure." 2019.

Jacobs, B. I., and K. N. Levy. 1988. "Disentangling Equity Return Regularities: New Insights and Investment Opportunities." *Financial Analysts Journal* 44 (3): 18–43.

——. 1989a. "The Complexity of the Stock Market." *The Journal of Portfolio Management* 16 (1): 19–27.

——. 1989b. "Forecasting the Size Effect." *Financial Analysts Journal* 45 (3): 38–54.

——. 1990. "Disentangling Equity Return Regularities: New Insights and Investment Opportunities." *Security Analysts Journal* 28 (3/4): 30–49, 29–50 (Japanese translation).

——. 2014a. "Investing in a Multidimensional Market." *Financial Analysts Journal* 70 (6): 6–12.

——. 2014b. "Smart Beta versus Smart Alpha." *The Journal of Portfolio Management* 40 (4): 4–7.

Jagannathan, R., and Z. Wang. 1996. "The Conditional CAPM and the Cross-Section of Expected Returns." *The Journal of Finance* 51 (1): 3–53.

King, B. F. 1966. "Market and Industry Factors in Stock Price Behavior." *The Journal of Business* 39 (1): 139–190.

Lewellen, J. 2015. "The Cross Section of Expected Stock Returns." *Critical Finance Review* 4 (1): 1–44.

Light, N., D. Maslov, and O. Rytchkov. 2017. "Aggregation of Information about the Cross Section of Stock Returns: A Latent Variable Approach." *The Review of Financial Studies* 30 (4): 1339–1381.

Markowitz, H. M. "Foreword." In *Equity Management: Quantitative Analysis for Stock Selection*, by B. I. Jacobs and K. N. Levy. New York: McGraw-Hill, 2000. (Chinese translation. Beijing: Chinese Machine Press, 2006).

——. "Foreword." In *Equity Management: The Art and Science of Modern Quantitative Investing*, 2nd ed., by B. I. Jacobs and K. N. Levy. New York: McGraw-Hill, 2017.

Markowitz, H. M., and E. van Dijk. "Risk–Return Analysis." In *Handbook of Asset and Liability Management Volume 1: Theory and Methodology*, edited by S. A. Zenios and W. T. Ziemba. Amsterdam: North-Holland, 2006.

McLean, R. D., and J. Pontiff. 2016. "Does Academic Research Destroy Stock Return Predictability?" *The Journal of Finance* 71 (1): 5–32.

MSCI. "The Barra US Equity Model (USE4)." 2011.

Nomura. "50 Shades of Factor Construction." 2012.

Petersen, M. A. 2009. "Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches." *The Review of Financial Studies* 22 (1): 435–480.

Rosenberg, B. 1974. "Extra-Market Components of Covariance in Security Returns." *Journal of Financial and Quantitative Analysis* 9 (2): 263–274.

Rosenberg, B., and V. Marathe. 1976. "Common Factors in Security Returns: Microeconomic Determinants and Macroeconomic Correlates." Working Paper No. 44, University of California Institute of Business and Economic Research, Research Program in Finance.

Rosenberg, B., K. Reid, and R. Lanstein. 1985. "Persuasive Evidence of Market Inefficiency." *The Journal of Portfolio Management* 11 (3): 9–16.

Schwartz, S. L., and W. T. Ziemba. "Predicting Returns on the Tokyo Stock Exchange." In *Security Market Imperfections in Worldwide Equity Markets*, edited by D. B. Keim and W. T. Ziemba. Cambridge, UK: Cambridge University Press, 2000.

Sharpe, W. F. 1964. "Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk." *The Journal of Finance* 19 (3): 425–442.

——. 1982. "Factors in New York Stock Exchange Security Returns, 1931–1979." *The Journal of Portfolio Management* 8 (4): 5–19.

Société Générale. "Global Style Counselling." 2017.

White, J. A. "How Jacobs and Levy Crunch Stocks for Buying—and Selling." *The Wall Street Journal*, March 20, 1991.

Wolfe Research. "Déjà Vu, Risk-On and Risk-Off (RORO) Again." 2020.

Ziemba, R. E. S., and W. T. Ziemba. *Scenarios for Risk Management and Global Investment Strategies*. Chichester, UK: Wiley, 2007.

Ziemba, W. T., and S. L. Schwartz. *Invest Japan*. Chicago: Probus Publishing Company, 1992.