Explaining Stock Returns: A Literature Survey

James L. Davis

I. Introduction

My objective in writing this survey is to provide an overview of the work that has been done in an important area of financial markets research—explaining the behavior of common stock returns. I have tried to make this survey as complete as possible, without getting bogged down in a lot of technical details. Since this area of research has been very active for the past several years, describing all of the work that has been done is not feasible. I have tried to include the most important research in my discussion, but in doing so, I have left out some very good papers. What follows is my attempt to adequately discuss all the main ideas in as concise a manner as possible.

The next section provides an overview of the financial theory that underlies the behavior of stock returns.[1] The remainder of the paper is concerned with the results of numerous empirical studies that have been published during the past quarter-century. Throughout this discussion of empirical results, the link back to financial theory is maintained. Interesting recent studies are included in the discussion, even though they have not yet received the level of attention that has been given to many of the older studies.

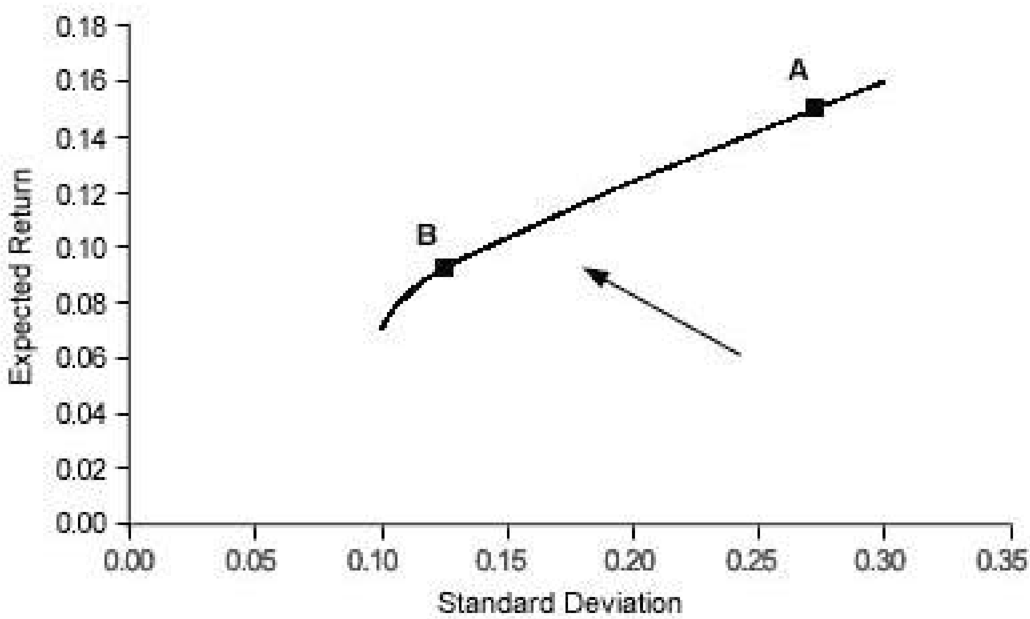
II. Theoretical Background

Markowitz Portfolio Selection

Any discussion of the theory of stock price behavior has to start with Markowitz (1952, 1959). The Markowitz model is a single-period model, where an investor forms a portfolio at the beginning of the period. The investor's objective is to maximize the portfolio's expected return, subject to an acceptable level of risk (or minimize risk, subject to an acceptable expected return). The assumption of a single time period, coupled with assumptions about the investor's attitude toward risk, allows risk to be measured by the variance (or standard deviation) of the portfolio's return. Thus, as indicated by the arrow in Figure 1, the investor is trying to go as far northwest as possible.

As securities are added to a portfolio, the expected return and standard deviation change in very specific ways, based on the way in which the added securities co-vary with the other securities in the portfolio. The best that an investor can do (i.e., the furthest northwest a portfolio can be) is bounded by a curve that is the upper half of a hyperbola, as shown in Figure 1. This curve is known as the efficient frontier. According to the Markowitz model, investors select portfolios along this curve, according to their tolerance for risk. An investor who can live with a lot of risk might choose portfolio A, while a more risk-averse investor would be more likely to choose portfolio B. One of the major insights of the Markowitz model is that it is a security's expected return, coupled with how it co-varies with other securities, that determines how it is added to investor portfolios.

Figure 1
Markowitz Portfolio Selection



Capital Asset Pricing Model

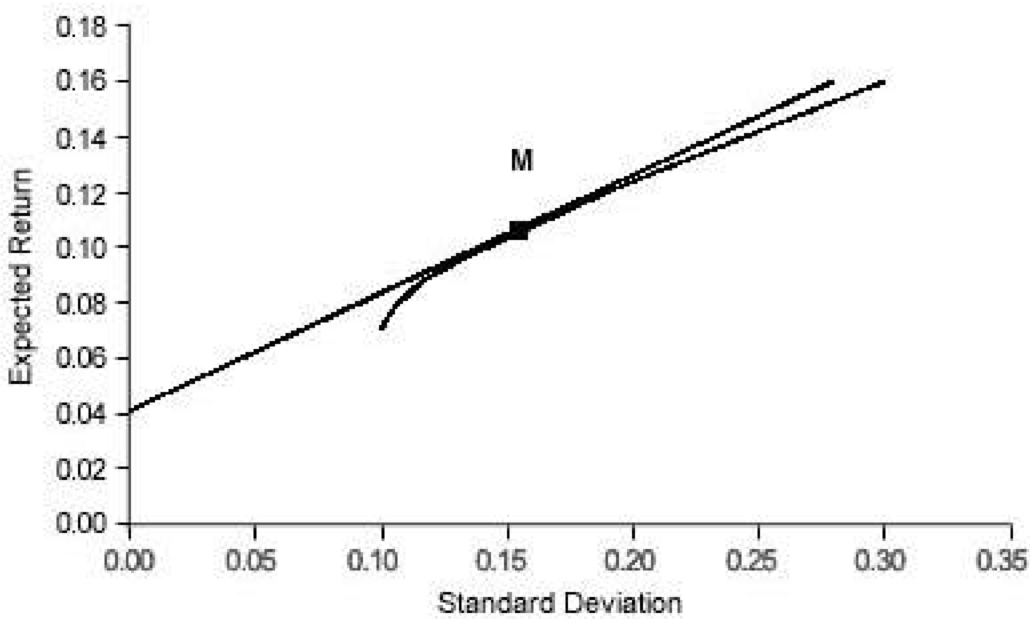
Building on the Markowitz framework, Sharpe (1964), Lintner (1965) and Mossin (1966) independently developed what has come to be known as the Capital Asset Pricing Model (CAPM). This model assumes that investors use the logic of Markowitz in forming portfolios. It further assumes that there is an asset (the risk free asset) that has a certain return. With a risk free asset, the efficient frontier in Figure 1 is no longer the best that investors can do. The straight line in Figure 2, which has the risk free rate as its intercept and is tangent to the efficient frontier, is now the northwest boundary of the investment opportunity set. Investors choose portfolios along this line (the capital market line), which shows combinations of the risk free asset and the risky portfolio M. In order for markets to be in equilibrium (quantity supplied = quantity demanded), the portfolio M must be the market portfolio of all risky assets. So, all investors combine the market portfolio and the risk free asset, and the only risk that investors are paid for bearing is the risk associated with the market portfolio. This leads to the CAPM equation:

(CAPM)

$$E(R_j) = R_f + \beta_j\ [E(R_m) - R_f]$$

$E(R_j)$ and $E(R_m)$ are the expected returns to asset j and the market portfolio, respectively, R_f is the risk free rate, and β_j is the beta coefficient for asset j. β_j measures the tendency of asset j to co-vary with the market portfolio. It represents the part of the asset's risk that cannot be diversified away, and this is the risk that investors are compensated for bearing. The CAPM equation says that the expected return of any risky asset is a linear function of its tendency to co-vary with the market portfolio. So, if the CAPM is an accurate description of the way assets are priced, this positive linear relation should be observed when average portfolio returns are compared to portfolio betas. Further, when beta is included as an explanatory variable, no other variable should be able to explain cross-sectional differences in average returns. Beta should be all that matters in a CAPM world.

Figure 2
Capital Market Line



Arbitrage Pricing Theory

While the CAPM is a simple model that is based on sound reasoning, some of the assumptions that underlie the model are unrealistic.[2] Some extensions of the basic CAPM were proposed that relaxed one or more of these assumptions (e.g., Black, 1972). Instead of simply extending an existing theory, Ross (1976a, 1976b) addresses this concern by developing a completely different model: the Arbitrage Pricing Theory (APT). Unlike the CAPM, which is a model of financial market equilibrium, the APT starts with the premise that arbitrage[3] opportunities should not be present in efficient financial markets. This assumption is much less restrictive than those required to derive the CAPM.

The APT starts by assuming that there are *n* factors which cause asset returns to systematically deviate from their expected values. The theory does not specify how large the number *n* is, nor does it identify the factors. It simply assumes that these n factors cause returns to vary together. There may be other, firm-specific reasons for returns to differ from their expected values, but these firm-specific deviations are not related across stocks. Since the firm-specific deviations are not related to one another, all return variation not related to the *n* common factors can be diversified away. Based on these assumptions, Ross shows that, in order to prevent arbitrage, an asset's expected return must be a linear function of its sensitivity to the *n* common factors:

(APT)

$$E(R_j) = R_f + \beta_{j1} \lambda_1 + \beta_{j2} \lambda_2 + \ldots.. + \beta_{jn} \lambda_n$$

$E(R_j)$ and R_f are defined as before. Each β_{jk} coefficient represents the sensitivity of asset j to risk factor k, and λ_k represents the risk premium for factor k. As with the CAPM, we have an expression for expected return that is a linear function of the asset's sensitivity to systematic risk. Under the assumptions of APT, there are *n* sources of systematic risk, where there is only one in a CAPM world.

Intertemporal Capital Asset Pricing Model

Both the CAPM and the APT are static, or single-period models. As such, they ignore the multi-period nature of participation in the capital markets. Merton's (1973) intertemporal capital asset pricing model (ICAPM) was developed to capture this multi-period aspect of financial market equilibrium. The ICAPM framework recognizes that the investment opportunity set (see Figures 1 and 2) might shift over time, and investors would like to hedge themselves against unfavorable shifts in the set of available investments. If a particular security tends to have high returns when bad things happen to the investment opportunity set, investors would want to hold this security as a hedge. This increased demand would result in a higher equilibrium price for the security (all else constant). One of the main insights of the ICAPM is the need to reflect this hedging demand in the asset pricing equation. The resulting model is:

(ICAPM)

$$E(R_j) = R_f + \beta_{jM}\,\lambda_M + \beta_{j2}\,\lambda_2 + \ldots.. + \beta_{jn}\,\lambda_n$$

Note that the form of the ICAPM is very similar to that of the APT. There are subtle differences, however. The first factor of the ICAPM is explicitly identified as being related to the market portfolio. Further, while the APT gives little guidance as to the number and nature of factors, the factors that appear in the ICAPM are those that satisfy the following conditions:

1. They describe the evolution of the investment opportunity set over time.
2. Investors care enough about them to hedge their effects.

For example, there might be a priced factor for unexpected changes in the real interest rate. Such a change would certainly shift the investment opportunity set (for example, the intercept of the line in Figure 2 would move), and the effect would be pervasive enough that investors would want to protect themselves from the negative consequences. We still don't know exactly how many factors there are, but the ICAPM at least gives us some guidance.

Consumption-Oriented Capital Asset Pricing Model

The consumption-based model of Breeden (1979) provides a logical extension of the previous work in asset pricing. Breeden's model is based on the intuition that an extra dollar of consumption is worth more to a consumer when the level of aggregate consumption is low. When things are going really well and many people can afford a comfortable standard of living, another dollar of consumption doesn't make us feel very much better off. But when times are hard, a few extra dollars to spend on consumption goods is very welcome. Based on this "diminishing marginal utility of consumption," securities that have high returns when aggregate consumption is low will be demanded by investors, bidding up their prices (and lowering their expected returns). In contrast, stocks that co-vary positively with aggregate consumption will require higher expected returns, since they provide high returns during states of the economy where the high returns do the least good.

Based on this line of reasoning, Breeden derives a consumption-based capital asset pricing model (CCAPM) of the form:

(CCAPM)

$$E(R_j) = R_f + \beta_{jC}\,[E(R_m) - R_f]$$

In this model, β_{jC} measures the sensitivity of the return of asset j to changes in aggregate consumption. β_{jC} is referred to as the consumption beta of asset j, and the CCAPM's main result is that expected returns should be a linear function of consumption betas.[4]

Despite the intuitive appeal of the consumption-based model, empirical tests have not supported its predictions (see Breeden, Gibbons and Litzenberger, 1989). Accordingly, consumption-based asset pricing has not received as much attention in practice as the other models discussed here. More will be said about the CCAPM later.

In spite of the unrealistic assumptions underlying the single-period CAPM, it still became the most widely used asset pricing model within a few years after its development. Its simplicity, coupled with empirical tests that supported most of its predictions (for example, see Fama and MacBeth, 1973), made it the most widely taught asset pricing model in schools of business. The APT was tested in a number of empirical studies, but the CAPM received most of the financial world's attention.

III. Early Empirical Work

Early cross-sectional studies of stock returns (e.g., Nicholson, 1960) did not receive a great deal of attention, due to the small samples used to conduct the empirical tests. It was not until the CRSP and Compustat[5] databases became available that researchers could construct samples large enough (and of sufficient quality) to produce reliable results. Consequently, for a few years after the development of the CAPM, there was no reliable way to test the model's predictions against variables like book-to-market equity or earnings/price.

Earnings / Price

One of the early studies that contradicted the predictions of the CAPM was Basu (1977). Using a sample period that stretched from April 1957 to March 1971, Basu showed that stocks with high earnings/price ratios (or low P/E ratios) earned significantly higher returns than stocks with low earnings/price ratios. His results indicated that differences in beta could not explain these return differences. In a follow-up study, Basu (1983) showed that this "E/P effect" is not just observed among small-cap stocks. A later study by Jaffe, Keim and Westerfield (1989) confirmed this finding and also showed that the E/P effect does not just appear in the month of January, as had been claimed by some researchers. The E/P effect is a direct contradiction of the CAPM; beta should be all that matters.

Firm Size

Banz (1981) uncovered another apparent contradiction of the CAPM by showing that the stocks of firms with low market capitalizations have higher average returns than large-cap stocks. Other researchers (e.g., Basu, 1983) showed that the size effect is distinct from the E/P effect discussed above. Small firms tend to have higher returns, even after controlling for E/P.

Proponents of the CAPM are quick to point out that small firms tend to have higher betas than large firms, so we would expect to see higher average returns for small firms. However, the beta differences are not large enough to explain the observed return differences. Once again, the CAPM predictions are violated.

Long-Term Return Reversals

DeBondt and Thaler (1985) identify "losers" as stocks that have had poor returns over the past three to five years. "Winners" are those stocks that had high returns over a similar period. The main result of DeBondt and Thaler is that losers have much higher average returns than winners over the next three to five years. Chopra, Lakonishok and Ritter (1992) show that beta cannot account for this difference in average returns. This tendency of returns to reverse over long horizons (i.e., losers become winners) is yet another contradiction of the CAPM. Losers would have to have much higher betas than winners in order to justify the return difference. Chopra, Lakonishok and Ritter (1992) show that the beta difference required to save the CAPM is not there.

Book-to-Market Equity

Rosenberg, Reid and Lanstein (1985) provide yet another piece of evidence against the CAPM by showing that stocks with high ratios of book value of common equity to market value of common equity (also known as book-to-market equity, or BtM) have significantly higher returns than stocks with low BtM. Since the sample period for this study is fairly short (1973–1984), the empirical results did not receive as much attention as some of the other studies discussed above. However, when Chan, Hamao and Lakonishok (1991) found similar results in the Japanese market, BtM began to receive serious attention as a variable that could produce dispersion in average returns.

Leverage

Bhandari (1988) finds that firms with high leverage (high debt/equity ratios) have higher average returns than firms with low leverage for the 1948–1979 period. This result persists after size and beta are included as explanatory variables. High leverage increases the riskiness of a firm's equity, but this increased risk should be reflected in a higher beta coefficient. Consequently, Bhandari's results are yet another deviation from the CAPM predictions.

Momentum

Jegadeesh (1990) found that stock returns tend to exhibit short-term momentum; stocks that have done well over the previous few months continue to have high returns over the next month. In contrast, stocks that have had low returns in recent months tend to continue the poor performance for another month. A study by Jegadeesh and Titman (1993) would later confirm these results, showing that the momentum lasts for more than just one month. Their study also indicates that the momentum is stronger for firms that have had poor recent performance. The tendency of recent good performance to continue is weaker. Note that the pattern here is the opposite of that found in the long-term overreaction papers. In those studies, long-term losers outperform long-term winners. In the momentum studies, short-term winners outperform short-term losers.

The studies discussed in this section cast doubt on the ability of the CAPM to explain equilibrium relationships in the financial markets. These other variables should not be able to explain average returns better than beta. Stocks with high E/P, high BtM, high leverage, etc. should not outperform other stocks to the extent that they have. To make matters worse, Reinganum (1981) shows that the positive relation between beta and return that was observed in earlier studies (e.g., Fama and MacBeth, 1973) has weakened in more recent years. In spite of all this negative evidence, the CAPM was still the default view for most financial economists and practitioners going into the 1990s. That was about to change.

IV. A Turning Point

In 1992, an influential paper was published that pulled together much of the earlier empirical work. Fama and French (1992) brought together size, leverage, E/P, BtM, and beta in a single cross-sectional study. Their results were controversial. First, they showed that the previously documented positive relation between beta and average return was an artifact of the negative correlation between firm size and beta. When this correlation is accounted for, the relation between beta and return disappears. Figures 3 and 4 show this result. Figure 3 plots beta and average return for twelve portfolios formed by ranking stocks on firm size. The positive relation between return and beta is highly linear, as predicted by the CAPM. Based on this evidence, it appears that the CAPM nicely explains the higher returns that small firms have earned. Figure 4 plots average return and beta for portfolios formed by ranking on both firm size and beta, so that each portfolio contains stocks that are similar in both their betas and their market values. This chart shows that when beta is allowed to vary in a manner unrelated to size, the positive, linear beta-return relation disappears. This result contradicts the central prediction of the single-period CAPM.

Given that beta does a poor job of explaining average returns, what variables can do a better job? This is the second main point of the Fama-French study. They compared the explanatory power of size, leverage, E/P, BtM, and beta in cross-sectional regressions that spanned the 1963–1990 period. Their results indicate that BtM and size are the variables that have the strongest relation to returns. The explanatory power of the other variables vanishes when these two variables are included in the regressions. The cross-section of average stock returns can be nicely described by two variables.

The Fama-French (1992) results dealt a severe blow to the view that the single-period CAPM is the way securities are actually priced. The model that has been taught more than any other in business school doesn't seem to work.

Figure 3
Beta and Average Return for Portfolios formed on Size

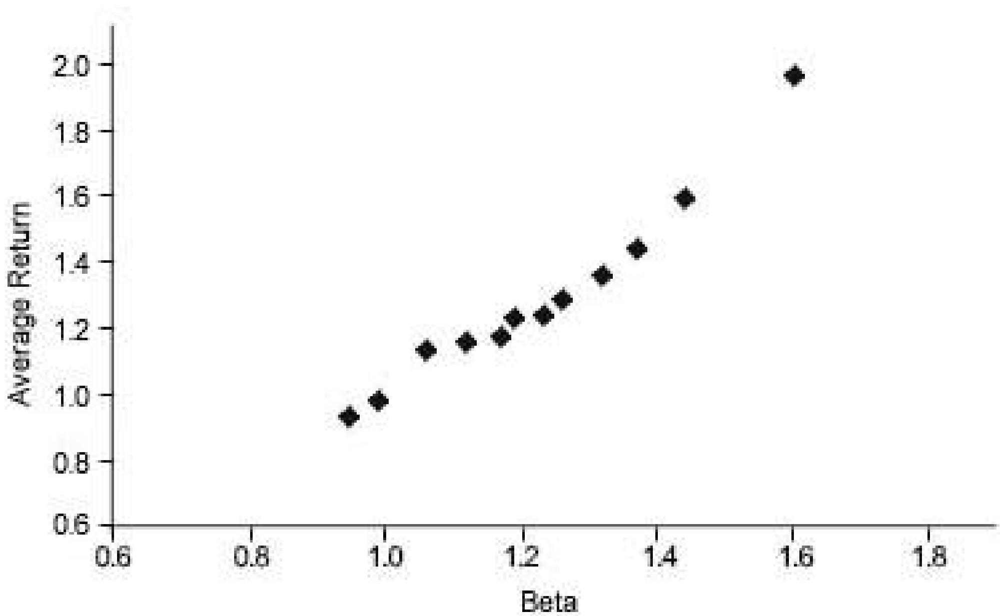
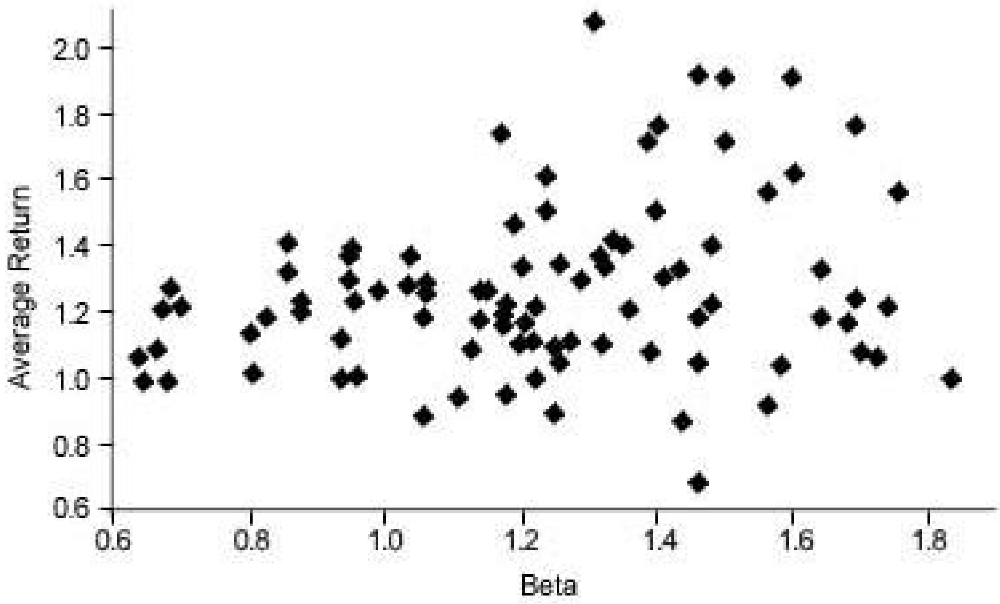


Figure 4
Beta and Average Return for Portfolios formed on Size and Beta



V. The Attack

Whenever a well-established paradigm is questioned, the reaction will be swift and often aggressive. It is no different in the world of academic finance. This is a good thing, as long as the reaction is honest and straightforward. Well-established prior beliefs should not be abandoned unless the contrary evidence is rigorously analyzed and found to be valid.

Data Mining

If an academic paper is judged by the amount of discussion that it generates, then Fama-French (1992) was an unparalleled success. The reaction was not timid. One of the first replies was from Black (1993a, 1993b), who suggested that the Fama-French results were likely an artifact of data mining. Hundreds of researchers, in an attempt to write publishable papers, spend a great deal of time looking for relationships between stock returns and other variables. Only the successful tests are submitted for publication; the unsuccessful ones never see the light of day. A few variables are bound to show a statistical relation to returns, just by chance. Since Fama and French chose their explanatory variables based on the results of earlier empirical studies, the observed explanatory power of these variables could be due to a massive data mining exercise on the part of the authors of these earlier studies. Based on this, Black contended that some of the statistical tests in Fama-French (1992) were not properly specified. He also suggested that, since the relations between returns and size, BtM, etc. were likely an artifact of data mining, they would disappear if another time period or another data source were analyzed. MacKinlay (1995) also mentions data mining as a potential cause of the observed results.

Another criticism of the Fama-French results came from Kothari, Shanken and Sloan (1995). Their attack proceeded along two main fronts: Survivorship bias and beta mis-measurement.

Survivorship Bias

It is a well known fact that the Compustat database suffers from a survivorship bias, due to the way firms are added to the database.[6] As described by Banz and Breen (1986), Breen and Korajczyk (1994), and Kothari, Shanken and Sloan (1995), firms are typically brought into the Compustat files with several years of historical data. Since the firms that are added to the database during a given year are firms that still exist, the backfilling of historical data for the previous several years biases the database towards firms that survived through those years. The firms that died during those years, and that were not already in the database, are never included. This "*ex post* selection bias" can have a significant effect on cross-sectional studies of stock returns. Kothari, Shanken and Sloan claim that the observed explanatory power of BtM is likely due to survivorship bias: Since many of the firms that are excluded from Compustat are firms that failed, it is likely that these firms had high BtM and low returns. Adding these firms to the database would reduce the explanatory power of BtM, possibly eliminating it.

Beta Estimation

The other main criticism of Fama and French (1992) put forth by Kothari, Shanken and Sloan (1995) is related to the estimation of beta. Levhari and Levy (1977) show that beta coefficients estimated with monthly returns are not the same as betas estimated with annual returns. Since they are different, the results of empirical studies will depend upon which beta estimation convention is used. Kothari, Shanken and Sloan argue that annual betas are more appropriate than monthly betas, since the investment horizon for a typical investor is probably closer to a year than a month. They show that the relation between beta and return is stronger when betas are estimated using annual returns.

Based on the data mining, selection bias, and beta estimation criticisms of Fama-French (1992), many researchers in the early-to-mid 1990s believed that the explanatory power of BtM should not be taken seriously. A number of authors argued that the CAPM was still the best model of expected returns, claiming that the empirical results contradicting the CAPM are unreliable.

VI. The Response

The give-and-take that followed Fama-French (1992) represents one of the more interesting strands of the academic finance literature. Many a graduate student found a dissertation topic buried in this debate. One of the nice aspects of this area of inquiry was the fact that most of the important questions could be answered, if the researcher could find the necessary data. The papers that were written in response to the criticisms of Fama-French (1992) have impacted both the practice of finance and the theoretical study of financial economics. Seldom has an area of academic inquiry had so much real-world application.

One of the early responses to the criticisms of Fama-French (1992) was Davis (1994), who constructed a database of book values for large U.S. industrial firms for the 1940-1963 period, a period for which the Compustat coverage is either poor or nonexistent. This database was constructed to be free of survivorship bias, and it covers a period that precedes the period studied by Fama and French. If the Fama-French results are a result of data mining, this independent time period should produce different results. A spurious relation in one period is not likely to carry over to a different period. Also, the beta coefficients in this study were estimated using annual returns to address one of Kothari, Shanken and Sloan's (1995) main points.

The results of Davis (1994) generally confirmed those of Fama and French (1992).[7] The explanatory power of BtM was observed in the 1940–1963 period, although the magnitude of the return dispersion was somewhat smaller. This is probably caused by the fact that the database for the "pre-Compustat" period contains only large firms. In addition, the relation between beta and average return was flat. Betas based on annual returns could not improve the CAPM's performance during the 1940-1963 period.

Chan, Jegadeesh and Lakonishok (1995) provide further evidence that the Fama-French (1992) results are not due to survivorship bias. Examining the 1968-1991 period, they find that, when firms on CRSP and Compustat are properly matched, there are not enough firms missing from Compustat to have a significant effect on the Fama-French results. They also form a dataset of large firms for this period that is free of survivorship bias. Using this dataset, they find a reliable BtM effect.

Barber and Lyon (1997) present a clever way to address the issue of data mining. Noting that empirical results that are caused by data mining should not carry over to other independent samples, they form a sample of financial firms for the 1973–1994 period and find a reliable BtM effect among these firms. Since financial firms were purposely excluded from the Fama-French sample, the results of Barber and Lyon provide independent evidence of the explanatory power of BtM.

Further independent evidence came from Fama and French (1998), who found a reliable BtM effect in several developed countries for the 1975-1995 period. They also found a reliable value premium in several emerging markets. Capaul, Rowley and Sharpe (1993) also found evidence of a BtM effect in the U.S. and five other developed countries for the 1981–1992 period. This international evidence casts even more doubt on the data mining criticisms of the U.S. results.

VII. The Explanations

Because of their controversial nature, the results of Fama and French (1992) were subjected to a high degree of scrutiny. Based on the papers that supported the Fama-French results, most researchers reached the conclusion that the size and book-to-market effects are real, since they have been observed over several decades in the U.S., and in other countries as well. The next topic to be debated is: Why? The issue is no longer *whether* size and BtM are able to produce cross-sectional dispersion in average returns, but *why*. The two primary explanations are risk and inefficiency.

The risk-based story starts with Fama and French (1993), who show that factors related to size and BtM are able to explain a significant amount of the common variation in stock returns. For the 1963–1991 period, they run three-factor regressions of the form:

$$R_{jt} - R_{ft} = a_j + b_j (R_{mt} - R_{ft}) + s_j SMB_t + h_j HML_t + e_{jt}$$

where R_{jt} is the return to portfolio j for month t, R_{ft} is the T-Bill return for month t, and R_{mt} is the return to the CRSP value weighted index for month t. SMB_t is the realization on a capitalization-based factor portfolio that buys small-cap stocks and sells large-cap stocks. Similarly, HML_t is the realization on a factor portfolio that buys high-BtM stocks and sells low-BtM stocks. The s_j and h_j coefficients measure the sensitivity of the portfolio's return to the small-minus-big and high-minus-low factors, respectively. Portfolios of value stocks will have a high value for h, while growth portfolios will have a negative h. Large-cap portfolios will load negatively on SMB (s_j will be negative), and small-cap portfolios will have a large positive value for s.

The Fama-French (1993) results support a risk-based explanation of the return dispersion produced by size and BtM. The 3-factor regressions tend to produce significant coefficients on all three factors, and regression R^2 values are close to 1 for most portfolios. This indicates that the three factors are capturing much of the common variation in portfolio returns. Therefore, it appears that SMB and HML are capturing independent sources of systematic risk. They behave like we would expect to see risk factors from the APT or ICAPM behave. The time series averages of SMB_t and HML_t can be interpreted as the average risk premiums for these risk factors (i.e., the λ's from the APT and ICAPM equations). According to the three-factor model, small-cap stocks and value stocks have high average returns because they are risky – they have high sensitivity to the risk factors that are being measured by SMB and HML.

In contrast to the risk-based story is the proposition that value stocks have higher returns than growth stocks because markets are not efficient. This position is well represented by Lakonishok, Shleifer and

Vishny (1994), who contend that investors naively extrapolate firms' past performance into the future. Value stocks typically have had poor past performance, and investors assume that this poor performance will continue. Then, when some of these poorly performing firms get things turned around, investors are surprised, and the stocks of these firms experience high returns. According to this hypothesis, the high returns to value stocks (and the low returns to growth stocks) are due to investors being systematically wrong about the future. An implication of this is that investors can increase returns *without increasing risk*, simply by buying value stocks and selling (or not buying) growth stocks.

Lakonishok, Shleifer and Vishny (1994) support their extrapolation-based story by showing that a two-way sort on cash flow/price and 5-year sales growth produces more dispersion in average returns than other variables (including BtM) for the 1968–1990 period. The sort on 5-year sales growth classifies firms according to past performance, and the sort on cash flow/price parses firms according to expected future performance. The extrapolation hypothesis says that firms with low sales growth and high cash flow/price should have the highest returns, since the poor historical performance (measured by sales growth) is projected into the future (reflected by high cash flow/price). Their results support this hypothesis. However, Davis (1994) shows that the two-way classification using sales growth and cash flow/price produces about the same return dispersion as a simple sort on cash flow/price for the 1940-1963 period. This out-of-sample evidence does not support the extrapolation hypothesis, which predicts that the two-way classification should produce more dispersion.

Several papers have been published in recent years supporting the risk-based story. Fama and French (1995) provide support for the risk hypothesis by showing that there are size and value factors in earnings as well as returns. This suggests that systematic variation in firms' cash flow streams may be associated with systematic variation in stock returns. Also, Fama and French (1996) show that the three-factor model can explain most of the departures from the CAPM predictions discussed in the recent financial literature, including the two-way sorts of Lakonishok, Shleifer and Vishny (1994). However, the three-factor model could not explain the short-term momentum in stock prices. The ability of the three-factor model to explain most of the observed cross-sectional empirical results supports a multi-factor risk model of expected returns. Still, it is not clear why the three-factor model cannot explain momentum.

Liew and Vassalou (2000) support the risk-based story by showing that SMB and HML are able to predict future GDP growth in some countries. However, the relation between these variables and GDP growth is weak in several countries, and it is nonexistent in the U.S. for the 1957-1998 period.

Daniel and Titman (1997) doubt the risk-based explanation. They contend that it is "characteristics, not covariances", that produce return dispersion. For example, the risk-based story says that high-BtM stocks have high average returns because they are sensitive to common variation in stock returns. In other words, the high returns are due to a high sensitivity to HML. In contrast, Daniel and Titman argue that high-BtM stocks have high returns due to some other reason (possibly overreaction), so that the high returns have nothing to do with systematic risk. In their opinion, it is the characteristic (high BtM) rather than the covariance (high sensitivity to HML) that is associated with high returns.

The cross-sectional correlation between BtM and HML sensitivity is quite high, so it is difficult to see which of these variables has more explanatory power for returns. Nevertheless, Daniel and Titman provide results suggesting that the characteristics-based story is more plausible for the 1973-1993 period. However, Davis, Fama and French (2000) show that the Daniel and Titman results are confined to their relatively short sample period. When the longer 1929-1997 period is examined, covariances show more explanatory power than characteristics. It is not clear why the shorter period produces different results, but the longer period should produce more reliable results, and these results favor the risk-based story.

VIII. Recent Developments

The research into stock price behavior and asset pricing continues, and a number of interesting results have surfaced recently. Perez-Quiros and Timmermann (2000) provide evidence that small firms have high average returns because they are more affected by tight credit market conditions. Small firms do not have the same access to domestic and international bond markets that are enjoyed by large firms. Since the availability of credit is tied to economic conditions, so that a credit contraction typically occurs near a recession, small firms would be very sensitive to systematic variation in credit market conditions. Thus, the high returns to small firms might be compensation for the high sensitivity to a credit-related risk factor.

A study by Elton, Gruber, Agrawal, and Mann (2001) reports a potentially important link between the equity and fixed income markets. If certain risk factors are pervasive enough to explain common variation in stock returns, it is reasonable to expect that these same risk factors would be at work in the bond market as well. Elton, et al. provide evidence that SMB and HML do just that. Their research isolates the portion of a bond's return that is due to changing risk premiums, and they show that this part of the bond's return is strongly related to SMB and HML. Not only does this result support the risk-based story, but it also suggests some interesting avenues for future research in fixed income portfolio management.

In an interesting recent study, Lettau and Ludvigson (2001) show that a consumption-oriented capital asset pricing model (CCAPM) that allows expected returns to vary over time provides a nice cross-sectional explanation of equity returns. They use the ratio of aggregate consumption to wealth as a "conditioning variable" to model the evolution of expected returns over time. The relation between the consumption/wealth ratio and expected returns is straightforward. If investors expect returns to be high in the future, they would be more likely to raise their consumption level (relative to their level of wealth). So, an increase in the consumption/wealth ratio would signal high expected returns. Lettau and Ludvigson also find that the variation in returns that is picked up by the Fama-French three-factor model appears to be related to the changing risk premium from the CCAPM.

Lettau and Ludvigson's results bolster the risk-based explanation of the size and value effects. SMB and HML capture common variation in returns because they seem to be related to variation in a consumption-based risk premium that changes over time.[8] Financial theory (the CCAPM) and empirical observation (the size and value effects) are linked in an intuitively appealing way. It would be ironic if the asset pricing model that received the least empirical support in the early years turned out to be the best description of expected returns.

Pastor and Stambaugh (2001) provide evidence that sensitivity to market-wide shifts in liquidity might be a priced risk factor. Stocks that are highly sensitive to shifts in market liquidity (they have a high "liquidity beta") have high average returns. This liquidity factor appears to be distinct from SMB and HML, suggesting an independent source of risk. However, it appears that liquidity betas are highly unstable, and there is substantial variation in the corresponding premium. While it is too early to conclude that there is a systematic liquidity factor in stock returns, more research is sure to be forthcoming in this area.

Finally, an indication of the acceptance of the three-factor model is the frequency with which it is now used as a benchmark for performance measurement. For example, Quigley and Sinquefield (2000) use a three-factor benchmark to analyze the performance of UK unit trusts, and Carhart (1997) and Davis (2001) use the Fama-French model in studies of US mutual fund performance (although Carhart adds a fourth factor to reflect momentum).

IX. Conclusions

The issue of whether the value and size premiums are caused by risk or inefficiency may never be resolved to everyone's satisfaction. Feelings run strong on both sides of the argument. For investors, there are two crucial points to remember. First, factors based on value and size have explained much of the common variation in U.S. stock returns for the past three-quarters of a century. Second, value and size premiums have been observed in several other countries, with the value premium being observed in nearly every country that has been studied. While these observations are consistent with a risk-based story, they do not prove anything. Nevertheless, something very fundamental would have to change in the financial markets in order for these premiums to disappear. Furthermore, the returns observed in the U.S. market during 1999 show that “value-minus-growth” is not a low-risk strategy.

The inability of the Fama-French three-factor model to explain stock price momentum is a problem for the model’s proponents. However, the problem may not be all that serious. Consider the following facts:

1. Pure momentum strategies involve very high turnover. Consequently, transactions costs and taxes can significantly erode momentum profits.
2. Most of the return to the “winner-minus-loser” momentum portfolio is due to the poor performance of the losers. So, in order to capture the bulk of the momentum effect, short positions are necessary. This is not feasible for some investors.
3. The momentum effect is stronger among small-cap stocks, which tend to be less liquid. Trying to implement a high-turnover strategy with small-cap stocks is unrealistic.

These facts suggest that momentum strategies probably do not represent a real opportunity for investors to earn abnormal returns, at least not to the extent implied by recent studies.

The helpful comments of Robert Dintzner, Ken French, Kate Hudson, Graham Lennon, and Weston Wellington are gratefully acknowledged.

1 *Readers who are already familiar with the theory of asset pricing may skip Section II. However, the remaining sections assume that the reader is familiar with the main implications of the various asset pricing models.*
2 *For example, all investors are assumed to have the same information, and this information is costless to gather and process. In addition, there are no taxes, transactions costs, or other “frictions”. It is also assumed that investors can readily borrow funds at the risk free rate of interest.*
3 *An arbitrage opportunity is an investment strategy that has the following properties: 1) the strategy’s cost is zero; 2) the probability of a negative payoff is equal to zero; and 3) the probability of a positive payoff is greater than zero. In other words, the costless strategy can’t lose, and it might win.*
4 *The beta coefficient in the CCAPM equation is actually the ratio of the consumption betas for asset j and the market portfolio. The main point is that high positive covariance with aggregate consumption requires a higher expected return.*
5 *CRSP is the Center for Research in Security Prices at the University of Chicago. The CRSP database contains stock market data (prices, returns, shares outstanding, etc.) for NYSE, AMEX, and NASDAQ common stocks. The Compustat database is produced by Standard & Poor’s Corporation and contains accounting data for U.S. (and some foreign) companies.*
6 *Some people have the mistaken belief that the CRSP database suffers from a survivorship bias. CRSP is free of survivorship bias; Compustat is not.*
7 *Davis, Fama and French (2000) would later provide additional evidence using a much larger database over a longer sample period.*
8 *It is also important to note that SMB and HML seem to be slightly more precise than the consumption-wealth ratio in describing the time-varying risk premium. So, empirical estimation of expected returns can still be done more precisely with the Fama-French model. The CCAPM does not replace the 3-factor model; it simply provides a theoretical justification for its use*

References

Banz, Rolf W., “The Relationship Between Return and Market Value of Common Stocks,” Journal of Financial Economics *9 (1981): 3-18.*

Banz, Rolf W. and William J. Breen, “Sample-Dependent Results Using Accounting and Market Data: Some Evidence,” Journal of Finance *41 (1986): 779-793.*

Barber, Brad M., and John D. Lyon, “Firm Size, Book-to-Market Ratio, and Security Returns: A Holdout Sample of Financial Firms,” Journal of Finance *52 (1997): 875-883.*

Basu, Sanjoy, “Investment Performance of Common Stocks in Relation to Their Price-Earnings Ratios: A Test of the Efficient Market Hypothesis,” Journal of Finance *32 (1977): 663-682.*

Basu, Sanjoy, “The Relationship between Earnings Yield, Market Value and Return for NYSE Common Stocks: Further Evidence,” Journal of Financial Economics *12 (1983): 129-156.*

Bhandari, Laxmi Chand, “Debt/Equity Ratio and Expected Common Stock Returns: Empirical Evidence,” Journal of Finance *43 (1988): 507-528.*

Black, Fischer, "Capital Market Equilibrium with Restricted Borrowing," Journal of Business *45 (1972): 444-455.*

Black, Fischer, "Beta and Return," Journal of Portfolio Management *20 (1993a): 8-18.*

Black, Fischer, "Estimating Expected Return," Financial Analysts Journal *49 (1993b): 36-38.*

Breeden, Douglas, "An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities," Journal of Financial Economics *7 (1979): 265-296.*

Breeden, Douglas, Michael Gibbons, and Robert Litzenberger, "Empirical Tests of the Consumption-Oriented CAPM," Journal of Finance *44 (1989): 231-262.*

Breen, William J., and Robert A. Korajczyk, "On Selection Biases in Book-to-Market Based Tests of Asset Pricing Models," working paper #167, Northwestern University (1994).

Capaul, Carlo, Ian Rowley and William F. Sharpe, "International Value and Growth Stock Returns," Financial Analysts Journal *49 (1993): 27-36.*

Carhart, Mark M., "On Persistence in Mutual Fund Performance," Journal of Finance *52 (1997): 57-82.*

Chan, Louis K.C., Narasimhan Jegadeesh, and Josef Lakonishok, "Evaluating the Performance of Value Versus Glamour Stocks: The Impact of Selection Bias," Journal of Financial Economics *38 (1995): 269-296.*

Chan, Louis K.C., Yasushi Hamao, and Josef Lakonishok, "Fundamentals and Stock Returns in Japan," Journal of Finance 46 (1991): 1739-1789.

Chopra, Navin, Josef Lakonishok, and Jay R. Ritter, "Measuring Abnormal Performance: Do Stocks Overreact," Journal of Financial Economics *31 (1992): 235-268.*

Daniel, Kent, and Sheridan Titman, "Evidence on the Characteristics of Cross-Sectional Variation in Stock Returns," Journal of Finance *52 (1997): 1-33.*

Davis, James L., "The Cross-Section of Realized Stock Returns: The Pre-COMPUSTAT Evidence," Journal of Finance *49 (1994): 1579-1593.*

Davis, James L., "Mutual Fund Performance and Manager Style," Financial Analysts Journal *57 (2001): 19-27.*

Davis, James L., Eugene F. Fama, and Kenneth R. French, "Characteristics, Covariances, and Average Returns: 1929 to 1997," Journal of Finance *55 (2000): 389-406.*

DeBondt, Werner F.M., and Richard H. Thaler, "Does the Stock Market Overreact," Journal of Finance *40 (1985): 557-581.*

Elton, Edwin J., Martin J. Gruber, Deepak Agrawal, and Christopher Mann, "Explaining the Rate Spread on Corporate Bonds," Journal of Finance *56 (2001): 247-277.*

Fama, Eugene F., and Kenneth R. French, "The Cross-Section of Expected Stock Returns," Journal of Finance *47 (1992): 427-465.*

Fama, Eugene F., and Kenneth R. French, "Common Risk Factors in the Returns on Stocks and Bonds," Journal of Financial Economics *33 (1993): 3-56.*

Fama, Eugene F., and Kenneth R. French, "Size and Book-to-Market Factors in Earnings and Returns," Journal of Finance *50 (1995): 131-155.*

Fama, Eugene F., and Kenneth R. French, "Multifactor Explanations of Asset Pricing Anomalies," Journal of Finance *51 (1996): 55-84.*

Fama, Eugene F., and Kenneth R. French, "Value Versus Growth: The International Evidence," Journal of Finance *53 (1998): 1975-1999.*

Fama, Eugene F., and James D. MacBeth, "Risk, Return and Equilibrium: Empirical Tests," Journal of Political Economy *81 (1973): 607-636.*

Jaffe, Jeffrey, Donald B. Keim, and Randolph Westerfield, "Earnings Yields, Market Values, and Stock Returns," Journal of Finance *44 (1989): 135-148.*

Jegadeesh, Narasimhan, "Evidence of Predictable Behavior of Security Returns," Journal of Finance *45 (1990): 881-898.*

Jegadeesh, Narasimhan, and Sheridan Titman, "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency," Journal of Finance *48 (1993): 65-91.*

Kothari, S.P., Jay Shanken, and Richard G. Sloan, "Another Look at the Cross-Section of Expected Returns," Journal of Finance *50 (1995): 185-224.*

Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny, "Contrarian Investment, Extrapolation, and Risk," Journal of Finance *49 (1994): 1541-1578.*

Lettau, Martin, and Sydney Ludvigson, "Resurrecting the (C)CAPM: A Cross-sectional Test when Risk Premia are Time-Varying," Journal of Political Economy *109 (2001): 1238-1287.*

Levhari, David, and Haim Levy, "The Capital Asset Pricing Model and the Investment Horizon," Review of Economics and Statistics *59 (1977): 92-104.*

Liew, Jimmy, and Maria Vassalou, "Can Book-to-Market, Size and Momentum be Risk Factors that Predict Economic Growth," Journal of Financial Economics *57 (2000): 221-245.*

Lintner, John, "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets," Review of Economics and Statistics *47 (1965): 13-37.*

MacKinlay, A. Craig, "Multifactor Models Do Not Explain Deviations from the CAPM," Journal of Financial Economics *38 (1995): 3-28.*

Markowitz, Harry, "Portfolio Selection," Journal of Finance *7 (1952): 77-91.*

Markowitz, Harry. Portfolio Selection: Efficient Diversification of Investment. *New York: John Wiley and Sons, 1959.*

Merton, Robert C., "An Intertemporal Capital Asset Pricing Model," Econometrica *41 (1973): 867-887.*

Mossin, Jan, "Equilibrium in a Capital Asset Market," Econometrica *34 (1966): 768-783.*

Nicholson, S. Francis, "Price/Earnings Ratios," Financial Analysts Journal *16 (1960): 43-45.*

Pastor, Lubos, and Robert F. Stambaugh, "Liquidity Risk and Expected Stock Returns," working paper, University of Chicago and the Wharton School, University of Pennsylvania (2001).

Perez-Quiros, Gabriel, and Allan Timmermann, "Firm Size and Cyclical Variations in Stock Returns," Journal of Finance *55 (2000): 1229-1262.*

Quigley, Garrett, and Rex A. Sinquefield, "Performance of UK Equity Unit Trusts," Journal of Asset Management *1 (2000): 72-92.*

Reinganum, Marc R., "A New Empirical Perspective on the CAPM," Journal of Financial and Quantitative Analysis *16 (1981): 439-462.*

Rosenberg, Barr, Kenneth Reid, and Ronald Lanstein, "Persuasive Evidence of Market Inefficiency," Journal of Portfolio Management *11 (1985): 9-17.*

Ross, Stephen A., "The Arbitrage Theory of Capital Asset Pricing," Journal of Economic Theory *13 (1976a): 341-360.*

Ross, Stephen A., "Risk, Return, and Arbitrage." In Risk and Return in Finance*, edited by Irwin Friend and James Bicksler, 189-218. Cambridge, Mass.: Ballinger, (1976b).*

Sharpe, William F., "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk," Journal of Finance *19 (1964): 425-442.*