

First Draft: February 1997
This draft: June 1997
Not for Quotation: Comments Welcome

Market Efficiency, Long-Term Returns, and Behavioral Finance

Eugene F. Fama*

Abstract

Market efficiency survives the challenge from the literature on long-term return anomalies. Consistent with the market efficiency hypothesis that the anomalies are chance results, apparent over-reaction to information is about as common as under-reaction, and post-event continuation of pre-event abnormal returns is about as frequent as post-event reversal. Most important, consistent with the market efficiency prediction that apparent anomalies can be due to methodology, most long-term return anomalies tend to disappear with reasonable changes in technique.

*Graduate School of Business, University of Chicago, 1101 East 58th St., Chicago, IL 60637. E-Mail eugene.fama@gsb.uchicago.edu. The comments of Kenneth French, David Hirshleifer, S.P. Kothari, Owen Lamont, Mark Mitchell, Hersh Shefrin, Rex Sinquefeld, Richard Thaler, Theo Vermaelen, Robert Vishny, and two referees have been helpful.

Event studies, introduced by Fama, Fisher, Jensen, and Roll (FFJR 1969), produce useful evidence on how stock prices respond to information. Many studies focus on returns in a short window (a few days) around a cleanly dated event. An advantage of this approach is that because daily expected returns are close to zero, the model for expected returns does not have a big effect on inferences about abnormal returns.

The assumption in studies that focus on short return windows is that any lag in the response of prices to an event is short-lived. There is a developing literature that challenges this assumption, arguing instead that stock prices adjust slowly to information, so one must examine returns over long horizons to get a full view of market inefficiency.

If one accepts their stated conclusions, many of the recent studies on long-term returns suggest market inefficiency, specifically, long-term under-reaction or over-reaction to information. It is time, however, to ask whether this literature, viewed as a whole, suggests that efficiency should be discarded. My answer is a solid no, for two reasons.

First, an efficient market generates categories of events that individually suggest that prices over-react to information. But in an efficient market, apparent under-reaction will be about as frequent as over-reaction. If anomalies split randomly between under- and over-reaction, they are consistent with market efficiency. We shall see that a roughly even split between apparent over-reaction and under-reaction is a good description of the menu of existing anomalies.

Second, and more important, if the long-term return anomalies are so large they cannot be attributed to chance, then an even split between over- and under-reaction is a pyrrhic victory for market efficiency. We shall find, however, that the long-term return anomalies are sensitive to methodology. They tend to become marginal or disappear when exposed to different models for expected (normal) returns or when different statistical approaches are used to measure them. Thus, even viewed one-by-one, most long-term return anomalies can reasonably be attributed to chance.

A problem in developing an overall perspective on long-term return studies is that they rarely test a specific alternative to market efficiency. Instead, the alternative hypothesis is vague, market inefficiency. This is unacceptable. Like all models, market efficiency (the hypothesis that prices fully reflect available information) is a faulty description of price formation. But following the

standard scientific rule, market efficiency can only be replaced by a better specific model of price formation, itself potentially rejectable by empirical tests.

Any alternative model has a daunting task. It must specify biases in information processing that cause the same investors to under-react to some types of events and over-react to others. The alternative must also explain the range of observed results better than the simple market efficiency story; that is, the expected value of abnormal returns is zero, but chance generates deviations from zero (anomalies) in both directions.

Since the anomalies literature has not settled on a specific alternative to market efficiency, to get the ball rolling, I assume reasonable alternatives must choose between over-reaction or under-reaction. Using this perspective, section I reviews existing studies, without questioning their inferences. My conclusion is that, viewed as a whole, the long-term return literature does not identify over-reaction or under-reaction as the dominant phenomenon. But the random split predicted by market efficiency holds up rather well.

Two recent papers, Barberis, Shleifer, and Vishny (1996) and Daniel, Hirshleifer, and Subramanyam (1997), present behavioral models that accommodate over-reaction and under-reaction. To their credit, these models present rejectable hypotheses. Section II argues that, not surprisingly, the two behavioral models work well on the anomalies they are designed to explain. Other anomalies are, however, embarrassing. The problem is that both models predict post-event return reversals in response to long-term pre-event abnormal returns. In fact, post-event return continuation is about as frequent as reversal -- a result that is more consistent with market efficiency than with the two behavioral models.

Section III examines the problems in drawing inferences about long-term returns. Foremost is an unavoidable bad-model problem. Market efficiency must be tested jointly with a model for expected (normal) returns, and all models show problems describing average returns. The bad-model problem is ubiquitous, but it is more serious in long-term returns. The reason is that bad-model errors in expected returns grow faster with the return horizon than the volatility of returns. Section III also argues that theoretical and statistical considerations alike suggest that formal inferences about long-term returns should be based on averages or sums of short-term abnormal returns (AARs or CARs) rather than the currently popular buy-and-hold abnormal returns (BHARs).

In categorizing studies on long-term returns, sections I and II do not question their inferences. Dissection of individual studies takes place in section IV. The bottom line is that the evidence against market efficiency from the long-term return studies is fragile. Reasonable changes in the approach used to measure abnormal returns typically suggest that apparent anomalies are methodological illusions.

I. Over-Reaction and Under-Reaction: An Overview

One of the first papers on long-term return anomalies is DeBondt and Thaler (1985). They find that when stocks are ranked on 3- to 5-year past returns, past winners tend to be future losers, and vice versa. They attribute these long-term return reversals to investor over-reaction. In forming expectations, investors give too much weight to the past performance of firms and too little to the fact that performance tends to mean-revert. DeBondt and Thaler seem to argue that over-reaction to past information is a general prediction of the behavioral decision theory of Kahneman and Tversky (1982). Thus, one could take over-reaction to be the prediction of a behavioral finance alternative to market efficiency. For the most part, however, the anomalies literature has not accepted the discipline of an alternative hypothesis.

An exception is Lakonishok, Shleifer, and Vishny (1994). They argue that ratios involving stock prices proxy for past performance. Firms with high ratios of earnings to price (E/P), cashflow to price (C/P), and book-to-market equity (BE/ME) tend to have poor past earnings growth, and firms with low E/P, C/P, and BE/ME tend to have strong past earnings growth. Because the market over-reacts to past growth, it is surprised when earnings growth mean reverts. As a result, high E/P, C/P, and BE/ME stocks (poor past performers) have high future returns, and low E/P, C/P, and BE/ME stocks (strong past performers) have low future returns.

I also classify the poor long-term post-event returns of initial public offerings (IPOs) [Ritter (1991), Loughran and Ritter (1995)] and seasoned equity offerings (SEOs) [Loughran and Ritter (1995), Spiess and Affleck-Graves (1995)] in the over-reaction camp. Mitchell and Stafford (1997) show that SEOs have strong stock returns in the three years prior to the issue. It seems safe to presume that these strong returns reflect strong earnings. It also seems safe to presume that IPOs have strong past earnings to display when going public. If the market does not understand that earnings growth tends to mean revert, stock prices at the time of the equity issue (IPO or SEO) are too

high. If the market only gradually recognizes its mistakes, the over-reaction to past earnings growth is corrected slowly in the future. Finally, Dharan and Ikenberry (1995) argue that the long-term negative post-listing abnormal stock returns of firms that newly list on the NYSE or AMEX, are due to over-reaction. Firms list their stocks to take advantage of the market's over-reaction to their recent strong performance.

If apparent over-reaction was the general result in studies of long-term returns, market efficiency would be dead, replaced by the behavioral alternative of DeBondt and Thaler (1985). In fact, apparent under-reaction is about as frequent. The grand-daddy of under-reaction events is the evidence that stock prices seem to respond to earnings for about a year after they are announced [Ball and Brown (1968), Bernard and Thomas (1990)]. More recent is the momentum effect identified by Jegadeesh and Titman (1993); stocks with high returns over the past year tend to have high returns over the following three to six months.

Other recent event studies also produce long-term post-event abnormal returns that suggest under-reaction. Cusatis, Miles, and Woolridge (1993) find positive post-event abnormal returns for divesting firms and the firms they divest. They attribute the result to market under-reaction to an enhanced probability that, after a spinoff, both the parent and the spinoff are likely to become mergers targets, and the recipients of premiums. Desai and Jain (1997) and Ikenberry, Rankine, and Stice (1996) find that firms that split their stock experience long-term positive abnormal returns both before and after the split. They attribute the post-split returns to market under-reaction to the positive information signaled by a split. Lakonishok and Vermaelen (1990) find positive long-term post-event abnormal returns when firms tender for their stock. Ikenberry, Lakonishok, and Vermaelen (1995) observe similar results for open-market share repurchases. The story in both cases is that the market under-reacts to the positive signal in share repurchases about future performance. Finally, Michaely, Thaler, and Womack (1995) find that stock prices seem to under-react to the negative information in dividend omissions and the positive information in initiations.

Some long-term return anomalies are difficult to classify. For example, Asquith (1983) and Agrawal, Jaffe, and Mandelker (1992) find negative long-term abnormal returns to acquiring firms following mergers. This might be attributed to market under-reaction to a poor investment decision [Roll (1986)] or over-reaction to the typically strong performance of acquiring firms in advance of

mergers [documented in Mitchell and Stafford (1997)]. Ikenberry and Lakonishok (1993) find negative post-event abnormal returns for firms involved in proxy contests. One story is that stock prices under-react to the poor performance of these firms before the proxy contest, but another is that prices over-react to the information in a proxy that something is likely to change.

Given the ambiguities in classifying some anomalies, and given that the review above is surely incomplete, I shall not do a count of under-reaction versus over-reaction studies. The important point is that the literature does not lean cleanly toward either as the behavioral alternative to market efficiency. This is not lost on behavioral finance researchers who acknowledge the issue:

“We hope future research will help us understand why the market appears to overreact in some circumstances and underreact in others.” Michaely, Thaler, and Womack (1995, p.606).

The market efficiency hypothesis offers a simple answer to this question -- chance. Specifically, the expected value of abnormal returns is zero, but chance generates apparent anomalies that split randomly between over-reaction and under-reaction.

Is the weight of the evidence on long-term return anomalies so overwhelming that market efficiency is not a viable working model even in the absence of an alternative that explains both under- and over-reaction? My answer to this question is no, for three reasons.

First, I doubt that the literature presents a random sample of events. Splashy results get more attention, and this creates an incentive to find them. That dredging for anomalies is a rewarding occupation is suggested by the fact that the anomalies literature shows so little sensitivity to the alternative hypothesis problem. The same authors, viewing different events, are often content with over-reaction or under-reaction, and willing to infer that both warrant rejecting market efficiency.

Second, some apparent anomalies may be generated by rational asset pricing. Fama and French (1996) find that the long-term return reversals of DeBondt and Thaler (1985) and the contrarian returns of Lakonishok, Shleifer, and Vishny (1994) are captured by a multifactor asset pricing model. In a nutshell, return covariation among long-term losers seems to be associated with a risk premium that can explain why they have higher future average returns than long-term winners. Fama and French (1996) discuss the quarrels with their multifactor model, but their results suffice to illustrate an important point: Inferences about market efficiency can be sensitive to the assumed model for expected returns.

Finally, but most important, a roughly even split between over-reaction and under-reaction would not be much support for market efficiency if the long-term return anomalies are so large they cannot possibly be attributed to chance. Section IV argues, however, that even viewed individually, most anomalies are shaky. They tend to disappear when reasonable alternative approaches are used to measure them.

II. Behavioral Models of Under-Reaction and Over-Reaction

Before examining individual long-term return studies, I first consider two behavioral models, recently proposed by Barberis, Shleifer, and Vishny (BSV 1996), and Daniel, Hirshleifer, and Subramanyam (DHS 1997), to explain how the judgement biases of investors can produce over-reaction to some events and under-reaction to others.

The BSV model is motivated by evidence from cognitive psychology of two judgment biases. (i) The representativeness bias of Kahneman and Tversky (1982): People give too much weight to recent patterns in the data and too little to the properties of the population that generates the data. (ii) Conservatism, attributed to Edwards (1968): The slow updating of models in the face of new evidence.

In the model of stock prices proposed by BSV to capture the two judgment biases, earnings are a random walk, but investors falsely perceive that there are two earnings regimes. In regime A, which investors assume is more likely, earnings are mean-reverting. When investors decide regime A holds, a stock's price under-reacts to a change in earnings because investors mistakenly think the change is likely to be temporary. When this expectation is not confirmed by later earnings, stock prices show a delayed response to earlier earnings. In regime B, which investors think is less likely, a run of earnings changes of the same sign leads investors to perceive that a firm's earnings are trending. Once investors are convinced that the trending regime B holds, they incorrectly extrapolate the trend and the stock price over-reacts. Because earnings are a random walk, the over-reaction is exposed by future earnings, leading to reversal of long-term returns.

Regime A in the BSV model is motivated by the evidence of short-term momentum in stock returns [Jegadeesh and Titman (1993)] and the evidence of delayed short-term responses of stock prices to earnings announcements [Ball and Brown (1968), Bernard and Thomas (1990)]. Regime B is meant to explain the long-term return reversals of DeBondt and Thaler (1985), and the returns to

the contrarian investment strategies of Lakonishok, Shleifer, and Vishny (1994). How does the model do on other anomalies?

The prediction of regime B is reversal of long-term abnormal returns. Specifically, persistent long-term pre-event returns are evidence of market over-reaction which should eventually be corrected in post-event returns. In addition to DeBondt and Thaler (1985) and Lakonishok, Shleifer and Vishny (1994), other events consistent with this prediction are seasoned equity offerings [Loughran and Ritter (1995), Mitchell and Stafford (1997)], new exchange listings [Dharan and Ikenberry (1995)], and returns to acquiring firms in mergers [Asquith (1983)]. All these events are characterized by positive long-term abnormal returns before the event, and negative abnormal returns thereafter.

But long-term return reversal is not the norm. Events characterized by long-term post-event abnormal returns of the same sign as long-term pre-event returns include dividend initiations and omissions [Michaely, Thaler, and Womack (1995)], stock splits [Ikenberry, Rankine, and Stice (1996), Desai and Jain (1997)], proxy contests [Ikenberry and Lakonishok (1993)], and spinoffs [Miles and Rosenfeld (1983), Cusatis, Miles, and Woolridge (1993)].

In short, not surprisingly, the BSV model does well on the anomalies it was designed to explain. But its prediction of long-term return reversal does not capture the range of long-term results observed in the literature. Viewed as whole, the long-term return literature seems more consistent with the market efficiency prediction that long-term return continuation and long-term return reversal are equally likely chance results.

The DHS model has different behavioral foundations than the BSV model. In DHS there are informed and uninformed investors. The uninformed are not subject to judgment biases. But stock prices are determined by the informed investors, and they are subject to two biases, overconfidence and biased self-attribution. Overconfidence leads them to exaggerate the precision of their private signals about a stock's value. Biased self-attribution causes them to downweight public signals about value, especially when the public signals contradict their private signals. Over-reaction to private information and under-reaction to public information tends to produce short-term continuation of stock returns, but long-term reversals as public information eventually overwhelms the behavioral biases. Thus, though based on different behavioral premises, the DHS predictions are close to those

of BSV, and the DHS model shares the empirical successes and failures of the BSV model. This last comment also applies to Hong and Stein (1997).

DHS make a special prediction about what they call selective events. These are events that occur to take advantage of the mispricing of a firm's stock. For example, managers announce a new stock issue when a firm's stock price is too high, or they repurchase shares when the stock price is too low. This public signal produces an immediate price reaction that absorbs some of the mis-pricing. But in the DHS model, the announcement period price response is incomplete because informed investors overweight their prior beliefs about the stock's value. (The conservatism bias of the BSV model would produce a similar result.) Eventually, the mispricing is fully absorbed as further public information confirms the information implied by the event announcement. The general prediction for selective events is thus momentum; stock returns after an event announcement will tend to have the same sign as the announcement period return.

Does the DHS prediction about selective events stand up to the data? Table 1 summarizes the signs of short-term announcement returns and long-term post-announcement returns for the major long-term return studies. Except for earnings announcements, all these events seem selective. As predicted by DHS, announcement and post-announcement returns have the same sign for SEOs, dividend initiations and omissions, share repurchases, stock splits, and spinoffs. But announcement and post-announcement returns have opposite signs for new exchange listings and proxy fights, and the negative post-event returns to acquiring firms in mergers are not preceded by negative announcement returns. Most embarrassing for the DHS prediction, the long-term negative post-event returns of IPOs (the premier long-term return anomaly) are preceded by positive returns for a few months following the event [Ibbotson (1975), Ritter (1991)].

Finally, given the demonstrated ingenuity of the theory branch of finance, and given the long litany of apparent judgment biases unearthed by cognitive psychologists [DeBondt and Thaler (1995)], it is safe to predict that we will soon see a menu of behavioral models that can be mixed and matched to explain specific anomalies. My view is that any new model should be judged (as above) on how it explains the big picture. The question should be: Does the new model produce rejectable predictions that capture the menu of anomalies better than market efficiency? For existing behavioral models, my answer to this question (perhaps predictably) is an emphatic no.

The main task that remains is to examine the long-term return anomalies one at a time to see if they deliver on their claims. We set the stage with a discussion of some of the general problems that arise in tests on long-term returns.

III. Drawing Inferences from Long-Term Returns

Fama (1970) emphasizes that market efficiency must be tested jointly with a model for expected (normal) returns. The problem is that all models for expected returns are incomplete descriptions of the systematic patterns in average returns during any sample period. As a result, tests of efficiency are always contaminated by a bad-model problem.

The bad-model problem is less serious in event studies that focus on short return windows (a few days) since daily expected returns are close to zero and so have little effect on estimates of unexpected (abnormal) returns. But the problem grows with the return horizon. A bad-model problem that produces a spurious abnormal average return of $x\%$ per month eventually becomes statistically reliable in cumulative monthly abnormal returns (CARs). The reason is that the mean of the CAR increases like N , the number of months summed, but the standard error of the CAR increases like $N^{1/2}$. In AARs (averages of monthly abnormal returns), the pricing error is constant at $x\%$, but the standard error of the AAR decreases like $N^{-1/2}$. Bad model problems are most acute in long-term buy-and-hold returns since BHARs compound (multiply) an expected-return model's problems in explaining short-term returns.

This section discusses various approaches that attempt to limit bad-model problems. It also discusses a related issue, the relevant return metric in tests on long-term returns. I argue that theoretical and statistical considerations alike suggest that CARs (or AARs) should be used, rather than BHARs.

A. Bad-Model Problems

Bad-model problems are of two types. (i) Any asset pricing model is just a model and so does not completely describe expected returns. For example, the CAPM of Sharpe (1964) and Lintner (1965) does not seem to describe expected returns on small stocks [Banz (1981)]. If an event sample is tilted toward small stocks, risk adjustment with the CAPM can produce spurious abnormal returns. (ii) Even if there were a true model, any sample period produces systematic

deviations from the model's predictions, that is, sample-specific patterns in average returns that are due to chance. If an event sample is tilted toward sample-specific patterns in average returns, a spurious anomaly can arise even with risk adjustment using the true asset pricing model.

One approach to limiting bad-model problems bypasses formal asset pricing models by using firm-specific models for expected returns. For example, the FFJR (1969) stock split study uses the market model to measure abnormal returns. The intercept and slope from the regression of a stock's return on the market return, estimated outside the event period, are used to estimate the stock's expected returns, conditional on market returns, during the event period. Masulis' (1980) comparison period approach uses a stock's average return outside the event period as the estimate of its expected return during the event period.

Unlike formal asset pricing models, the market model and the comparison period approach produce firm-specific expected return estimates; that is, a stock's expected return is estimated without constraining the cross-section of expected returns. Thus, these approaches can be used to study the reaction of stock prices to firm-specific events (splits, earnings, etc.). But they cannot identify anomalies in the cross-section of average returns, like the size effect of Banz (1981), since such anomalies must be measured relative to predictions about the cross-section of average returns.

The hypothesis in studies that focus on long-term returns is that the adjustment of stock prices to an event may be spread over a long post-event period. For many events, long periods of unusual pre-event returns are common. Thus, the choice of a normal period to estimate a stock's expected return or its market model parameters is problematic. Perhaps because of this problem, event studies often control for expected returns with approaches that constrain the cross-section of expected returns. An advantage of these approaches is that they do not require out-of-sample parameter estimates. A disadvantage is that constraints on the cross-section of expected returns always produce imperfect descriptions of average returns, and so can lead to bad-model problems.

For example, one approach estimates an abnormal return as the difference between an event firm's return and the return on a non-event firm or portfolio that is similar on characteristics known to be related to average returns. The hope in this matching approach is to control for cross-firm variation in average returns due both to differences in expected returns and to chance sample-specific patterns in average returns. For example, following Banz' (1981) evidence that small stocks have

higher average returns than predicted by the CAPM, an event stock's abnormal return is often estimated as the difference between its return and the return on non-event stocks matched to the event stock on size. Following the evidence of Fama and French (1992) that average stock returns are also related to book-to-market equity (BE/ME), it is now common to estimate abnormal returns by matching event stocks with non-event stocks similar in terms of size and BE/ME.

When we analyze individual event studies, we shall see that matching on size can produce much different abnormal returns than matching on size and BE/ME. And size and BE/ME surely do not capture all relevant cross-firm variation in average returns due to expected returns or sample-specific patterns in average returns. In short, the matching approach is not a panacea for bad-model problems in studies of long-term abnormal returns.

Another method of estimating abnormal returns is to use an asset pricing model to estimate expected returns. Early studies of long-term abnormal returns [Jaffe (1974), Mandelker (1974), Asquith (1983)], use the CAPM. Some recent studies use the three-factor model of Fama and French (FF 1993). Like all asset pricing models, however, the CAPM and the FF model are incomplete descriptions of average returns. The shortcomings of the CAPM are well-known [FF (1992)]. And FF (1993) show that their three-factor model does not even provide a full explanation of average returns on portfolios formed on size and BE/ME, the dimensions of average returns that the model's risk factors are designed to capture.

In short, bad-model problems are unavoidable, and they are more serious in tests on long-term returns. When we review individual studies in section IV, the tracks of the bad-model problem will be clear. Different models for expected returns produce different estimates of long-term abnormal returns. And a reasonable change of models often causes an anomaly to disappear. I argue that when this happens, the anomaly is not much evidence against market efficiency.

B. The Return Metric

Studies of long-term returns are also sensitive to the way the tests are done. Average monthly abnormal returns (AARs or CARs) can produce different inferences than buy-and-hold abnormal returns (BHARs). Equal-weight returns produce different results than value-weight returns. And failure to account for the cross-correlation of event firm returns during long post-event periods can affect inferences. These implementation issues are discussed next.

The Return Metric: Theoretical Issues -- In principle, the model of market equilibrium jointly tested with market efficiency specifies the unit of time for returns. For example, if the model defines equilibrium in terms of monthly expected returns, average monthly returns should be the metric used to test market efficiency. To examine how prices respond over periods longer than a month, one can average (AARs) or sum (CARs) the average monthly abnormal returns. Beginning with FFJR (1969), AARs and CARs are a common approach to examining long-term returns.

A criticism of this approach is that an average monthly return does not accurately measure the return to an investor who holds a security for a long post-event period. Long-term investor experience is better captured by compounding short-term returns to obtain long-term buy-and-hold (BH) returns. Much of the recent literature tests buy-and-hold abnormal returns (BHARs) for periods up to five years after an event.

Investor experience is interesting, and long-term BHARs are thus interesting. But formal tests for abnormal returns should use the return metric called for by the model invoked to estimate expected (normal) returns. The problem, of course, is that discrete-time asset pricing models are silent on the relevant interval for expected returns. Nevertheless, there are at least three theoretical reasons to lean toward shorter intervals.

(i) Asset pricing models, like the Sharpe (1964) - Lintner (1965) CAPM and the discrete-time version of Merton's (1973) ICAPM, commonly assume normally distributed returns. Normality is a better approximation for short horizons like a month than for longer horizons, where skewness becomes increasingly important [Fama (1976, 1996)].

(ii) The empirical tests of asset pricing models, invoked to justify applying the models in tests of market efficiency, typically use monthly returns. I know of no tests of asset pricing models on five-year returns.

(iii) Mitchell and Stafford (1997) point out that BHARs can give false impressions of the speed of price adjustment to an event. The reason is that BHARs can grow with the return horizon even when there is no abnormal return after the first period. For example, suppose returns for the first year after the event are 10% for event firms and zero for benchmark firms, so the first year abnormal return is 10%. Suppose event and benchmark firms both have a 100% BH return over the next four years. Although there is no abnormal return after the first year, the BHAR after five years grows to 20%, that is $(1.1 \times 2.0) - (1.0 \times 2.0)$.

The Return Metric: Statistical Issues -- AARs and CARs also pose fewer statistical problems than long-term BHARs. Barber and Lyon (1997) provide the most complete discussion of the inference problems in tests on long-term returns. [See also Kothari and Warner (1997).] Barber and Lyon favor BHARs, but their tests show that inferences are less problematic for average monthly returns (AARs or CARs). In a follow-up paper, Barber, Lyon, and Tsai (1996) develop elaborate

techniques for correcting some of the inference problems of BHARs. But they acknowledge that their improved methods for BHARs produce inferences no more reliable than simpler methods applied to monthly AARs or CARs. The reason is that average monthly returns avoid the problems (e.g., extreme skewness) produced by compounding monthly returns to get long-term BHARs.

Brav (1997) emphasizes that all existing methods for drawing inferences from BHARs, including those in Barber, Lyon, and Tsai (1996), fail to correct fully for the correlation of returns across events that is not absorbed by the model used to adjust for expected returns. The problem is more severe in long-term BHARs because more firms have events within, say, a given 5-year window than within a 3-day window. Brav (1997) presents an elaborate scheme to adjust for the cross-correlation of long-term BHARs in special cases (e.g., when it is due to industry effects). But a full solution is not typically available because the number of return covariances to be estimated is greater than the number of times-series observations.

In contrast, if average monthly returns are used, there has long been a full solution to the cross-correlation problem. Suppose the post-event period of interest is five years. For each calendar month, calculate the abnormal return on each stock that had an event in the last five years. (Abnormal returns can be estimated in any reasonable way, for example, with a matching firm portfolio approach, or with a formal asset pricing model.) Then average the abnormal returns for the calendar month across stocks to get the abnormal return for the month on the portfolio of stocks with an event in the last five years. Re-form the portfolio every month. The time-series variation of the monthly abnormal return on this portfolio accurately captures the effects of the correlation of returns across event stocks missed by the model for expected returns. The mean and variance of the time series of abnormal portfolio returns can be used to test the average monthly response of the prices of event stocks for five years following the event. The approach can also be refined to allow for heteroscedasticity of the portfolio's abnormal return due to changes through time in the composition of the portfolio. This rolling portfolio approach (with refinements) was first used by Jaffe (1974) and Mandelker (1974).

A referee suggests that the portfolio approach described above can cause an anomaly to be understated if events bunch in time because firms exploit pricing errors during windows of opportunity. For example, Loughran and Ritter (1995) suggest that IPOs bunch because particular

industries tend to be overvalued at specific times. This criticism of the portfolio approach is valid if monthly abnormal portfolio returns are weighted equally in calculating an overall average monthly abnormal return. But the problem disappears if the abnormal return for a month is first adjusted for heteroscedasticity, that is, if it is weighted by an estimate of the standard deviation of that month's abnormal return, and the overall abnormal return is then estimated by averaging the standardized monthly abnormal returns. More generally, one can weight the portfolio abnormal return for a month in any way that captures the economic hypothesis of interest.

The Return Metric: Value Weights versus Equal Weights -- In the review of individual studies that follows, we find that apparent anomalies in long-term post-event returns typically shrink a lot and often disappear when event firms are value-weighted rather than equal-weighted. One can argue that value-weight returns give the right perspective on an anomaly because they more accurately capture the total wealth effects experienced by investors. But I am more concerned about bad-model problems. All the common asset pricing models, including the Fama-French (1993) three-factor model, have systematic problems explaining the average returns on categories of small stocks. Since equal-weight portfolio returns give more weight to small stocks, bad-model problems are more severe in inferences from equal-weight returns.

Many readers of suggest that more serious mispricing of small stocks is a general prediction of behavioral finance. It is thus worth noting that the two behavioral pricing models reviewed in section II do not produce this prediction. In Barberis et al (1996), pricing is dominated by a representative investor, and there is no prediction that the judgment biases of this investor are more severe for small stocks. In Daniel et al (1997), pricing is dominated by informed investors subject to judgment biases. Uninformed investors have no such biases. Thus, if large stocks attract more interest from informed investors (e.g., security analysts), mispricing problems might be more severe for large stocks.

Most important, the cognitive psychology evidence does not seem to say that different classes of people are more subject to judgment biases. The same biases that plague college students (the subjects of most cognitive psychology experiments) also occur among experts (see the references in Barberis et al). Thus, cognitive psychology, the basis of behavioral finance, does not seem to provide a basis for the common presumption that small stocks are more likely to be mispriced.

IV. The Reliability of Individual Studies

The summary of long-term return studies in sections I and II accepts the conclusions of the papers at face value. Now the gloves come off. Examining long-term return anomalies one at a time, I argue that most are fragile. Abnormal returns often disappear with reasonable changes in the way they are measured.

IPOs and SEOs -- Among the more striking of the long-term return anomalies is the study of initial public offerings (IPOs) and seasoned equity offerings (SEOs) of Loughran and Ritter (1995). They find that the total wealth generated at the end of five years if one invests \$1 in each IPO or SEO immediately following the event is about 70% of that produced by the same buy-and-hold (BH) strategy applied to a sample of stocks matched to the IPOs and SEOs on size.

IPOs and SEOs clearly have poor long-term returns during the Loughran-Ritter sample period (1970-1990). The interesting question is whether the returns are really abnormal or whether they are shared with non-event firms similar on characteristics related to average returns. During the Loughran-Ritter period, variables known to be related to average stock return include size and book-to-market equity (BE/ME) [Fama and French (1992)], and short-term past return [Jegadeesh and Titman (1993)]. Since the long-term buy-and-hold returns in Loughran and Ritter only control for size, their results might be affected by other variables that are systematically related to average return.

Following up on this possibility, Brav and Gompers (1996) compare five-year BH returns on IPOs with the returns on portfolios that match the IPOs on size and BE/ME but exclude SEOs as well as IPOs. The five-year wealth relative (the ratio of five-year BH wealth for IPOs to five-year BH wealth for the benchmarks) rises from about .7 with the size benchmarks in Loughran-Ritter to a bit more than 1.0 (that is, the anomaly disappears) when the benchmarks control for BE/ME as well as size. Similarly, Brav, Geczy, and Gompers (1995) find that the five-year BH returns on SEOs are close to those of non-event portfolios matched on size and BE/ME.

Brav (1997) and Mitchell and Stafford (1997) show that IPOs and SEOs are typically small growth stocks. Fama and French (1993) show that such stocks have low returns during the post-1963 period. The results of Brav and Gompers (1996) and Brav et al (1995) then suggest that explaining the IPO-SEO anomaly reduces to explaining why small growth stocks in general have poor returns

during the IPO-SEO sample period. In other words, if there is a mispricing problem, it is not special to IPO-SEO stocks.

Brav and Gompers (1996) and Brav et al (1995) also find that when IPOs and SEOs are value-weighted, five-year abnormal BH returns shrink a lot, whatever the benchmark. For IPOs, value-weight five-year wealth relatives are .86 or greater for all benchmarks, and four of six benchmarks produce wealth relatives in excess of .9. For SEOs, value-weight five-year wealth relatives are .88 or greater for all benchmarks; three of six are in excess of .98. The message is that many IPO and SEO stocks are tiny, and they are influential in the returns observed when sample firms are equal weighted. This result is general. We shall see that apparent anomalies typically shrink a lot when viewed in terms of value-weight returns.

Loughran and Ritter (1995), Brav and Gompers (1996), and Brav et al (1995) do not engage in the treacherous game of drawing statistical inferences from long-term BH returns. Their inferences are based on average monthly returns. Every month they calculate the return on a portfolio that contains all firms with an IPO or SEO in the last five years. The three-factor model of Fama and French (FF 1993) is then used to estimate the portfolio's abnormal returns. The average monthly abnormal return during the five-year post-event period is the intercept, α_p , of the time-series regression,

$$(1) \quad R_{pt} - R_{ft} = a_p + b_p [R_{Mt} - R_{ft}] + s_p \text{SMB} + h_p \text{HML} + \varepsilon_{pt},$$

where R_{pt} is the monthly return on the IPO or SEO portfolio, R_{ft} is the one-month Treasury bill rate, R_{Mt} is the monthly return on a value-weight market portfolio of NYSE, AMEX, and NASDAQ stocks, SMB is the difference between the returns on portfolios of small and big stocks (below or above the NYSE median), and HML is the difference between the returns on portfolios of high- and low-BE/ME stocks (above and below the .7 and .3 fractiles of BE/ME).

Brav et al (1995) estimate the intercepts in (1) for equal- and value-weight portfolios of the SEOs of their 1975-92 sample

period. The intercept for the equal-weight portfolio is $-.42\%$ per month ($t=-4.8$), but the intercept for the value-weight portfolio is $-.14\%$ per month ($t=-1.18$). Similarly, Brav and Gompers (1996) find that the intercepts for equal- and value-weight portfolios of IPOs that are not backed by venture capitalists are $-.52\%$ ($t=-2.80$) and $-.29\%$ ($t=-1.84$). For IPOs backed by venture capitalists, the intercepts are slightly positive. Loughran and Ritter (1995) only show regressions that combine IPOs and SEOs, but their results are similar; when value-weight portfolios are used, abnormal returns are economically and statistically close to zero.

Since inferences about abnormal returns from estimates of (1) on rolling post-event portfolio returns are common in the recent anomalies literature, it is important to note three potential problems.

First, since the firms in the event portfolio change through time, the true slopes on the risk factors in (1) are time-varying. Mitchell and Stafford (1997) confirm that for three important events (mergers, share repurchases, and SEOs), changes in the composition of the event portfolio generate substantial variation in the slopes in (1). For SEOs, they find that the intercepts for their 1960-93 period for equal- and value-weight portfolios drop from $-.38$ ($t=4.47$) and $-.14$ ($t=-1.51$) in the constant slope regressions, to $-.24$ ($t=-3.64$) and $-.07$ ($t=-.81$) in the regressions that allow the slopes to vary through time.

Second, the number of firms in the event portfolio changes through time, creating residual heteroscedasticity that can affect inferences about the intercept. [Solutions to this problem like

those in Jaffe (1974) and Mandelker (1974) are apparently a lost technology.]

Third, but perhaps most important, FF (1993) show that the three-factor model is not a perfect story for average returns. This bad-model problem can produce spurious anomalies in event studies. For example, IPOs and SEOs tend to be small low-BE/ME firms. FF (1993) show that the three-factor model overestimates average returns on such firms during the IPO-SEO sample periods. This bad-model problem can explain why estimates of (1) on equal-weight IPO and SEO portfolios produce reliably negative intercepts, but estimates on value-weight portfolios produce intercepts close to zero. It can also explain why the intercepts in (1), which control for loadings on risk factors related to size and BE/ME, suggest abnormal post-event returns for equal-weight IPO and SEO portfolios. But with direct benchmark matching on size and BE/ME, the abnormal returns largely disappear.

I emphasize, however, that the results for IPOs and SEOs do not imply that benchmark matching on size and BE/ME is always superior to estimating abnormal returns as the intercepts from (1). All methods for estimating abnormal returns are subject to bad-model problems, and no method is likely to minimize bad-model problems for all classes of events. The important general message from the IPO-SEO results is one of caution: Two approaches that seem closely related (both attempt to control for variation in average returns related to size and BE/ME) can produce much different estimates of long-term abnormal returns.

In sum, I read Brav and Gompers (1996) and Brav et al (1995) as showing that the poor long-term BH returns following IPOs and SEOs are not a special anomaly. The low returns are shared with

other firms similar on two dimensions, size and BE/ME, known to be related to average return. Moreover, when IPOs and SEOs are value-weighted, abnormal returns shrink for all benchmarks, and they are not reliably different from zero. Thus, if there is an IPO-SEO anomaly, it seems to be largely restricted to tiny firms.

Mergers-- Asquith (1983) and Agrawal, Jaffe, and Mandelker (1992) find negative abnormal returns for acquiring firms for up to five years following merger announcements. Using a comprehensive sample of mergers for 1960-93, Mitchell and Stafford (1997) also find negative long-term abnormal returns for acquiring firms. Since these studies produce similar results, I focus on Mitchell and Stafford (MS 1997).

MS find that the three-year post-event BH return for equal-weighted acquiring firms is on average 4% lower than for portfolios matched to acquiring firms on size and BE/ME. In economic terms, this is not a dramatic anomaly. For formal inferences, MS estimate the three-factor model (1) on the monthly returns on a rolling portfolio that includes firms with acquisitions during the preceding three years. When the acquirers are equal-weighted, the intercept in (1), that is, the average monthly abnormal return for the three years after a merger, is .25% per month (-25 basis points, $t=-3.49$), which is larger than but roughly consistent with the BH returns. When acquiring firms are value-weighted, the intercept in (1) drops to a puny -.11% per month ($t=-1.55$). Thus, if there is an anomaly, it is more important for smaller acquiring firms. Finally, MS and Loughran and Vijh (1997) show that abnormal post-announcement average returns to acquiring firms are limited to mergers financed with stock, that is, mergers that are also SEOs. When mergers are

financed without issuing stock, the negative abnormal post-event returns disappear. This suggests that there is no distinct merger anomaly. Any merger anomaly may be the SEO anomaly in disguise.

Stock Splits -- Desai and Jain (1997) and Ikenberry, Rankine, and Stice (1996) find that for the 17-year 1975-91 period, stock splits are followed by positive abnormal returns of about 7% in the year after the split. Abnormal returns are calculated relative to benchmarks that control for size, BE/ME, and, in Desai and Jain, past one-year return.

One way to test whether such an anomaly is real or the sample-specific result of chance is to examine a different sample period. FFJR (1969) examine splits during the 33-year 1927-59 period. They find no drift in cumulative abnormal returns during the 30 months following splits. Since the split anomaly fails the out-of-sample test provided by FFJR, it seems reasonable to conclude that the 1975-91 anomaly is not real, unless the market has recently become inefficient.

Desai and Jain (1997) and Ikenberry et al (1996) do provide neat evidence on one of the pitfalls in using buy-and-hold abnormal returns (BHARs) to judge the long-term return drift associated with an event. As noted earlier, Mitchell and Stafford (1997) point out that BHARs are likely to grow with the return horizon even when there is no abnormal return after the first period. For the 1975-91 period, the abnormal return is about 7% for the first year following stock splits, but it is close to zero in the following two years (slightly negative in the second and slightly positive in the third). BHARs, however, rise from 7% to about 12% after three years.

One way to avoid the distorted perspective on long-term drift produced by BHARs is to examine ratios rather than differences of the cumulative wealths generated by event and benchmark firms [as in Ritter (1991) and Loughran and Ritter (1995)]. Another is the time-worn CAR approach of FFJR (1969), which sums returns rather than compounding them, or the rolling portfolio average monthly abnormal return (AAR) approach of Jaffe (1974) and Mandelker (1974).

Self-Tenders and Share Repurchases -- Lakonishok and Vermaelen (1990) examine long-term returns following self-tender offers (tenders by firms for their own shares) during the 1962-86 period. Ikenberry, Lakonishok, and Vermaelen (1995) examine long-term returns following share repurchases during the 1980-90 period. Mitchell and Stafford (1997) study both self-tenders and repurchases for the 1960-93 period. Since the MS results are similar to but more comprehensive than those in the earlier papers, the discussion focuses on them.

MS find that three-year post-event BHARs, computed relative to matching portfolios that control for size and BE/ME, are 9% for self-tenders (475 events) and 19% for the much larger sample of 2542 repurchases. When they estimate the three-factor regression (1) for the monthly returns on an equal-weight portfolio that contains all self-tenders and repurchases in the last three years, however, the average abnormal monthly return is puny, .11% per month ($t=1.62$). Any hint of significance, economic or statistical, disappears entirely when the stocks in the rolling portfolio are value-weighted. The intercept for the value-weight portfolio of self-tenders and repurchases is $-.03\%$ (-3 basis points per month, $t=-.34$). In short, viewed from the perspective

of the three factor model of (1), there is no share repurchase anomaly.

Note that, once again, two apparently similar methods for estimating abnormal returns, (i) a matching portfolio control for size and BE/ME and (ii) an asset pricing model that adjusts for sensitivity to risk factors related to size and BE/ME, produce somewhat different results. This again illustrates that estimates of long-term abnormal returns can be sensitive to apparently small changes in technique.

Exchange Listings -- Dharan and Ikenberry (1995) find that during the 1962-90 period, stocks that newly list on the NYSE, or move from NASDAQ to AMEX, have negative post-listing abnormal returns. When returns are risk-adjusted using matching portfolios formed on size and BE/ME, the three-year average abnormal return is -7.02%. The t-statistic for this CAR is -2.78, but this is without a full adjustment for the correlation of abnormal returns across firms. Moreover, Dharan and Ikenberry show that the negative post-listing abnormal returns are limited to firms below the NYSE-AMEX median in size. Thus, once again, an apparent anomaly is limited to small stocks.

Mitchell and Stafford (1997) offer concrete perspective on how significance levels can be overstated because of the failure to adjust for the correlation across firms of post-event abnormal returns. Using the three-factor model (1), they calculate the standard deviations of abnormal returns for portfolios of firms with an event during the most recent 36 months. The proportions vary somewhat through time and across their three event classes (mergers, share repurchases, and SEOs), but on average the covariances of event-firm abnormal returns account for about half

the standard deviation of the event portfolio's abnormal return. Thus, if the covariances are ignored, the standard error of the abnormal portfolio return is too small by about 50%. This estimate need not apply intact to the exchange listings of Dharan and Ikenberry (1995), but it suggests that a full adjustment for the cross-correlation of post-listing abnormal returns could cause the statistical reliability ($t=-2.78$) of their -7.02% post-event three-year CAR to disappear.

The story that Dharan and Ikenberry offer to explain their negative post-listing abnormal returns is also suspect. They argue that firms are opportunistic. They take advantage of the market's over-reaction to their recent good times by listing. But what are the benefits of listing, given that any over-reaction to past performance has already occurred and will soon be reversed? Moreover, standard signaling theory [e.g., Ross (1977)] does not predict that firms will incur costs to make a false signal whose price effects are soon obliterated. On the contrary, since listing involves costs, it should be a signal that the firm is under-valued.

Dividend Initiations and Omissions -- Michaely, Thaler, and Womack (1995) find that during the 1964-88 period, firms that initiate dividends have positive abnormal stock returns for three years after the event, and firms omitting dividends have negative abnormal returns. For the same sample, Brav (1997) finds that the three-year post-event abnormal return following initiations disappears with benchmarks that control for size and BE/ME. Michaely et al (1995) show that the negative three-year abnormal returns following omissions [confirmed by Brav (1997)] are largely concentrated in the second half of their 1964-88 sample period.

All this suggests that inferences about long-term returns following changes in dividends should probably await an out-of-sample test.

The finding that stock prices under-react to dividend announcements is suspicious on other grounds. It seems reasonable that under-reaction would occur because the market under-estimates the information in dividends about future earnings. However, from Watts (1973) to Benartzi, Michaely, and Thaler (1996), there is little evidence that changes in dividends predict changes in earnings.

Spinoffs -- Cusatis, Miles, and Woolridge (1993) study the post-event returns of spinoffs and their parents for the 1965-88 period. The benchmarks are firms matched to the event firms on size and industry, and abnormal returns are BHARs. Both parents and spinoffs have positive abnormal returns in the three years after the event. The abnormal returns are, however, limited to event firms (parents and spinoffs) acquired in mergers. The conclusion is that the market did not properly assess the increased probability of takeover (and the attendant buyout premiums) following spinoffs.

The t-statistics for the three-year BHARs for spinoffs range from .58 to 2.55, hardly overwhelming. Moreover, in calculating the t-statistics, the BHARs of the event firms are assumed to be independent. It would not take a large adjustment for cross-correlation to produce t-statistics that suggest no real anomaly.

Proxy Contests -- Ikenberry and Lakonishok (1993) examine stock returns following proxy contests during the 1968-87 period. They find negative post-event abnormal returns relative to benchmarks that control for market β and size. In the results for all proxy contests, the post-event abnormal returns are not statistically reliable. The negative post-event returns are only statistically reliable for the 50 odd proxy contests where the dissidents win board representation. Since this result is not an ex ante prediction, the weak evidence for the overall sample seems more relevant, and it does not suggest a reliable anomaly. This is more or less the conclusion of the authors.

Summary -- If a reasonable change in the method of estimating abnormal returns causes an anomaly to disappear, the anomaly is on shaky footing, and it is reasonable to suggest that it is an illusion. Included in this category are IPOs, SEOs, self tenders, share repurchases, and dividend

initiations. Moreover, the doubts about these anomalies are the result of replication and robustness checks that followed publication of the original studies. Other anomalies will likely fall prey to the same process.

Other long-term return anomalies are economically or statistically marginal. The negative post-event abnormal returns to acquiring firms in mergers are economically small. For exchange listings, spinoffs, and proxy contests, a full correction for the cross-correlation of long-term post-event abnormal returns could easily reduce them to former anomalies.

Some anomalies do not stand up to out-of-sample replication. Foremost (in my mind) is the stock split anomaly observed after 1975, which is contradicted by the earlier FFJR (1969) study. The long-term negative post-event returns of dividend omitting firms also seem sensitive to sample period.

Whenever value-weight returns are examined, apparent anomalies shrink a lot and typically become statistically unreliable. At a minimum, this suggests that anomalies are largely limited to small stocks. But a reasonable alternative explanation is that small stocks are just a sure source of bad-model problems. Small stocks always pose problems in tests of asset pricing models, so they are prime candidates for bad-model problems in tests of market efficiency on long-term returns.

Which anomalies are above suspicion? The post-earnings-announcement drift first reported by Ball and Brown (1968) has survived robustness checks, including extension to more recent data [Bernard and Thomas (1990), Chan, Jegadeesh, and Lakonishok (1996)]. Again, though, the anomaly is stronger for small stocks. The short-term continuation of returns documented by Jegadeesh and Titman (1993) is also an open puzzle, but it is still rather new and further tests are in order.

V. Conclusions

The recent finance literature seems to produce many long-term return anomalies. Subjected to scrutiny, however, the evidence does not suggest that market efficiency should be abandoned. Consistent with the market efficiency hypothesis that the anomalies are chance results, apparent over-reaction of stock prices to information is about as common as under-reaction. And post-event continuation of pre-event abnormal returns is about as frequent as post-event reversal. Most important, the long-term return anomalies are fragile. They tend to disappear with reasonable changes in the way they are measured.

References

- Agrawal, Anup, Jeffrey F. Jaffe, and Gershon N. Mandelker, 1992, The post-merger performance of acquiring firms: A re-examination of an anomaly, *Journal of Finance* 47, 1605-1621.
- Asquith, Paul, 1983, Merger bids, uncertainty and stockholder returns, *Journal of Financial Economics* 11, 51-83.
- Ball, Ray and Philip Brown, 1968, An empirical evaluation of accounting income numbers, *Journal of Accounting Research* 6, 159-178.
- Banz, Rolf W., 1981, The relationship between return and market value of common stocks, *Journal of Financial Economics* 9, 3-18.
- Barber, Brad M., and John D. Lyon, 1996, Detecting long-horizon abnormal stock returns: The empirical power and specification of test statistics, *Journal of Financial Economics* 43, 341-72.
- Barber, Brad M., John D. Lyon, and Chih-Ling Tsai, 1996, Improved methods for tests of long-run abnormal returns, manuscript, Graduate School of Management, U-C Davis, December.
- Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, 1996, A model of investor sentiment, CRSP Working Paper 443, Graduate School of Business, University of Chicago, December
- Benartzi, Shlomo, Roni Michaely, and Richard Thaler, 1996, Do Dividend Changes Signal the Future or the Past, manuscript, Graduate School of Business, University of Chicago.
- Bernard, Victor L., and Jacob K. Thomas, 1990, Evidence that stock prices do not fully reflect the implications of current earnings for future earnings, *Journal of Accounting and Economics* 13, 305-40.
- Brav, Alon, 1997, Inference in Long-Horizon Event Studies: A Re-evaluation of the Evidence, manuscript, Graduate School of Business, University of Chicago, January.
- Brav, Alon, and Paul A. Gompers, 1996, Myth or reality? The long-run underperformance of initial public offerings: Evidence from venture and nonventure capital-backed companies, manuscript, Graduate School of Business, University of Chicago, September.
- Brav, Alon, Christopher Geczy, and Paul A. Gompers, 1995, The long-run underperformance of seasoned equity offerings revisited, manuscript, Graduate School of Business, University of Chicago, November.
- Chan, Louis K.C., Narasimhan Jegadeesh, and Josef Lakonishok, 1996, Momentum strategies, *Journal of Finance* 51, 1681-1713.
- Cusatis, Patrick J., James A. Miles, and J. Randall Woolridge, 1993, Restructuring through spinoffs, *Journal of Financial Economics* 33, 293-311.
- Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, 1997, A theory of overconfidence, self-attribution, and security market under- and over-reactions, manuscript, February.
- DeBondt, Werner F. M., and Thaler, Richard H., 1985, Does the stock market overreact, *Journal of Finance* 40, 793-805.
- DeBondt, Werner F. M., and Thaler, Richard H., 1995, Financial decision-making in markets and firms: A behavioral perspective, in R. Jarrow et al., Eds., *Handbooks in OR and MS*, Vol. 9, Elsevier Science B.V., Chapter 13, 385-410.

- Desai, Hemang, and Prem C. Jain, 1997, Long-run common stock returns following splits and reverse splits, *Journal of Business* 70, 409-433.
- Dharan, Bala G., and David L. Ikenberry, 1995, The long-run negative drift of post-listing stock returns, *Journal of Finance* 50, 1547-1574.
- Edwards, Ward, 1968, Conservatism in Human Information Processing, in B. Kleinmütz (ed.), *Formal Representation of Human Judgement*, New York, John Wiley and Sons.
- Fama, Eugene F., 1976, *Foundations of Finance*, New York, Basic Books.
- Fama, Eugene F., 1970, Efficient capital markets: A review of theory and empirical work, *Journal of Finance* 25, 383-417.
- E. Fama, 1996, Discounting under Uncertainty, *Journal of Business* 69, 415-428.
- Fama, Eugene F., Lawrence Fisher, Michael Jensen, and Richard Roll, 1969, The adjustment of stock prices to new information, *International Economic Review* 10, 1-21.
- Fama, Eugene F., and Kenneth R. French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427-465.
- Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3-56.
- Fama, Eugene F., and Kenneth R. French, 1996, Multifactor explanations of asset pricing anomalies, *Journal of Finance* 51, 55-84.
- Hong, Harrison, and Jeremy C. Stein, 1997, A unified theory of underreaction, momentum trading, and overreaction in asset markets, manuscript, MIT, May
- Ibbotson, Roger, 1975, Price performance of common stock new issues, *Journal of Financial Economics* 2, 235-272.
- Ikenberry, David, and Josef Lakonishok, 1993, Corporate governance through the proxy contest: Evidence and implications, *Journal of Business* 66, 405-435
- Ikenberry, David, Josef Lakonishok, and Theo Vermaelen, 1995, Market underreaction to open market share repurchases, 1995, *Journal of Financial Economics* 39, 181-208.
- Ikenberry, David, Grame Rankine, Earl K. Stice, 1996, What do stock splits really signal?, *Journal of Financial and Quantitative Analysis* 31, 357-377.
- Jaffe, Jeffrey F., 1974, Special information and insider trading, *Journal of Business* 47, 410-428.
- Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65-91.
- Kahneman, Daniel, and Amos Tversky, 1982, Intuitive predictions: Biases and corrective procedures. Reprinted in Kahneman, Slovic, and Tversky, *Judgement under Uncertainty: Heuristics and Biases* Cambridge University Press, Cambridge, England.
- Kothari, S.P., and Jerold B. Warner, 1997, Measuring long-horizon security price performance, *Journal of Financial Economics* 43, 301-339.
- Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny, 1994, Contrarian investment, extrapolation, and risk, *Journal of Finance* 49, 1541-1578.

- Lakonishok, Josef, and Theo Vermaelen, 1990, Anomalous price behavior around repurchase tender offers, *Journal of Finance* 45, 455-477.
- Lintner, John, 1965, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, *Review of Economics and Statistics* 47, 13-37.
- Loughran, Tim, and Jay R. Ritter, 1995, The new issues puzzle, *Journal of Finance* 50, 23-51.
- Loughran, Tim, and Anand M. Vijh, 1997, Do long-term shareholders benefit from corporate acquisitions?, *Journal of Finance*, forthcoming.
- Mandelker, Gershon, 1974, Risk and return: the case of merging firms, *Journal of Financial Economics* 1, 303-36.
- Masulis, Ronald W., 1980, The effects of capital structure changes on security prices: A study of exchange offers, *Journal of Financial Economics* 8, 139-177.
- Merton, Robert C., 1973, An intertemporal capital asset pricing model, *Econometrica* 41, 867-887.
- Michaely, Roni, Richard H. Thaler, and Kent L. Womack, 1995, Price reactions to dividend initiations and omissions, *Journal of Finance* 50, 573-608.
- Miles, James A., and James D. Rosenfeld, 1983, The effect of voluntary spinoff announcements on shareholder wealth, *Journal of Finance* 38, 1597-1606.
- Mitchell, Mark L., and Erik Stafford, 1997, Managerial decisions and long-term stock price performance, manuscript, June.
- Ritter, Jay R., 1991, The long-term performance of initial public offerings, *Journal of Finance* 46, 3-27.
- Roll, Richard, 1986, The hubris hypothesis of corporate takeovers, *Journal of Business* 59, 197-216.
- Ross, Stephen, 1977, The determinants of financial structure: The incentive signaling approach, *Bell Journal of Economics* 8, 23-40.
- Sharpe, William F., 1964, Capital asset prices: a theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425-42.
- Spiess, D. Katherine, and John Affleck-Graves, 1995, Underperformance in long-run stock returns following seasoned equity offerings, *Journal of Financial Economics* 38, 243-267.
- Vermaelen, Theo, 1981, Common stock repurchases and market signaling: An empirical study, *Journal of Financial Economics* 9, 139-83.
- Watts, Ross, 1973, The Information Content of Dividends, *Journal of Business* 46, 191-211.

Table 1 -- Signs of long-term pre-event, announcement, and long-term post-event returns for various long-term return studies			
Event	Long-Term Pre-Event Return	Announcement Return	Long-Term Post-Event Return
Initial Public Offerings (IPOs) [Ibbotson (1975), Loughran and Ritter (1995)]	Not Available	+	-
Seasoned Equity Offerings [Loughran and Ritter (1995)]	+	-	-
Mergers (Acquiring Firm) [Asquith (1983) Agrawal, Jaffe, and Mandelker (1992)]	+	0	-
Dividend Initiations [Michaely, Thaler, and Womack (1995)]	+	+	+
Dividend Omissions [Michaely, Thaler, and Womack (1995)]	-	-	-
Earnings Announcements [Ball and Brown (1968), Bernard and Thomas (1990)]	Not Available	+	+
New Exchange Listing [Dharan and Ikenberry (1995)]	+	+	-
Share Repurchases (Open Market) [Ikenberry, Lakonishok, and Vermaelen (1995), Mitchell and Stafford (1997)]	0	+	+
Share Repurchases (Tenders) [Lakonishok and Vermaelen (1990), Mitchell and Stafford (1997)]	0	+	+
Proxy fights [Ikenberry and Lakonishok (1993)]	-	+	- (or 0)
Stock Splits [Dharan and Ikenberry (1995), Ikenberry, Rankine, and Stice (1996)]	+	+	+

Spinoffs [Miles and Rosenfeld (1983), Cusatis, Miles, and Woolridge (1993)]	+	+	+ (or 0)
---	---	---	----------