STOCK PRICE REACTIONS TO
EARNINGS ANNOUNCEMENTS:
A SUMMARY OF RECENT ANOMALOUS EVIDENCE
AND POSSIBLE EXPLANATIONS

Working Paper #703
Victor L. Bernard*
The University of Michigan

Forthcoming: Advances in Behavioral Finance, Richard Thaler, Editor
1992: Russell Sage Foundation, Publisher

*Price Waterhouse Professor of Accounting, University of Michigan and Marvin Bower
Visiting Fellow, Harvard University. I am grateful for useful comments and suggestions
from Ray Ball, Werner DeBondt, John Hand, Gene Imhoff, Prem Jain, S.P. Kothari, Josef
Lakonishok, Krishna Palepu, Stephen Penman, Jay Ritter, Terry Shevlin, Thomas Stober,
Richard Thaler, Jacob Thomas, Jim Wahlen, and Paul Zarowin, and for unpublished
evidence supplied by Jim McKeown and Hans Stoll.

FOR DISCUSSION PURPOSES ONLY
This material may not be quoted or reproduced without the expressed permission of the
Division of Research.

COPYRIGHT 1992
The University of Michigan, School of Business Administration,
Ann Arbor, Michigan 48109-1234
1. INTRODUCTION

No firm-specific performance measure is more widely reported, followed, and analyzed than accounting earnings. Thus, although claims of incomplete initial stock price reactions to earnings announcements have existed in the finance and accounting literature for more than 20 years, they have often been greeted with skepticism (e.g., Ball [1978]). The skepticism stems not just from the economic logic underlying the efficient markets hypothesis, but also from direct concerns about potential research design weaknesses, including a failure to control fully for risk. Some of the recent evidence, however, makes it more difficult than ever to dismiss earnings-related anomalies as the product of research design flaws.

This paper reviews recent evidence on market efficiency with respect to accounting earnings. The survey includes evidence indicating that the average initial response to earnings announcements is an underreaction, as well as other evidence that has been interpreted to indicate extreme stock price movements may represent overreactions to earnings.

There are several categories of anomalous evidence consistent with an initial underreaction to earnings reports. First, there are more than twenty (non-independent) studies of "post-earnings-announcement drift": the phenomenon whereby estimated cumulative abnormal returns continue to drift upward (downward) after the announcement of earnings increases (decreases). A second well known anomaly, the so-called "Value Line enigma" (Copeland and Mayers [1982]), has recently been shown to be driven primarily by post-announcement drift (Affleck-Graves and Mendenhall [1990]). If indeed stock prices underreact to earnings data, the phenomenon could also potentially explain at least a portion of a third anomaly, the "PE effect" (Basu [1977]).

Among the recent evidence on the possibility of an incomplete initial response to earnings announcements, the most puzzling appears in Bernard and Thomas [1989,1990], Freeman and Tse [1989], Mendenhall [1991], and Wiggins [1991]. Each of these studies entertains the possibility that post-announcement drift arises because stock prices fail to reflect fully what current earnings imply, on average, about earnings in subsequent quarters. As a result, when the subsequent quarters' earnings are announced, stock prices appear to reflect some surprise to earnings changes
that should have been predictable in advance. Specifically, the short-term (three-day) reactions to
the announcement of earnings for each of the next four quarters are partially predictable, based on
the current quarter's earnings (Bernard and Thomas [1990] and Wiggins [1991]). Even more
surprisingly, the signs and magnitudes of the predictable three-day reactions are related to the
autocorrelation structure of earnings, as if stock prices failed to reflect the extent to which each
firm's earnings series differs from a simple seasonal random walk. Several features of the
evidence are difficult to explain as the product of a failure to control for risk, or transactions costs.

In seeming contrast to the above-mentioned studies, evidence in DeBondt and Thaler
[1987] and Ou and Penman [1989a] appears consistent with stock prices overreacting to current
changes in earnings. Specifically, both studies document positive (negative) estimated abnormal
stock returns, for portfolios that previously generated inferior (superior) stock price and earnings
performance—as if the prior period stock price behavior constituted an overreaction to the earnings
developments. However, Ou and Penman find no evidence of overreaction to earnings on the
average, and conclude that their evidence is most appropriately characterized as an underreaction to
financial statement data that in certain cases predicts current earnings changes are transitory.

Both DeBondt and Thaler [1987] and Ou and Penman [1989] have sparked continuing
research. The interpretations of DeBondt and Thaler [1987] are challenged by Ball and Kothari
[1989], Zarowin [1989], and Klein [1990], but are supported by DeBondt and Thaler [1990] and
Chopra, Lakonishok, and Ritter [1991]. The robustness or interpretation of the Ou-Penman
results have been questioned by Greig [1990], Holthausen and Larcker [1990], and Stober [1991],
but Holthausen and Larcker identify a closely related portfolio strategy that appears much more
robust in its ability to predict abnormal returns. While it is too early to draw firm conclusions from
a literature in such flux, the recent evidence in Chopra, Lakonishok, and Ritter [1991] appears to
challenge conventional theories of market efficiency and asset pricing. Whether the evidence
should be characterized as an overreaction to earnings is, however, uncertain.

The remainder of this paper is organized as follows. Section 2 is devoted to studies of
apparent incomplete responses to earnings announcements. Within this section, all studies
accumulated prior to 1990 can all be characterized as examinations of possible underreaction to earnings, but the most recent evidence cannot be categorized so neatly. For this reason, the title of this section refers to incomplete initial responses to earnings announcements, thus leaving open the possibility that subsequent corrections can go in a direction opposite to the initial response. In contrast, section 3 refers only to research focused on possible overreaction to earnings. I summarize some debates found in recent reexaminations of Ou and Penman [1989a] and DeBondt and Thaler [1987], and discuss how such evidence can be reconciled with the larger body of work on underreaction to earnings. Section 4 presents a brief summary of other recently studied earnings-related anomalies (Ball and Kothari [1991] and Hand [1990]). Concluding remarks appear in section 5.

2. EVIDENCE SUGGESTING AN INCOMPLETE INITIAL RESPONSE TO EARNINGS ANNOUNCEMENTS

It has long been documented that there is much unusual stock price activity immediately surrounding earnings announcements. Beaver [1968] showed that unusual volatility and volume was largely concentrated within the week of the announcement. Patell and Wolfson [1984] showed that abnormal returns to knowledge of earnings were greatest within the 30 minutes of the announcement, with most of that return occurring within the first 5 to 10 minutes. The question left unanswered by this research, however, is whether this initial quick response is too large or too small. If the initial reaction represents an over- or underreaction that is corrected over a long period of time (perhaps as expectations about future earnings fail to materialize), the correction might not be apparent over short post-announcement intervals. Below, I review research consistent with the initial reaction being (on average) too small, and being completed over a period of at least six months.
2.1 Post-earnings-announcement drift: evidence accumulated prior to 1990

The basic phenomenon

Ball and Brown [1968] were the first to document that, subsequent to the announcement of earnings, cumulative abnormal returns (CARs) continue to drift up for "good news" firms and down for "bad news" firms. (In this case, good (bad) news can be defined as an earnings increase (decrease).) The same basic phenomenon was also documented by Jones and Litzenberger [1970], Joy, Litzenberger, and McEnally [1977], Watts [1978], Givoly and Lakonishok [1979], Latane and Jones [1979], Rendleman, Jones, and Latane [1982], Foster, Olsen, and Shevlin [1984], and at least 15 other studies.

The early studies of post-earnings announcement drift suffered from a variety of limitations (in part due to data constraints) that could have biased the results. (See Ball [1978], Foster, Olsen, and Shevlin [1984], and Bernard and Thomas [1989] for discussions.) However, even after the known research design flaws were corrected, post-earnings announcement drift remained apparent. A graphical depiction of the drift from a recent study (Bernard and Thomas [1989]) appears in Figure 1. The figure is based on a sample of approximately 85,000 observations for NYSE and AMEX stocks over the 1974-1986 period.

In Figure 1, firms are assigned to one of ten portfolios on the basis of their standardized unexpected earnings, or SUE. The procedure for calculating SUEs involves producing a statistical forecast of earnings, based on the model proposed by Foster [1977].\(^1\) The difference between actual earnings and the forecast is scaled by the historical standard deviation of the forecast errors to arrive at the SUE. For a given quarter, a firm's SUE is then compared to the distribution of all sample firms' SUEs from the prior quarter to place the firm in a decile portfolio. Abnormal (size-adjusted) returns for each portfolio are then cumulated beginning the day after the earnings announcement to estimate the post-announcement drift.

---

\(^1\) The model is a univariate first-order autoregressive model in seasonal differences. A forecast based on a naive seasonal random walk model yields similar results.
Figure 1 indicates that post-earnings-announcement abnormal returns increase monotonically across the ten SUE decile portfolios. Over the 60 trading days subsequent to the earnings announcement, firms with extreme good earnings news (SUE decile portfolio 10) experience a mean abnormal return of nearly 2 percent, while firms with extreme bad news (SUE decile portfolio 1) experience a negative abnormal return of approximately the same absolute magnitude. Thus, a movement of funds from the SUE 1 portfolio to the SUE 10 portfolio appears to generate a 60-day incremental abnormal return of 4.2 percent (before transactions costs), or about 18 percent on an annualized basis. (These amounts are based on summed returns that implicitly assume daily rebalancing, but are nearly identical when based on buy-and-hold portfolios.)

Larger abnormal returns are estimated when the research design is altered in particular ways. First, the absolute magnitude of the drift is larger for smaller firms (Foster, Olsen, and Shevlin [1984]). Second, the drift continues beyond the 60-trading-day interval displayed in Figure 1. Figure 2 (based on the same sample underlying Figure 1) shows that over a 180-trading-day post-announcement interval, a combined long position in SUE portfolio 10 and a short position in SUE portfolio 1 generates an abnormal return of approximately 10 percent, 9 percent, and 4.5 percent for small, medium, and large firms, respectively. Third, the drift is about 50 percent larger when SUEs are measured relative to analysts' forecasts, rather than the statistical forecasts used in Figures 1 and 2 (Freeman and Tse [1989, Table 7]).

The plausibility of risk as an explanation for post-announcement drift

While the evidence in Figures 1 and 2 is consistent with a market inefficiency—specifically, a delayed response to earnings announcements—it might also be explained by a failure to control fully for risk. That is, the good news firms may be riskier, and the bad news firms may be less risky, than the benchmark portfolios to which they are compared. Ball [1978] explains why such a difference in risk might be expected in this context. Ball, Kothari, and Watts [1990]
set forth the economic logic that would predict temporary risk increases (decreases) for firms in the good news (bad news) category, and provide evidence that such risk shifts do occur.

The Bernard and Thomas [1989] study was largely devoted to assessing the plausibility that post-earnings-announcement drift can be explained by a failure to control fully for risk. There are several aspects of the evidence in that study which cast doubt on a risk-based explanation.

First, Bernard and Thomas tested the Ball-Kothari-Watts hypothesis that betas shift up (down) during years with good (bad) earnings news. The premise of the hypothesis is that, for firms experiencing increases (decreases) in risk, a portion of the higher (lower) risk will be passed on to customers and/or suppliers in the form of price changes, thus translating into contemporaneous earnings increases (decreases). Using the Ball-Kothari-Watts methodology, Bernard and Thomas found evidence of beta shifts in the predicted direction, and noted that a portion of the shift persisted into the post-announcement period. However, the shifts were only about 8 to 13 percent as large as necessary to explain fully the magnitude of the drift.

Second, there was no evidence to indicate that long (short) post-announcement positions in extreme good (bad) news stocks were risky along any of five dimensions specified in the Chen, Roll, and Ross [1986] test of arbitrage pricing theory.

Third, there was little evidence that any risk to which the SUE portfolios were exposed had surfaced in the form of a loss. That is, estimated abnormal returns to a zero-investment portfolio with long (short) positions in extreme good (bad) news stocks were consistently positive, regardless of macroeconomic conditions. Over the 13 years from 1974 through 1986, the estimated abnormal returns were positive all 13 times. Over 50 quarters, the estimated abnormal returns were positive as many as 46 times, and the gains in these quarters exceeded the cumulative losses in the four remaining quarters by a factor of 35 to 1.

Fourth, mean raw (total) returns on extreme bad news stocks were so low as to raise doubts about whether declines in risk of any kind could plausibly explain their magnitude.

---

2 Bernard and Thomas's examination of risk was motivated in part by the first version of the Ball, Kothari, and Watts study, written in 1988.
Specifically, the raw returns were less than Treasury bill rates during the week after the earnings announcement, and were only slightly greater than Treasury bill rates during the first two months of the post-announcement period. While capital asset pricing theory does not rule out equilibrium expected returns on risky assets that are less than the risk-free rate, such low returns are permitted only under special conditions that many would find implausible in this context. Specifically, extreme bad news stocks would have to offer some hedge, the value of which exceeds the cost of any other risk to which the stocks are exposed.

Perhaps the most compelling result in the Bernard-Thomas [1989] study is the most simple: that the returns on zero-investment portfolios based on SUE are consistently positive. That result can be aligned with a claim that a SUE strategy is risky only if (1) the infrequency of losses in the 1974-1986 period is extremely unusual, relative to what would be observed over a longer period, or (2) the risk associated with the SUE strategy gives rise to infrequent but catastrophic losses, none of which was observed within this 13-year time span.

Figure 3 combines the evidence from three studies of post-announcement drift to examine whether the infrequency of losses noted above is unusual. The Latane, Jones, and Rieke [1974] study, the Rendleman, Jones, and Latane [1982] study, and the Bernard and Thomas [1989] are used here because they disclose returns to SUE strategies on a period-by-period basis. The evidence must be used cautiously, because the methods vary across the three studies (e.g., in terms of how extreme are the SUEs underlying the test portfolios), and the earliest study might have contained biases not present in later work.\(^3\) Notwithstanding these caveats, it is interesting to see that the combined studies indicate that long (short) positions in extreme high (low) SUE stocks generated positive estimated abnormal returns in each year from 1965 through 1986. If the SUE strategy were risky along some dimension not fully controlled for in the studies, it is difficult to understand why these zero-investment portfolios did not generate a loss for 22 consecutive years.

---

\(^3\) Latane, et al. assume that quarterly earnings were made public by two months after the fiscal quarter end, which could induce some upward bias in abnormal returns. (On the other hand, since most firms announce earnings well before this date, the Latane, et al study may have missed a significant portion of the drift.)
Despite the above evidence, the Bernard-Thomas conclusion that risk shifts could at best explain a small portion of post-announcement drift is not unchallenged. Ball, Kothari, and Watts [1990] estimate that over the nine months following the presumed earnings announcement month, the difference in betas between top decile and bottom decile earnings performers is .29, or about twice as large as the average of the differences estimated by Bernard and Thomas for the first three post-announcement quarters. After controlling for beta, the Ball-Kothari-Watts estimate of the difference in abnormal returns between the extreme decile portfolios is only 2.98 percent over the nine months—certainly smaller than the drift estimated elsewhere. However, the issue of post-announcement drift is not the primary focus of Ball, Kothari, and Watts, and they "caution against making fine comparisons," due to the variety of differences in research design between their study and others. Perhaps most importantly, they form portfolios on the basis of changes in annual earnings, rather than quarterly earnings. Given that much of the change in annual earnings is "old news" by the time it is announced, this approach tends to reduce the power to capture the full magnitude of post-announcement drift. At the same time, so long as beta shifts have any degree of permanence, using annual data increases the power to identify firms with large beta shifts. Since the design magnifies the importance of beta risk premiums, relative to post-announcement drift, the study cannot provide clear indications about the extent to which the drift is explained by risk.

Implementation problems and transactions costs as an explanation for post-announcement drift

Bernard and Thomas [1989, section 3.2.4] find no implementation problems that would prevent traders from generating abnormal returns based on SUEs, at least before considering transactions costs and short-selling restrictions. They examine a strategy that (1) followed the approach of Holthausen [1983] to assure that SUE portfolio assignments were based entirely on information available prior to the assignment; (2) took positions in stocks only when both extreme good news and extreme bad news stocks were available, to ensure that the strategy requires zero
initial investment;\textsuperscript{4} (3) involved computing buy-and-hold (compounded) returns, as opposed to cumulative returns that implicitly assume daily rebalancing; and (4) avoided the use of control portfolios that would implicitly assume the feasibility of taking new positions in hundreds of stocks each day. The estimated abnormal returns to this strategy are similar to those already described above, in terms of magnitude, insensitivity to known risk factors, and consistency through time.

Whether transactions costs would offset gains to trading on SUEs is more difficult to assess, but it seems unlikely that they would. Transactions costs for the \textit{average} transaction can be imputed by comparing the sum of economy-wide commission income, market-maker trading gains (including specialists' profit from the bid-ask spread), and underwriting profits with the aggregate trading volume on all stock exchanges. Using this approach, Stoll [1991] estimates mean round-trip transactions costs to be 1.2 percent. For institutions, the average roundtrip costs would of course be lower—perhaps on the order of 0.75 percent. Even after doubling these average costs to account for positions involving both good news and bad news firms, they are still much lower than the abnormal returns presented in Figure 2 for the nine-month post-announcement interval: 4.5 percent, 8.9 percent, and 9.9 percent for large, medium, and small stocks, respectively.\textsuperscript{5}

Although \textit{mean} roundtrip transactions costs fall far short of the amount necessary to offset gains from post-announcement drift, the costs could be much higher for those who trade aggressively and thus exert price pressure or bear a higher share of the bid-ask spread. However, given that post-announcement drift lasts for months, it is not clear why a trader would act on it aggressively. Moreover, in so far as the marginal returns to trading on SUEs are concerned,

\textsuperscript{4} On 86 percent of the trading days, both extreme good news and extreme bad news stocks were newly available, so positions could be taken immediately. In 97 percent of all cases, a match of good and bad news stocks could be made by waiting less than three days.

\textsuperscript{5} The costs and restrictions pertaining to short selling of bad news stocks are also relevant to those who do not already hold the bad news stocks. However, since bad news stocks could always be sold by those holding them without bearing these costs, it is not clear why short selling should play any role in explaining post-announcement drift.
transactions costs are irrelevant for other traders already committed to buying or selling for reasons unrelated to earnings; it is not clear why their actions would not eliminate the drift.

There are more fundamental questions raised by this discussion, however. First, how could transactions costs of any magnitude explain post-announcement drift, given that trading occurs throughout the period of the drift? That is, given that traders have agreed to transact, why do they not transact at a price that appears to reflect all publicly available earnings information? Second, even if transactions costs could "explain" the drift, would we then label the stock market efficient with respect to earnings? Ball [1990] argues that a meaningful definition of efficiency should require that frictions such as transactions costs and trading restrictions do not influence price. (If instead, efficiency requires only the absence of trading profits net of transactions costs, even a market that is closed for trading would be considered efficient!)

2.2 Post-earnings-announcement drift: recent evidence

Evidence that stock prices reflect naive earnings expectations

Given that a long series of studies had failed to produce a satisfactory explanation for post-announcement drift based on research designs flaws, incomplete controls for risk, or transactions costs, recent research turned to alternative possibilities involving market inefficiency. Following up on the suggestion of Rendleman, Jones, and Latane [1987] and initial evidence produced by Bernard and Thomas [1989] and Freeman and Tse [1989], two recent studies (Bernard and Thomas [1990] and Wiggins [1991]) investigated the possibility that post-announcement drift arises because stock prices fail to reflect the implications of current earnings for future earnings.

The essential hypothesis tested by Bernard and Thomas [1990] and Wiggins [1991] is that prices do not fully reflect the extent to which current changes in quarterly earnings (relative to the comparable quarter of the prior year) signal future changes in earnings. More specifically, stock prices appear to reflect, at least partially, a naive earnings expectation: one based on a seasonal random walk, where expected earnings are simply earnings for the corresponding quarter from the
previous year. It is well known that earnings forecast errors based on such a naive model are correlated through time (e.g., Freeman and Tse [1989]). In contrast, in a market that fully impounds all prior earnings information, forecast errors should not be autocorrelated (by definition). What the recent studies examine is the possibility that market prices can be described partially as reflections of naive expectations, and that as a result, the reactions of prices to future earnings are predictable, just as the forecast errors of a naive expectation model are predictable.

For an investor who relies on a naive seasonal random walk forecast of earnings, forecast errors would be autocorrelated in a pattern that has remained consistent since at least 1946, the beginning of the first period for which the data have been studied (Watts [1975], Foster [1977], Bernard and Thomas [1990]). The cumulative evidence indicates that the pattern of autocorrelations can be viewed as including two components. First, there is a positive autocorrelation between seasonal differences (i.e., seasonal random walk forecast errors) that is strongest for adjacent quarters, but that remains positive over the first three lags. Thus, a change in earnings of quarter t (relative to the comparable quarter of the prior year) tends to be followed by progressively smaller changes of the same sign in quarters t+1, t+2, and t+3. Second, there is a negative autocorrelation between seasonal differences that are four quarters apart. That is, a portion (about one-fourth, on average) of the change for quarter t is "reversed" in quarter t+4; only the remaining portion of the initial change represents a "permanent" shock.

To offer a more specific description of the time series behavior of earnings, Table 1 presents summary statistics for the samples studied by Foster [1977], Hopwood and McKeown [1986], and Bernard and Thomas [1990]. Together, the three samples span the 1946-1986 period. In all three samples, the autocorrelations at the first three lags are positive but declining, and the autocorrelation at the fourth lag is negative. The general pattern shown in Table 1 is quite consistent across a variety of industries, and holds regardless of fiscal quarter alignment (Bernard and Thomas [1990]).

Under the hypothesis that stock prices fail to reflect the extent to which the earnings process tends to deviate from a seasonal random walk, abnormal stock returns would depend on
forecast errors that are autocorrelated and therefore predictable; the autocorrelations would mimic
the same pattern observed in Table 1. For example, if an earnings increase over the prior year
were announced in quarter t, the market would be "surprised" to learn that earnings tend to increase
again (over the prior year) in quarters t+1, t+2, and t+3, and thus would (on average) react
positively to the earnings announcements for those quarters. The magnitude of this surprise would
be declining over the three quarters, just as the corresponding autocorrelations decline. The
hypothesis also predicts that the market would be surprised that a portion of the quarter t earnings
increase tends to be reversed in quarter t+4, and thus would (on average) react negatively to the
announcement for that quarter. Note that the predictions of market reactions to announcements for
quarters t+1 through t+4 are based solely on earnings information available as of quarter t. Note
also that the prediction pertaining to quarters t+1 through t+3 is consistent with the view that stock
prices initially underreact to quarter t earnings information, but the prediction for quarter t+4 is
more readily characterized in terms of a reversal of a prior overreaction.

The evidence in Bernard and Thomas [1990] and Wiggins [1991] is surprisingly consistent
with the above predictions. In Bernard and Thomas [1990], a long (short) position in quarter t's
extreme good news (bad news) firms yields a 3-day abnormal return surrounding the next three
earnings announcements that is positive and declining: 1.32 percent, 0.70 percent, and 0.04
percent, respectively. The same position yields a 3-day abnormal return around the fourth
subsequent announcement that is negative: -0.66 percent.

A graphical depiction of the main result in Bernard and Thomas [1990] appears in Figure 4.
The figure describes the behavior of the combined long (short) positions in quarter t's good news
(bad news) stocks throughout the four quarters subsequent to the announcement for quarter t. The
data are aligned so that the announcement dates for each subsequent announcement coincide for all
firms in the portfolio, leaving inter-announcement windows of varying lengths. The pattern of the
three-day stock price movements—positive but declining reactions around the first three

6 Good news (bad news) is defined in Bernard and Thomas [1990] in terms of SUEs where the earnings expectation
is based on a seasonal random walk with drift.
subsequent announcements, and a negative reaction at the fourth announcement—is consistent with the market being surprised that earnings deviate from a seasonal random walk. Abnormal returns measured over quarter-long intervals show the same pattern, but the concentration of abnormal returns around earnings announcements suggests that "corrections" of the earnings expectations embedded in stock prices tend to occur more frequently on those days.

The trading strategy suggested by the above would involve holding positions in the zero-investment portfolios in Figure 4 through the third subsequent announcement and then reversing the position during the fourth quarter. The indicated abnormal returns are approximately 11.5 percent, 10 percent, and 5 percent for small, medium, and large firms, respectively, for the four-quarter holding period. Bernard and Thomas [1990] describe another strategy that takes advantage of the concentration of drift around subsequent announcements; it involves taking positions 15 days prior to the expected date of a given subsequent earnings announcement (assumed equal to the announcement date of the prior year) and holding the position through the announcement. Long (short) positions are maintained in stocks that generated an extremely high (low) SUE within the last three quarters, or an extremely low (high) SUE four quarters prior. The indicated abnormal return is 4.2 percent for an average holding period of 15 days.

One might posit that the evidence is explainable as a function of the market's justifiable expectation of low levels of autocorrelation in seasonally-differenced earnings, and resulting justifiable surprise at the actual degree of autocorrelation. However, two features of the data cast doubt on this possibility. First, the estimates in Table 1 suggest that the degree of autocorrelation was if anything lower for the 1974-1986 period studied by Bernard and Thomas [1990] than in the prior years, although noncomparability across samples renders this observation tentative. Second, the behavior of abnormal returns described above for the overall 1974-1986 period holds consistently for each of the 13 years studied. Although the market might err in its expectations

---

7 The predicted signs are obtained in every year, with the exception of the third subsequent announcement, for which the relation was not predicted to be strong.
about the degree of autocorrelation in earnings, it is difficult to explain how it could justifiably err in the same direction year after year.

Overall, the evidence offers the following implications. First, it creates several added obstacles to contentions that the drift might ultimately be explained by errors in the methodology used to estimate expected returns. For example, an explanation based on failure to control for risk would now have to argue that good news firms experience delayed *increases* in risk over 3-day intervals that coincide with each of the next three earnings announcements, and then a *decrease* in risk over the 3-day interval coinciding with the fourth subsequent announcement. The converse would have to hold for bad news firms. At a minimum, any rationale for such behavior must have a more complex structure than explanations suggested to date. Ball, Kothari, and Watts [1990] conclude that "it is unlikely that risk changes can explain even a small proportion of the estimated abnormal returns that Bernard and Thomas [1989, 1990] attribute to specific days, such as 3 to 5 days immediately after the earnings announcement and those surrounding the following quarters' earnings announcements."\(^8\)

Second, by linking what appears to be the elimination of a discrepancy between prices and fundamentals to prespecified information events (i.e., earnings announcements), the evidence is perhaps the most direct indication to date that a market-efficiency anomaly is rooted in a failure of information to flow completely into price. Ball [1990] cautions reasonably that whether the results are consistent with efficiency or inefficiency is simply not knowable, at least today, but does note that the Bernard-Thomas [1989, 1990] evidence "points to the delayed reaction hypothesis."

\(^8\) The evidence also makes it more difficult than ever to argue that transactions costs are the sole cause of post-announcement drift. Even if transactions costs cause "sluggishness" in prices, it is hard to understand why the resulting "mispricing" would tend to be eliminated around subsequent earnings announcements according to a pattern that is related to the time series properties of earnings. It is particularly difficult to reconcile "price sluggishness" with the "return reversal" detected upon the announcement of earnings for quarter t+4.
But why would stock prices appear to reflect naive earnings expectations?

To suggest that stock prices reflect naive earnings expectations does not explain the anomaly with which we started; it simply raises another unanswered question. Under what kind of equilibrium could stock prices reflect naive earnings expectations?

Construction of an equilibrium explanation for the evidence would be difficult at best. However, it is less difficult to explain why, if stock prices are partially influenced by naive earnings expectations, at least some informed market participants would be unable to exploit the apparent profit opportunity. Specialists on the exchange can exploit what they believe to be mispricing only by exposing themselves to firm-specific risk (Bagehot [1971]), not to mention rendering themselves vulnerable to charges of a failure to maintain an orderly market. Even though transactions costs and other trading costs appear unlikely to offset the drift for all traders, they may preclude at least some speculators from fully exploiting SUE-based strategies. The possibility that SUE strategies may involve some systematic risk (Ball, Kothari, and Watts [1990]) would further limit their attractiveness.

Even if speculation based directly on SUEs cannot entirely eliminate any discrepancies between prices and fundamental values that are related to prior earnings, it remains to be explained how such discrepancies could arise in the first place. In other words, why wouldn't all available earnings information be taken into account even by traders who are not explicitly speculating on SUEs? Liquidity traders might be willing to trade at prices that fail to reflect all available information, but their seemingly random trades could not lead to a relation between mispricing and prior earnings with the systematic patterns described earlier. A full explanation requires a systematic tendency for some traders to maintain earnings expectations that are anchored too heavily on the comparable earnings of the prior year, in combination with the costs or risks mentioned above that might prevent speculators from fully exploiting this tendency.

---

9 Indeed, most (not all) of the fund managers that trade on SUE-related signals and with whom I have spoken do not attempt to sell short on bad news, and trade only within a universe of the 500 or 1000 largest stocks on the NYSE, where transactions costs are lower, and where large positions can be taken without much concern about price pressure.
Andreassen [1987, 1990] is among those who have hypothesized that systematic psychological forces could influence stock price behavior. Certainly, there is a large psychological literature documenting a tendency for individuals in certain prediction contexts to "anchor" on some value and place little weight on recent changes in a series.\(^\text{10}\) Andreassen [1987] and Andreassen and Kraus [1990] note that this is particularly true unless, for those making the prediction, the recent changes are salient and can be attributed to a stable underlying cause.

Whether such basic psychological forces could survive among professionals in competitive markets is unanswered, and there alternative reasons market participants might update forecasts only gradually. In the case of one key group of players, financial analysts, such behavior might be encouraged by incentive structure. That is, there may be little to gain and much to lose by adjusting a forecast to a level far from the current consensus, if the analyst believes his/her forecast is already likely to be the most accurate.

Though the causes of partial and gradual adjustment to earnings news are not clear, there is evidence that analysts underreact to recent earnings changes in predicting future earnings. Mendenhall [1991] shows that Value Line quarterly earnings forecasts are revised insufficiently to account for the most recent Value Line forecast error (i.e., Value Line forecast errors are positively autocorrelated at the first lag). Abarbanell and Bernard [1991] confirm Mendenhall's result and document declining but positive autocorrelations over the first three lags, consistent with what would be expected if the forecasts were influenced in part by the seasonal random walk model underlying the tests in Bernard and Thomas [1990].

Abarbanell and Bernard also conduct tests to indicate that even though the Value Line analysts underreact to recent earnings, the expectations embedded in stock prices underreact to a much greater extent and more closely resemble naive (seasonal random walk) earnings expectations. Thus, post-announcement drift cannot be fully explained by inefficiencies in analysts

\(^\text{10}\) Another set of literature documents "recency effects" or "nonregressive predictions," where recent changes in a series receive heavy emphasis; this is the literature to which DeBondt and Thaler [1985] refer in their discussion of possible overreaction in the stock market. Andreassen [1987] and Andreassen and Kraus [1990] are among those who have attempted to discriminate between contexts where predictions are regressive or nonregressive.
forecasts. A full explanation requires an understanding of not only why analysts' forecasts tend to underreact to earnings information, but also why market prices underreact to analysts' forecasts.

2.2 Other recent evidence consistent with an incomplete initial response to earnings

*The Value Line enigma*

It has long been recognized that investment strategies based on Value Line rankings of stocks appear to generate positive risk-adjusted abnormal returns (Black [1973], Copeland and Mayers [1982]). The evidence constitutes an anomaly, because Value Line purports to base its rankings solely on information that should be publicly available before the rankings are published.

Three of the indicators underlying the Value Line rankings are based on earnings information: quarterly earnings surprise, earnings momentum, and a function that relates earnings to prices. Thus, the apparent success of the Value Line strategy is consistent with stock prices failing to completely impound recent earnings information. This raises the question of whether the so-called Value Line enigma and post-earnings-announcement drift constitute the same basic anomaly. Indeed, Affleck-Graves and Mendenhall [1990] conclude that "the Value Line enigma is actually a manifestation of post-earnings-announcement drift."

There are indications that the funds created by Value Line to trade on their own rankings have generated performance that may fall considerably short of the "paper portfolio" profits studied in the academic literature (Bodie, Kane, and Marcus [1989, pp. 514-517]). The discrepancy raises questions about whether the costs of implementing such a trading strategy (at least on a stand-alone basis) are larger than they appear on the surface, and/or about how closely Value Line has followed their own rankings.

---

11 Huberman and Kandel [1990] show that the apparent abnormal returns have a time series behavior consistent with an autoregressive risk premia, but do not attempt to identify the source of such risk.
The EP effect

If stock prices underreact to earnings information, or if prices reflect valuation errors that are unrelated to earnings and corrected over time, then earnings-price (EP) ratios should serve as positive predictors of abnormal returns. Indeed, one of the earliest documented anomalies was the "EP effect": stocks with high earnings-price (EP) ratios generate higher returns than stocks with low EP ratios (Latane and Tuttle [1967], Latane, Tuttle and Jones [1969], Basu [1977]). While there has been considerable debate about whether the EP effect survives adjustments for the size effect (Reinganum [1981], Basu [1983], Cook and Rozeff [1984]) and various sources of hindsight bias (Banz and Breen [1986]), the most comprehensive examination to date suggests the answer is yes (Jaffe, Keim, and Westerfield [1989]).

Although EP ratios are certainly correlated with SUEs (Kim [1987]), there are three related bits of evidence that suggest the EP effect is not driven by a delayed response to earnings information. First, Basu [1983] finds that the EP effect is just as strong when portfolios are formed six months after the presumed month of the earnings announcement, as it is when they are formed immediately after the presumed announcement date. Second, Jaffe, Keim, and Westerfield [1989] find a much larger EP effect in January, which is 10 or 11 months after the announcement of earnings on which the EP ratio is based. Third, Kim [1987] finds that most of the EP effect is independent of post-announcement drift (and vice versa).

What causes the EP effect is and probably will continue to be an open question. One possible driving force is mean reversion in the denominator—price—that could reflect the correction of prior valuation errors unrelated to earnings. If so, then the EP effect would be related to the literature to be summarized in the following section, which investigates long-term reversals of prior period stock price movements. Indeed, the relation of the EP effect to the January effect and the size effect (Jaffe, Keim, and Westerfield [1989]) is something it has in common with the long-term reversals studied by DeBondt and Thaler [1985].

Another distinct possibility is that the EP effect is largely explainable as a risk premium. Basu's [1983] evidence suggests that, after controlling for size, the remaining EP effect continues
unabated for at least 12 months after the position is taken, as would be suggested by the risk-based explanation advanced by Ball [1978]. Although the EP effect is not eliminated by controlling for systematic risk (Jaffe, Keim, and Westerfield [1989]), other risk factors could be at work. Unlike portfolios built to take advantage of post-announcement drift, zero-investment portfolios designed to exploit the EP effect do not produce estimated abnormal returns that are consistently positive over time (Kim [1987]). That makes a risk-based explanation at least plausible, if not likely.

3. EVIDENCE CONSISTENT WITH AN OVERREACTION TO EARNINGS ANNOUNCEMENTS

While most of the evidence discussed thus far is suggestive of an underreaction to earnings information, other evidence has been interpreted as indicating an overreaction. At least some of this research was motivated by a desire to identify the source of long-run mean reversion in stock prices documented by DeBondt and Thaler [1985] and others. DeBondt and Thaler showed that firms with prior extreme negative stock price performance (the "losers") seem to outperform those with prior extreme positive performance (the "winners"), as if a portion of the prior stock price movement constituted a deviation from fundamental values. If this is the appropriate interpretation of the evidence, it raises an interesting question: could the deviations be driven by an overreaction to prior years' earnings trends? For example, could market prices fail to reflect that extreme annual earnings changes tend to be reversed subsequently, as documented by Brooks and Buckmaster [1976] and Freeman, Ohlson, and Penman [1982]?

The first study to entertain the possibility of overreaction to earnings was DeBondt and Thaler [1987]. Ou and Penman [1989a] also present evidence that has features consistent with

---

12 Long-run mean reversion in stock prices was also documented by Fama and French [1988] and Poterba and Summers [1988], among others. However, the evidence most pertinent here is that pertaining to the firm-specific mean reversion documented by DeBondt and Thaler, as opposed to mean reversion at more aggregate levels.
overreaction (although as discussed below, the authors have a different interpretation). In this section we review the evidence and discuss two questions. First, to what extent is the evidence consistent with overreaction, versus alternative explanations? Second, how can the evidence be reconciled with the previously described results suggesting underreaction to earnings?

Some central features of the DeBondt-Thaler [1987] and Ou-Penman [1989] evidence are summarized in Table 2. DeBondt and Thaler form "winner" and "loser" portfolios on the basis of market-adjusted stock returns over the prior four years; the overreaction hypothesis predicts a partial reversal of those returns during the subsequent test period. Table 2 Panel A shows that, consistent with this hypothesis, losers outperform winners by 37 percent over the four-year test period. In support of the possibility that the result could be driven by overreaction to earnings news, DeBondt and Thaler point to evidence (also summarized in Panel A) that indicates a pattern in earnings similar to that found in returns. That is, after losers (winners) experience earnings declines (increases) during the formation period, earnings move in the opposite direction in the subsequent test period. The evidence appears consistent with a failure of stock prices to reflect that annual earnings do not strictly follow a random walk, but show some mean reversion in the tails (Brooks and Buckmaster [1976]). In that sense the evidence seems related to the previously discussed finding that stock prices appear to reflect only incompletely the mean reversion in quarterly earnings that surfaces in the form of a negative autocorrelation in seasonal differences at the fourth lag (Bernard and Thomas [1990]).

Table 2 Panel B shows that patterns of returns and earnings reported for Ou and Penman's portfolios have much in common with those of DeBondt and Thaler. Ou and Penman form portfolios on the basis of a "Pr" measure, which is essentially the outcome of a computerized fundamental analysis; Pr represents an estimate of the probability of an annual earnings increase in the coming year, based on a function of financial statement variables identified and estimated using only historical data. A key factor contributing to the success of Pr as a predictor of future earnings changes is mean reversion in earnings scaled by equity. Firms with recent earnings declines have high Prs and subsequently increasing earnings; the opposite earnings pattern holds for low Pr
firms. In this sense, the high (low) Pr firms correspond to DeBondt and Thaler's "losers" ("winners"). Table 2 Panel B shows that the correspondence is apparent not only in earnings behavior, but also in returns behavior. That is, high Pr firms' stocks underperform low Pr firms during the portfolio formation period, just as DeBondt and Thaler's losers underperformed winners; the pattern is reversed in the test period.

**Alternative explanations for the evidence: DeBondt and Thaler**

Whether the DeBondt-Thaler evidence should be interpreted as evidence of overreaction is a subject of continuing debate. DeBondt and Thaler discuss features of the data that would not follow directly from the overreaction hypothesis: in particular, the concentration of the return reversals in the month of January. Ball and Kothari [1989] suggest that prior winners (losers) experience decreases (increases) in beta that are sufficiently large to explain nearly all of the estimated abnormal returns to the DeBondt-Thaler strategy. Specifically, Ball and Kothari document that when one controls for beta shifts, the DeBondt-Thaler strategy generates an abnormal return that is never statistically significant in any of five post-ranking years.

The recent work of Chopra, Lakonishok, and Ritter [1991] also examines the overreaction hypothesis while controlling for beta shifts. Chopra, Lakonishok, and Ritter argue that Ball and Kothari underestimate the abnormal returns by assuming that the compensation per unit of beta risk

---

13 Andreassen [1990] advances a behavioral explanation that could lead to a seasonal in overreaction.

14 DeBondt and Thaler [1987] and Chan [1988] also control for beta in a way that does not assume beta-stationarity across the formation and test periods. Ball and Kothari find larger beta shifts and smaller abnormal returns than DeBondt and Thaler; the source of the discrepancy is unexplained but may be in part related to Ball and Kothari's use of buy-and-hold returns, rather than the cumulated returns used by DeBondt and Thaler. (Chopra, Lakonishok, and Ritter [1991] replicate the DeBondt-Thaler methods but substitute buy-and-hold returns for cumulated returns, and obtain beta estimates very similar to those of Ball and Kothari.) The beta shifts estimated by Ball and Kothari are also greater than those estimated by Chan [1988], in large part because Ball and Kothari use annual data (rather than monthly data) to estimate betas, and allow betas to shift annually. (The advantages of using annual data in this context are discussed in Handa, Kothari, and Wasley [1989]).

15 The mean annual abnormal return over the five years is 3.1 percent, which Ball and Kothari indicate may be biased upward as the result of a survivorship bias that is not present in DeBondt and Thaler's sample. In contrast, DeBondt and Thaler [1987, Table III] report an average annual return of 9.2 percent before controlling for beta shifts, and 5.9 percent after instituting such controls.
is equal to the excess return on the market portfolio. Although this is the amount implied by the Sharpe-Lintner CAPM, several studies have documented empirical estimates of the risk premium that are substantially lower (e.g., Fama and MacBeth [1973] and Tinic and West [1984]). When Chopra, Lakonishok, and Ritter rely on empirical estimates of the market price of risk rather than assumed values, and while controlling for size, they find that prior extreme losers outperform prior extreme winners by nearly 5 percent per year (on average across five years); the amount is larger (11 percent) for firms in the smallest decile. The abnormal returns are more pronounced in January, but are also some evidence of positive abnormal returns in the February-December period. It is interesting that the size-adjusted abnormal returns are essentially zero for large firms, and are attributable primarily to small firms, as if overreaction occurs only outside the segment of the market dominated by institutions.

In addition to re-examining the Ball-Kothari [1989] tests, Chopra, Lakonishok, and Ritter present another analysis that points to overreaction. They show that during the five years following portfolio formation (and after controlling for beta and size), a disproportionate share (about 20 percent) of the apparent superior performance of losers over winners is concentrated within the three-day intervals surrounding earnings announcements. The evidence suggests that earnings expectations for losers (winners) are unduly pessimistic (optimistic), and thus the market is surprised by the subsequent earnings announcements. The magnitude of the mean three-day announcement period abnormal return—.27 percent for extreme losers minus extreme winners—is small enough that, given the evidence of other anomalous announcement-day return behavior (Ball and Kothari [1991]), some caution must be exercised in its interpretation. Nevertheless, the evidence appears difficult to explain except as a correction of a prior overreaction.

---

16 Ball and Kothari [1991] show that mean abnormal returns are positive around earnings announcement days for small firms. The three-day abnormal return for the smallest quintile is about 1 percent. However, in zero-investment portfolios, this effect should "cancel out", unless the long and short positions are mismatched on size. The Chopra-Lakonishok-Ritter result survives a size control, suggesting it is not explained by the general phenomenon documented by Ball and Kothari.
Even if extreme stock price movements do constitute overreactions, and even if the overreactions tend to be corrected upon subsequent earnings announcement days, it remains to be shown that the genesis of the overreaction is a misinterpretation of past earnings numbers. Zarowin [1989] focuses directly on this possibility. He acknowledges that subsequent reversals of earnings among extreme prior period stock return performers are consistent with prior period overreactions to earnings. He suggests, however, that demonstration of overreaction to earnings requires predicting future period stock returns based on prior period earnings realizations. Zarowin thus forms portfolios of prior period extreme winners and losers defined in terms of changes in annual earnings. He hypothesizes that if stock prices overreact to earnings, there should be positive (negative) subsequent abnormal stock returns for prior losers (winners).

Although Zarowin finds evidence consistent with the overreaction hypothesis when he analyzes market-adjusted returns, he finds none after controlling for beta and the size effect. Thus, the Zarowin evidence casts some doubt on the overreaction hypothesis—at least as a phenomenon occurring for the average extreme earnings winner or average extreme earnings loser. However, if overreaction is not widespread and varies across firms with similar earnings experience, Zarowin's approach of identifying winners and losers on the basis of prior period earnings changes would lack power, relative to the DeBondt-Thaler approach to identifying winners and losers in terms of prior period stock returns. In fact, when one moves to Zarowin's approach from DeBondt and Thaler's approach, the resulting mean annual test-period market-adjusted return declines to 5.3 percent from 9.1 percent—consistent with a substantial loss of power. Of course, the decline in market-adjusted returns might also reflect a reduction in exposure to uncontrolled risks.

17 Chopra, Lakonishok, and Ritter [1991] argue that Zarowin's results may reflect a survivorship bias that would obscure an overreaction effect. However, Zarowin [p. 1387] discusses the issue and points to evidence that suggests the bias should not be serious.
18 Zarowin [p. 1391] reports a three-year market-adjusted return of 16.6 percent for long (short) positions in lowest (highest) quintile stocks based on prior earnings changes; this represents an average of 5.3 percent per year. DeBondt and Thaler [p. 572] report a four-year market-adjusted return of 24.6 percent (-11.7 percent) for positions in lowest (highest) quintile stocks, based on prior stock return changes; this represents an average of 9.1 percent for the combination of the two positions. The Zarowin test periods range from 1972 to 1984; DeBondt and Thaler's test periods range from 1970 to 1983.
Klein [1990] also examines the hypothesis that investors overreact to earnings numbers. She finds evidence more consistent with *underreaction* than overreaction. Specifically, she finds that analysts of loser firms (identified on the basis of prior years' stock returns) are unduly *optimistic* about future earnings prospects. (Analysts of winner firms have approximately unbiased expectations.) Klein's evidence may *appear* inconsistent with that of DeBondt and Thaler [1990], who show that analysts' forecasts of earnings changes tend to be too extreme and therefore "consistent with generalized overreaction". However, Abarbanell and Bernard [1991], who replicate the DeBondt-Thaler result, conclude that "whatever is causing the overreaction, it could not be earnings." Abarbanell and Bernard [1991] explain that the key to reconciling the seemingly disparate results in Klein [1990] and DeBondt-Thaler [1990] is to recognize that the extreme forecasts driving the DeBondt-Thaler overreaction conclusion tend to be those where analysts expect a *departure* from the recent earnings trends—not the continuation that an overreaction to earnings would imply. With respect to earnings, the evidence is actually more consistent with an *underreaction*, as Klein [1990] would suggest.

Firm conclusions about the proper interpretation of the DeBondt-Thaler [1985, 1987] evidence are difficult to draw currently, given that the literature in this area is still in flux. On the one hand, when one allows for shifts in risk and imposes the standard restrictions implied by the Sharpe-Lintner CAPM, the evidence suggests no statistically significant overreaction (Ball and Kothari [1989]). On the other hand, one of the key restrictions of the CAPM is inconsistent with the data, and when it is relaxed, evidence of overreaction again surfaces, especially for small firms (Chopra, Lakonishok, and Ritter [1991]).

What is more clear is that even if overreaction exists, it appears too complicated to be characterized as a simple function of recent earnings changes. That is a topic to which we return later, in our reconciliation of the evidence on under- and overreaction to earnings.
Alternative explanations for the evidence: Ou and Penman

Ou and Penman [1989a] note the similarity between their evidence and that of DeBondt and Thaler, but report some evidence to suggest the two anomalies are distinct. (Most important, the Ou-Penman abnormal returns are not concentrated in January.) Examination of Table 2 also indicates that the Ou-Penman Pr portfolios are far more extreme in terms of prior year earnings performance—suggesting that the Pr portfolios must have different content than the DeBondt-Thaler winner/loser portfolios.

Precisely how the Ou and Penman evidence should be explained—as indicating market inefficiency or something else—is still undergoing scrutiny. Consistent with market inefficiency, Stober [1991] finds that the Pr factors work best as predictors of stock returns in those cases where analysts' earnings forecasts appear to ignore some of the information in Pr. On the other hand, the robustness of the Ou and Penman result is questioned in preliminary work by Greig [1990]. In addition, Stober [1991] uses data not completely available at the time Ou and Penman conducted their work to show that Pr predicts abnormal returns for at least six years, with little or no sign of diminution. If indeed the abnormal returns last for many years, Pr would appear to be capturing a permanent risk shift rather than a discrepancy between prices and fundamentals that should be corrected over time.19

Holthausen and Larcker [1990] present perhaps the most serious challenge to the Ou and Penman Pr "strategy", by showing that it performs poorly after 1983. The year 1983 was the last year of the Ou-Penman tests, and even Ou and Penman noted that the strategy may not have been profitable in 1982 or 1983. However, when Holthausen and Larcker refine the Ou-Penman strategy only slightly, so that the "computerized fundamental analysis" is focused directly on predictions of stock returns (rather than earnings changes), the resulting estimated abnormal

---

19 One of the empirical issues that makes it difficult to interpret the Ou-Penman evidence, as well as much of the previously described evidence on overreaction, concerns controlling for size effects. Since small firms include a disproportionate share of prior "losers" and high Pr firms, the size effect may be caused in part by the DeBondt-Thaler overreaction effect, and/or the Ou-Penman PR effect. If so, controlling for size through the introduction of size control portfolios may "throw the baby out with the bath water." The Ou-Penman result is much stronger when based on market-adjusted returns than on size-adjusted returns.
returns are consistently positive. The authors "find it surprising that a statistical model, derived without consideration of any economic foundations, can earn excess returns of the magnitude (they) document," and have no explanation for the result.

*How can the evidence consistent with overreaction to earnings be reconciled with the previously described results suggesting underreaction to earnings?*

Evidence in DeBondt and Thaler [1987] and Ou and Penman [1989a] may appear in conflict with the evidence in section 2 suggesting an underreaction to earnings. The two sets of results are, however, potentially reconcilable through several avenues.

One possibility is that, even if overreactions occur in the stock market, they cannot be characterized as an overreaction to earnings *per se*. A predictable reversal of prior period extreme price movements is consistent with a wide variety of market inefficiencies—including random deviations of prices from fundamental values—and need not be caused by any systematic misinterpretation of earnings information. Note that even where evidence of overreaction is found (DeBondt and Thaler [1985, 1987] and Chopra, Lakonishok, and Ritter [1991]), there is no direct linkage between the anomalous returns and an overemphasis on earnings. In the two studies that did test for such a linkage (Klein [1990] and Zarowin [1989]), there was no support for the overreaction hypothesis. In fact, Klein concludes the evidence is more consistent with underreaction. Beyond this, while Abarbanell and Bernard [1991] confirm the DeBondt-Thaler [1990] finding of overly extreme analyst forecasts, they show that the phenomenon could not be due to analyst overreaction to recent earnings; with respect to earnings, analysts tend to underreact.

A second possibility is that both underreactions and overreactions to earnings occur. Specifically, stock prices could underreact, *on average*, to earnings, while overreactions occur only under conditions too complex to be captured by a simple partition on prior period earnings changes. As noted earlier, this possibility could explain why research designs like that of Bernard and Thomas [1989], which partition on the basis of prior earnings changes, indicate underreaction, while the designs of DeBondt and Thaler [1987] and Chopra, Lakonishok and Ritter [1991],
which partition on the basis of prior returns, still suggest overreaction. Such evidence is, however, also consistent with the first possibility raised above; the second possibility could be distinguished from the first only by identifying some connection between any existing overreactions and the interpretation of prior earnings numbers. For example, there might be conditions under which stock prices would fail to reflect that current earnings trends are likely to be reversed. Indeed, Ou and Penman [1989b] characterize their own evidence as indicating an underreaction to non-earnings accounting data that suggest the current year's earnings change is transitory. Another way to say the same thing is that overreactions to earnings occur when such earnings are predictably transitory, even though they do not occur on average.

A third possibility is that the market's response to earnings announcements defies a simple characterization as underreaction or overreaction. If prices reflect earnings expectations that are naive in the sense described by Bernard and Thomas [1990]—that is, based on a simple seasonal random walk—then post-announcement abnormal returns would continue to move in the same direction as the earnings news only for three quarters. The drift to this point would appear like a completion of an initial underreaction. After that point, as prices reflect the tendency for earnings to gravitate to a level less extreme than the originally announced earnings, a reversal of some the earlier abnormal returns would occur—thus taking on the appearance of a correction of an earlier overreaction. Such a pattern is evident in Figure 2, based on Bernard and Thomas [1989], and to some extent, in the year-by-year behavior of abnormal returns in Zarowin [1989, Table 3].

4. OTHER RECENT EVIDENCE ON EARNINGS-RELATED ANOMALIES: THE RESURRECTION OF THE FUNCTIONAL FIXATION HYPOTHESIS

Other recent anomalous results concerning stock price reactions to earnings cannot be classified neatly among the aforementioned studies. Hand [1990] resurrects and extends the functional fixation hypothesis in an examination of how stock prices reacted to earnings announcements for quarters in which firms had undertaken a debt-equity swap. Debt-equity swaps
were common during 1981-1984, because they could be undertaken with no tax consequence. Swaps gave rise to accounting gains that averaged 21 percent of quarterly earnings excluding the swap gain. Hand's context provides an interesting test of functional fixation because one can argue that, upon the announcement of the earnings that include the gain, an efficient market should not react to the gain. First, the gain is "artificial" in the sense that it represents a previously unrealized economic gain that occurred prior to the swap, when interest rates rose. Second, the amount of the gain was either disclosed publicly, or accurately calculable based on public information, prior to the earnings announcement. Thus, the earnings announcement date simply constitutes the "re-announcement" of the swap gain. Unless stock prices are fixated mechanically on earnings numbers, there should be no reaction to the re-announcement of the swap gain at the earnings announcement date.

Hand reports a statistically significant positive response to the swap gain, consistent with the functional fixation hypothesis. He also finds the positive response is larger for firms that are small, and/or have less institutional following, which he interprets as consistent with the notion that the result is driven by "unsophisticated" traders. The magnitude of the positive response is small, on average—less than 0.5 percent—but the result is nevertheless puzzling.

Ball and Kothari [1991] investigate yet another earnings-related anomaly, and offer an alternative interpretation of Hand's evidence. Ball and Kothari show that, even after controlling for the possibility of beta shifts and some other potentially confounding factors, the estimated abnormal returns in the days surrounding earnings announcements are on average positive.\textsuperscript{20} For large firms, the mean abnormal return is nearly zero, but for small firms, it exceeds 0.5 percent on the announcement day, and exceeds 1 percent for the three days surrounding the announcement.

Ball and Kothari also re-analyze Hand's data, and interpret his evidence as a manifestation of the more general phenomenon of positive mean announcement-period returns for small firms, as opposed to functional fixation. This interpretation leaves unanswered the question of why Hand

\textsuperscript{20} Chari, Jagannathan, and Ofer [1988] were the first to study this anomaly intensively, although they did not control for all the factors examined in Ball and Kothari. Evidence of positive mean announcement-period returns is also apparent in earlier papers, including Beaver [1968, Table 6].
finds a significant result for his sample only during the swap quarter; furthermore, in his response to Ball and Kothari, Hand [1991] discusses several additional features of the evidence that support the functional fixation hypothesis. Nevertheless, when Ball and Kothari combine their evidence with skepticism about several aspects of Hand's theory, they have "serious doubts" about Hand's attribution of his anomalous result to functional fixation.

There is at least some added support for functional fixation in other studies of reactions to earnings announcements. Dietrich [1984] finds an anomalous positive reaction to accounting gains resulting from bond-for-bond refinancing, even though those accounting gains may not correspond to "economic gains", and in any case should be known or accurately estimable from public information prior to the earnings release.

Givoly and Hayn [1991] find a statistically significant positive reaction to "paper gains" resulting from the change in deferred tax accounting required by Financial Accounting Standard No. 96. Consistent with Hand's theory, the positive reaction is driven by stocks for which institutional holding is relatively low and thus "unsophisticated" traders are presumably more prevalent. The result survives a control for size, suggesting the Ball-Kothari effect is not the underlying explanation. Nevertheless, Givoly and Hayn express some discomfort with the notion of functional fixation in competitive markets, and label their interpretation of the evidence "speculative."

5. CONCLUDING REMARKS

Nearly every major piece of evidence discussed in this review has been the subject of controversy, and several of the controversies remain the focus of ongoing research. Any conclusions can only be speculative, and must be drawn with the recognition that in the final analysis, one can never overcome the ultimate quandary in testing market efficiency. That is, any conclusions about market efficiency cannot be divorced from some assumed model of market equilibrium—the correctness of which is not only unknown, but unknowable (Ball [1990]).
One can attempt, however, to categorize the evidence in terms of how much it challenges the imagination to reconcile it with standard theories of asset pricing in competitive markets. In this sense, much of the evidence reviewed here presents quite a challenge.

Whether an incomplete initial response to earnings is the true cause of post-earnings announcement drift has long been debated. However, the recent evidence has thus far defied explanation in terms of potential research design flaws, including a failure to control fully for risk. The concentration of the abnormal returns around subsequent earnings announcements, in patterns that are predicted by a model in which stock prices reflect naive earnings expectations, is particularly difficult to explain. Brennan [1991] characterizes this evidence (along with that of Ou and Penman [1989]) as "perhaps the most severe challenge to financial theorists posed by the recent work on market reactions to earnings announcements." The ancillary evidence suggesting that even professional analysts at Value Line underreact to recent earnings news provides further support for the notion that stock prices may not be responding completely and immediately to earnings announcements.

The recent debate on the existence of overreaction in the stock market has been less one-sided, with several studies offering evidence both for and against the overreaction hypothesis. Ball and Kothari [1989] find no reliable indications of overreaction. On the other hand, when Chopra, Lakonishok, and Ritter [1991] relax an assumption within the Ball-Kothari analysis that is rejected by the data, they produce intriguing support for overreaction.

Evidence of underreactions to earnings and overreactions among extreme stock price movements is not necessarily inconsistent, as it may at first appear. After all, even though some evidence suggests overreactions do occur, no study to date has yet linked the overreaction directly to an overemphasis on recent earnings changes. A possibility consistent with all the data is mispricing due to a variety of unidentified causes, which is most likely to surface among stocks with extreme price changes and be labeled an overreaction, in combination with incomplete initial responses to fundamentals like earnings.
When viewed as a whole, the evidence summarized here presents a more serious challenge to the efficient markets hypothesis than could have been anticipated a few years ago. It is difficult to understand how the simple trading rules described here could apparently generate positive abnormal returns. Only time will tell whether attempts to resolve the anomalies within the existing paradigm will prove successful. In the meantime, we should remain open to unconventional approaches to understanding how prices might deviate from fundamental values in what appear to be extremely competitive markets.
Table 1

Time series behavior of quarterly earnings

Panel A: Autocorrelations in seasonally differenced quarterly earnings, 1946-1974 (from Foster [1977], based on 69 firms):

<table>
<thead>
<tr>
<th>Lag</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>.45</td>
<td>.24</td>
<td>.13</td>
<td>-.12</td>
<td>.01</td>
<td>.02</td>
<td>-.02</td>
<td>-.03</td>
</tr>
</tbody>
</table>

Panel B: Autocorrelations in seasonally differenced quarterly earnings, 1962-1978 (from sample used by Hopwood and McKeown [1986], based on 267 firms*):

<table>
<thead>
<tr>
<th>Lag</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (1962-1978)</td>
<td>.45</td>
<td>.21</td>
<td>.02</td>
<td>-.19</td>
<td>-.07</td>
<td>-.02</td>
<td>-.02</td>
<td>-.03</td>
</tr>
<tr>
<td>Mean (1962-1970II)</td>
<td>.36</td>
<td>.17</td>
<td>.05</td>
<td>-.18</td>
<td>-.07</td>
<td>-.05</td>
<td>-.07</td>
<td>-.07</td>
</tr>
<tr>
<td>Mean (1970III-1978)</td>
<td>.42</td>
<td>.17</td>
<td>-.03</td>
<td>-.22</td>
<td>-.11</td>
<td>-.05</td>
<td>-.06</td>
<td>-.06</td>
</tr>
</tbody>
</table>

Panel C: Autocorrelations in seasonally differenced quarterly earnings, 1974-1986 (from Bernard and Thomas [1989], based on 2,626 firms):

<table>
<thead>
<tr>
<th>Lag</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>.34</td>
<td>.19</td>
<td>.06</td>
<td>-.24</td>
<td>-.08</td>
<td>-.07</td>
<td>-.07</td>
<td>-.06</td>
</tr>
<tr>
<td>25th percentile</td>
<td>.14</td>
<td>.05</td>
<td>-.10</td>
<td>-.46</td>
<td>-.26</td>
<td>-.24</td>
<td>-.24</td>
<td>-.25</td>
</tr>
<tr>
<td>Median</td>
<td>.36</td>
<td>.18</td>
<td>.06</td>
<td>-.29</td>
<td>-.09</td>
<td>-.08</td>
<td>-.06</td>
<td>-.06</td>
</tr>
<tr>
<td>75th percentile</td>
<td>.57</td>
<td>.35</td>
<td>.21</td>
<td>-.07</td>
<td>.08</td>
<td>.08</td>
<td>.09</td>
<td>.11</td>
</tr>
</tbody>
</table>

*These statistics are not reported in Hopwood and McKeown [1986], but were furnished to me in personal correspondence with Jim McKeown.
Table 2

Summary of evidence from DeBondt-Thaler [1987] and Ou-Penman [1989b]

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>Basis for construction</th>
<th>Median change in earnings-per-share</th>
<th>Market-adjusted returns</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Portfolio formation period Test period</td>
<td>Last year of formation period</td>
</tr>
<tr>
<td>Losers</td>
<td>CAR over prior four years in lowest quintile(^a)</td>
<td>-49% over 3 yrs b 27% over 4 yrs</td>
<td>-12% 6%</td>
</tr>
<tr>
<td>Winners</td>
<td>CAR over prior four years in highest quintile</td>
<td>20% over 3 yrs b -5% over 4 yrs</td>
<td>8% -3%</td>
</tr>
</tbody>
</table>

Panel A: DeBondt and Thaler [1987]
(4-year test periods range from 1970-1973 to 1980-1983)

Panel B: Ou and Penman [1989b]

| High Pr   | Probability of earnings increase in highest decile\(^c\) | -168% over 1 yr not reported over 2 yrs | -168% 81% | -42% over 1 yr 6% over 2 yrs |
| Low Pr    | Probability of earnings decrease in lowest decile\(^c\) | 66% over 1 yr not reported over 2 yrs | 66% -11% | 37% over 1 yr -22% over 2 yrs |

Sources: DeBondt and Thaler [1987, Tables V and VIII]; Ou and Penman [1989, Tables 6 and 7].

\(^a\) CAR is a cumulative abnormal (market-adjusted) return, measured over four-year period ending on last trading day of a given year. The test period begins on following trading day.

\(^b\) Although the portfolio formation period covers four years, data are reported only for last three years.

\(^c\) Pr is an estimate of the probability of an annual earnings increase in coming year, based on a function of financial statement variables identified and estimated using only historical data. The design assumes that financial data are available to calculate Pr within three months of the end of the fiscal year; the test period begins on first day of fourth month following fiscal year end.
Figure 1

Cumulative abnormal returns (CAR) for SUE portfolios
84,792 earnings announcements, 1974 - 1986

Earnings announcements are assigned to deciles based on standing of standardized unexpected earnings (SUE) relative to prior quarter SUE distribution. SUE represents forecast errors from the Foster [1977] first-order autoregressive earnings expectation model (in seasonal differences) scaled by their estimation-period standard deviation. CARs are the sums over pre and post-announcement holding periods (beginning day -59 and day 1, respectively) of the difference between daily returns and returns for NYSE-AMEX firms of the same size decile, based on January 1 market values of equity.
Figure 2

Post-earnings-announcement drift:
By firm size, for 360 trading days beyond announcement

Size-adjusted post-announcement return to long (short) position in top (bottom) SUE decile

Small firms

Medium firms

Large firms

Trading days after earnings announcement

Source of data: Bernard and Thomas [1989, Table 1]. Earnings announcements are assigned to deciles based on standing of standardized unexpected earnings (SUE) relative to prior quarter SUE distribution. (84,792 announcements, 1974 - 1986.) SUE represents forecast errors from the Foster [1977] first-order autoregressive earnings expectation model (in seasonal differences) scaled by their estimation-period standard deviation. Size-adjusted returns are the sums over post-announcement holding periods of the difference between daily returns and returns for NYSE-AMEX firms of the same size decile, based on January 1 market values of equity. Small, medium, and large firms are in size deciles 1 to 4, 5, to 7, and 8 to 10, respectively, based on January 1 market value of equity for all NYSE and AMEX firms.
Figure 3

Returns to SUE strategy by year, 1965-1986:
A combination of results reported in three studies

<table>
<thead>
<tr>
<th>Annual Return</th>
<th>Rendleman, Jones, and Latane 1972-1979</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Latane, Jones, and Rieke 1965-1971</td>
</tr>
<tr>
<td>0.50</td>
<td>(not true calendar years)</td>
</tr>
<tr>
<td>0.45</td>
<td>Bernard and Thomas 1974-1986</td>
</tr>
<tr>
<td>0.40</td>
<td>(true calendar years)</td>
</tr>
<tr>
<td>0.35</td>
<td></td>
</tr>
<tr>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>0.20</td>
<td></td>
</tr>
<tr>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>0.00</td>
<td>'65 '66 '67 '68 '69 '70 '71 '72 '73 '74 '75 '76 '77 '78 '79 '80 '81 '82 '83 '84 '85 '86</td>
</tr>
</tbody>
</table>

Latane, Jones, and Rieke [1974]: Long (short) positions in stocks with SUE greater than (less than) 1.5 are assumed two months after end of fiscal quarter, and held for six months. Returns are not adjusted for risk. Estimated holding period betas are 1.14 for long position and 1.09 for short position. Returns shown above are calculated first by annualizing six-month returns taken after each of four quarters. Return for 1971 is based on sum of only two six-month positions.

Rendleman, Jones, and Latane [1982]: Long (short) positions are assumed in 20 stocks with highest (lowest) SUEs among universe that grows from 170 to 972. Universe includes only stocks announcing earnings within one month of fiscal quarter close; positions are taken at the beginning of the next month and held for three months. Portfolios are balanced to assure pre-holding-period beta of zero, but no control for holding period beta. Returns from four three-month holding periods are summed to arrive at annual return.

Bernard and Thomas [1989]: Long (short) positions are assumed in stocks among the top (bottom) quintile of SUE, relative to prior quarter distribution. Each $1 long position is always offset by a $1 short position in stock(s) of similar size (small, medium, or large). Balancing in this way sometimes requires waiting after earnings announcements until an offsetting match is available. Returns are not adjusted for risk, but controls for holding-period beta and five other risk factors have little impact on the returns. Returns reported for calendar quarters and summed to arrive at annual return. Return for 1986 is based on only nine months.
Figure 4
Cumulative abnormal returns for SUE portfolios:
Returns aligned by subsequent earnings announcements

Portfolio CAR is the cumulative abnormal return over holding periods beginning after the earnings announcement day for quarter t, for a portfolio invested long (short) in the highest (lowest) decile of standardized unexpected earnings (SUE) at quarter t. SUE represents forecast error from the seasonal random walk (with trend) earnings expectation model scaled by its estimation-period standard deviation. Abnormal returns are the differences between daily returns for individual firms in an SUE decile portfolio, and returns for NYSE-AMEX firms of the same size decile, based on January 1 market values of equity. Small, medium and large firms are in size deciles 1 to 4, 5 to 7, and 8 to 10, respectively. Holding periods are obtained by splitting the period between adjacent earnings announcement dates into a 3-day pre-announcement window (day -2 to day 0) and an inter-announcement window. While the actual inter-announcement windows vary in length, the mean value of 60 days is used to illustrate the differential price responses occurring in the two windows.
REFERENCES


