THE JOURNAL OF FINANCE • VOL. XLIV, NO. 5 • DECEMBER 1989

Does the Stock Market Overreact to Corporate Earnings Information?

PAUL ZAROWIN*

ABSTRACT

This paper tests whether the stock market overreacts to extreme earnings, by examining firms' stock returns over the 36 months subsequent to extreme earnings years. While the poorest earners do outperform the best earners, the poorest earners are also significantly smaller than the best earners. When poor earners are matched with good earners of equal size, there is little evidence of differential performance. This suggests that size, and not investor overreaction to earnings, is responsible for the "overreaction" phenomenon, the tendency for prior period losers to outperform prior period winners in the subsequent period.

RECENTLY, AN INTRIGUING EFFICIENT markets anomaly has emerged. DeBondt and Thaler (1985, 1987) report that equities that experience the lowest (highest) returns over a prior period (2–5 years in their studies), overperform (underperform) the market in the subsequent period. They call this violation of the efficient markets hypothesis the "overreaction" phenomenon since it suggests that investors overreact in the initial period and subsequently correct themselves.[1]

DeBondt and Thaler (1987) hypothesize that the reason for the overreaction lies in the market's inefficient response to earnings information. Research has found that the random walk model, in which all changes are permanent, provides a good description of the time-series behavior of firms' annual earnings, *on average*.[2] The primary exception to the random walk model comes from firms that have experienced extreme good or bad years, relative to their "normal" performance. The earnings changes of such "outlier" firms are temporary; i.e., earnings regress to the mean.[3] If market participants incorrectly perceive the

* Leonard N. Stern Graduate School of Business, New York University. I want to thank Ray Ball, Victor Bernard, John Bildersee, Stephen Brown, Colin Camerer, Martin Gruber, Cole Kendall, Gary King, April Klein, S. P. Kothari, Josef Lakonishok, Joshua Livnat, Mike Long, Wayne Mikkelson, Paul Schultz, Keith Weigelt, an anonymous reviewer, and the co-editor, David Mayers, for their helpful comments. The financial support of the Peat Marwick Foundation is gratefully acknowledged. All errors remain my responsibility.

[1] As support for the overreaction hypothesis, DeBondt and Thaler cite evidence from research in experimental psychology which suggests that, in violation of Bayes' rule, people tend to overweight recent information and to underweight prior data. (See Kahneman and Tversky (1982)).

[2] Annual earnings appear to behave as a random walk if the information set is restricted to be the time series of past earnings. In the presence of other publicly available information, the best prediction of this year's earnings is not necessarily last year's. For a discussion, see Griffin (1982, Chapter 5).

[3] Brooks and Buckmaster (1976, 1980) document the mean reversion behavior of earnings at the tails. See Foster (1986) for a review of the literature on the time-series behavior of corporate earnings and on the stock market's response to earnings news.

extreme earnings changes of such firms to be permanent, they would overreact by bidding the stock prices of the good (bad) performers up (down) too high (low). When subsequent earnings realizations are not extreme, the market learns of its mistake and stock prices correct.

To support the "overreaction to earnings" hypothesis, DeBondt and Thaler (1987, Table VIII) show that the *subsequent* earnings changes of extreme *prior period* stock return performers show a reversal pattern. However, the fundamental testable implication of the overreaction to earnings hypothesis concerns stock returns *subsequent* to earnings realizations, *not prior* to them; i.e., the basis of a trading rule is returns *following*, not *preceding*, an observed event. Thus, their findings are consistent with, but not evidence of, an efficient markets anomaly due to earnings myopia.

The research presented here examines the "overreaction to earnings" phenomenon by testing a trading rule based on firms that have experienced extremely good and bad earnings years. Consistent with the overreaction hypothesis, stock returns of the poorest earners outperform those of the best earners over the 36 months subsequent to the extreme earnings year. The poorest earners, however, are significantly smaller firms than the best earners at the end of the "outlier" year. When both groups are matched by size, the return discrepancy disappears. This suggests that the overreaction to earnings phenomenon is another manifestation of the size phenomenon. Although we still do not completely understand the size phenomenon, our findings are important because they show that the stock market is not characterized by inefficiency due to earnings myopia. Furthermore, while stock market overreaction is an efficient markets anomaly, the size phenomenon is more likely a CAPM anomaly. (See Amihud and Mendelson (1986).)[4]

Chan (1988) explains the winner-loser effect as an artifact of beta nonstationarity since the standard form of the overreaction hypothesis assumes risk to be time invariant. Fama and French (1986) also find that the size effect may largely subsume the return reversal effect. Thus, our results are consistent with Chan's and Fama and French's, providing additional evidence against the overreaction hypothesis.

The paper is organized as follows. Section I describes the data, the sample selection procedure, and our measure of earnings performance. Section II discusses initial evidence that provides prima facie support for the overreaction to earnings hypothesis, while Section III provides evidence on risk differences between extreme (good vs. bad) earners. Section IV discusses the results for portfolios that are matched by size but differ in earnings performance and for portfolios that are matched by earnings performance but differ in size. Section V concludes the paper with a summary of findings and implications.

I. Data Description and Sample Selection

The basic strategy is to form portfolios of firms that are characterized by extreme (good vs. bad) current period earnings performance and to compare the subse-

[4] See Schwert (1983) for a review of the size anomaly evidence in finance.

Table I

Size and Return Characteristics of the Sample

Number of firms (N) in the entire sample for each year, their mean market value of equity at the end of the year in 000's (MV), and their mean monthly stock return during the year (RET). Mean market value of equity of all firms on the CRSP monthly file at the end of the year in 000's ($CRSPMV$) and the mean monthly return of the CRSP equally weighted index during the year ($CRSPRET$), 1971–1981.

Year	*N*	*MV*	*RET*	*CRSPMV*	*CRSPRET*
1971	551	867,390	0.015	505,894	0.016
1972	579	987,388	0.009	567,508	0.007
1973	597	806,410	−0.019	443,836	−0.026
1974	611	558,596	−0.019	311,936	−0.023
1975	629	728,515	0.038	424,601	0.044
1976	650	889,440	0.032	526,196	0.033
1977	669	813,076	0.004	493,227	0.008
1978	696	820,300	0.010	508,964	0.013
1979	721	907,300	0.024	600,435	0.027
1980	756	1,123,685	0.023	780,403	0.025
1981	788	1,020,537	0.004	726,896	0.006

quent stock returns of the poorest earners versus the best earners. The CRSP monthly return file and the Compustat Annual Industrial file are the databases for this study.[5]

Each year, from 1971 to 1981, all firms meeting the following data requirements are included in the sample for that year:

1) availability of the six consecutive prior years and the current year of earnings before extraordinary items and discontinued operations (data item No. 18) on Compustat,
2) December 31 fiscal year end, and
3) availability of price per share and number of shares outstanding on the CRSP monthly file at year end.

Table I shows the number of firms that satisfy the data requirements each year, along with their mean size at year end and their average monthly stock return during the year. Also shown are the average sizes and monthly returns for

[5] Although this study does not use the Compustat Research File, McElreath and Wiggins (1984) argue that delistings are probably too few in number to cause a significant survivorship bias. For example, they report that 16 firms were deleted from the NYSE due to bankruptcy from 1970–1979, with half of these years experiencing no bankruptcy delistings. This is an extremely small percentage of the number of firms used for this study. See Table I.

Since delistings might be due to either bad news (e.g,, bankruptcy) or good news (e.g., merger), it is not obvious in which direction such a survivorship bias might be. Pastena and Ruland (1986) note that many takeover targets are firms that might otherwise have gone bankrupt. Thus, poor performers may be more prevalent than good performers among delisted firms. In that case, excluding such firms would bias our tests in favor of the overreaction hypothesis since poor performers who went bankrupt (or might have, but got taken over) are delisted but those who improved remain.

all CRSP firms in each year. The average firm used in this study is somewhat larger than the average CRSP firm but has similar average return behavior.

For each firm in the sample, an earnings performance measure is calculated:

$$PERF_{iT} = \frac{\Delta x_{iT}}{\sigma_{\Delta x_i}}, \tag{1}$$

where

x_T = earnings before extraordinary items and discontinued operations (i.e., data item No. 18) for year T,

$\Delta x_T = x_T - x_{T-1}$ = the change in earnings from last year to this year, and

$\sigma_{\Delta x}$ = the standard deviation of the firm's earnings changes over the five previous years.

$PERF_{iT}$, similar to the measure used by Brooks and Buckmaster (1976), is an appealing earnings performance metric for this study because it captures to what extent a firm's earnings change is an "outlier" observation.[6] Since, for most firms, annual earnings are a nonstationary variable, computing $PERF_{iT}$ based on prior period *changes in*, rather than *levels of*, earnings is appropriate. The choice of a five-year period to calculate the standard deviation represents a tradeoff between having enough observations to calculate this statistic versus having enough to conduct the tests (since a longer estimation period reduces the number of available test periods). Furthermore, since firms' earnings processes undergo structural changes (see Foster (1986)), using a longer estimation period may introduce error.

Earnings before extraordinary items and discontinued operations (rather than, for example, net income) is the appropriate income variable for this study because we want to examine the market's ability to distinguish between permanent and temporary earnings shocks. Since extraordinary items are, at least in theory, temporary (and are segregated from other earnings components on the income statement), including them would not provide a good test of the market's efficiency in this area.[7] The fiscal year requirement is imposed in order to make our results comparable to DeBondt and Thaler's, whose excess return accumulation period begins in January,[8] and the CRSP shares outstanding and price per

[6] The tests reported in this paper were also conducted with a measure of earnings performance defined as ΔX_T divided by the beginning of year T market value of equity. The results of these tests were consistent with those reported in the paper and are available from the author on request.

[7] DeBondt and Thaler also used earnings before extraordinary items.

[8] Although a firm's annual earnings are not known as of January 1, Ball and Brown (1968) showed that a majority of the security revaluation process due to those earnings has occurred by that date. Based on the first three quarterly reports and supplemental information, the market can distinguish the extremely good from the extremely poor earnings performers as of December 31 (i.e., the end of the extreme performance year).

If there is any bias in our results from beginning the test period on January 1 (as DeBondt and Thaler do), it is likely to be extremely small and *in favor of the overreaction hypothesis* because prior period losers outperform prior period winners in January. As shown later, our results do not support the overreaction to earnings hypothesis in spite of this.

share data are necessary to measure firms' sizes at the beginning of the portfolio formation period. The eleven-year sample period is determined by the Compustat file used for this study, which covers the twenty years 1965–1984. Since we require six years of earnings to compute $\sigma_{\Delta x}$, 1971 is the first year for which we could compute $PERF_{iT}$. Our test period ends in 1981, in order to have three subsequent years to observe mean reversion behavior in earnings.

Each year of the test period, firms are ranked by $PERF_{iT}$ and sorted into five portfolios, with an equal number of firms per portfolio. (DeBondt and Thaler (1987, Section II) use quintiles.) Portfolio 1 includes the firms with the worst *PERF* measures, and portfolio 5 includes the firms with the best earnings performance. This study concentrates on the "extreme" portfolios, 1 and 5.[9]

During the extreme-earnings year, the market's best earners (the top quintile of firms ranked by *PERF*) outperform the market's worst earners (the bottom quintile) by an average cumulative return difference of 29.6% ($t = 9.1$). According to the efficient markets hypothesis, there should be no subsequent predictable pattern in the security returns of the two portfolios. If, however, these returns represent an overreaction to earnings information, we expect reversals in the subsequent performances of the two portfolios; i.e., the poorest earners should outperform the best earners.[10]

Table II shows the average firm size (market value of equity) for the extreme earner portfolios for each year of the sample period. For all years, the poor earners are significantly smaller than the best earners at the date of portfolio formation.[11] This is not surprising since the poor earners have experienced large market value declines relative to the best earners. The point here is that, if we find that poor earners outperform good earners subsequent to the extreme earnings year, this could be due to either stock market overreaction or the size effect.

The last two columns of Table II show the average market values of equity for extreme size quintile portfolios. Comparing columns 2 and 3 with columns 5 and 6 shows that *the loser and winner firms in this sample are neither extremely small nor large.* This is what DeBondt and Thaler (1987, Table V, Panels A and B) find, and this evidence leads them to conclude (1987, p. 571), "The winner-loser anomaly cannot be accurately described as a small firm phenomenon." However, the t-test results show that it may be a mistake to dismiss the size effect as an explanation for the winner-loser effect because, in fact, losers *are* significantly smaller than winners.

[9] Consistent with the results of Brooks and Buckmaster, the earnings of the extreme performers in our sample exhibit regression to the mean behavior, with the mean reversion of losers more pronounced than that of winners.

[10] The post-earnings announcement drift anomaly, wherein firms with positive (negative) unexpected earnings in year T experience positive (negative) abnormal returns in year $T + 1$, has implications opposite to those of the overreaction hypothesis. See Foster (1986) for a description of this anomaly.

[11] DeBondt and Thaler (1987, Table V, Panel A) also show that loser firms are, on average, smaller than winner firms at the time of portfolio formation; however, they perform no tests for differences.

Table II

Size Characteristics of Extreme Quintile Portfolios

Average market values of equity in 000's at the end of the year of the bottom quintile of firms ranked by the earnings performance measure (Size of Loser Quintile), of the top quintile of firms ranked by the earnings performance measure (Size of Winner Quintile), of the bottom quintile of firms ranked by market value of equity (Size of Smallest Quintile), and of the top quintile of firms ranked by market value of equity (Size of Largest Quintile), and t-statistic (t) to test the null hypothesis that the mean market value of equity of firms in the loser quintile equals the mean market value of equity of firms in the winner quintile, 1971–1981. The earnings performance measure is defined as the change in earnings (before extraordinary items and discontinued operations) from the previous year to the current year, divided by the standard deviation of annual earnings changes over the five previous years.[a]

Year	Size of Loser Quintile	Size of Winner Quintile	t	Size of Smallest Quintile	Size of Largest Quintile
1971	378,873	1,331,679	−2.63*	40,461	3,282,803
1972	587,203	2,033,326	−7.67*	37,433	3,873,451
1973	210,771	2,155,007	−9.96*	22,429	3,253,588
1974	307,162	795,081	−7.55*	15,102	2,241,929
1975	620,028	702,758	−4.23*	21,769	2,868,560
1976	542,075	1,440,121	−4.26*	31,731	3,416,096
1977	548,069	1,447,715	−6.35*	34,631	3,043,911
1978	569,828	1,249,678	−5.79*	34,477	3,073,720
1979	724,473	1,418,232	−9.07*	40,505	3,331,764
1980	618,477	1,557,780	−4.77*	46,225	4,179,061
1981	885,403	1,299,604	−4.98*	45,766	3,706,887

[a] The t-tests are performed on the log of market value of equity, because market value of equity is (approximately) lognormally distributed.

* Statistically significant at $\alpha = 0.01$ level.

II. The Overreaction to Earnings Hypothesis: Initial Tests

A. Methodology

To examine whether the stock market overreacts to extreme earnings news, we compare the excess returns of the two extreme earnings portfolios over the 36 months subsequent to the extreme earnings year.

For each month we calculate a firm's excess return as

$$XS_{it} = R_{it} - R_{mt}, \tag{2}$$

where

R_{it} = the return on firm i for month t and
R_{mt} = the return on the CRSP equally weighted index for month t.[12]

[12] We use market adjusted excess returns calculated with respect to the CRSP equally weighted index, as do DeBondt and Thaler.

We calculate the average excess return and cumulative average excess return for the firms in portfolio j ($j = 1$ or 5) for month t as

$$AR_{mjt} = \frac{\sum_{i=1}^{n_{jt}} XS_{it}}{n_{jt}} \tag{3}$$

and

$$CAR_{mj\tau} = \sum_{t=1}^{36} AR_{mjt}, \tag{4}$$

where n_{jt} is the number of observations in portfolio j for month t, $m = 1$ to 11 for the eleven years 1971–1981, and $t = 1$ is January after the extreme earnings year.

Our results are based on the averages of the $CAR_{mj\tau}$ over the eleven years of the sample period, i.e.,

$$\overline{CAR}_{j\tau} = \frac{\sum_{m=1}^{11} CAR_{mj\tau}}{11}. \tag{5}$$

The 36-month test period is chosen to make our analysis comparable to DeBondt and Thaler's, who find that the stock market's correction occurs over the 36 months subsequent to the period of extreme performance.[13] We compare the excess returns of the two portfolios by conducting a two-sample t-test.

B. Empirical Evidence

On average, over the 11-year sample period, the poorest earner portfolio outperforms the best earner portfolio by a statistically significant cumulative excess return difference of 16.6% ($t = 2.9$) over the 36 months. DeBondt and Thaler (1985) find a return difference of 24.6% over a three-year period. That our results are less extreme is not surprising since they include only 35 stocks in each portfolio, whereas we include an average of over 100 (see Table I), allowing for less extreme prior period performers. Also, we rank into portfolios by *PERF*, whereas they rank by prior period return. Nevertheless, our results conform to theirs, and it appears that the stock market overreacts to extreme earnings news.[14]

Aside from stock market overreaction, DeBondt and Thaler discuss two alternative explanations for their results: differences in risk (beta) and differences in size between loser and winner firms. Since the excess returns analyzed here are market adjusted, but not risk adjusted, poor earners being riskier than good earners could account for the superior performance. As shown in Table II, poor earners are significantly smaller than good earners at the portfolio formation date. Since smaller firms are known to outperform larger firms, this could also

[13] Although DeBondt and Thaler (1985) provide results for other return intervals, their main results (e.g., Section I and Figure 1) are for 36 months. Their conclusions are insensitive to the length of the return interval, beyond three years.

[14] Also similar to DeBondt and Thaler, we find a January effect, as the three Januaries of the test period have statistically significant excess return differences of 4.5%, 5.7%, and 3.2%, respectively.

Table III
Regression Results for Arbitrage Portfolios

Results of regressions of monthly arbitrage portfolio returns (R_{At}) against the market risk premium ($R_{mt} - R_{ft}$), for each of the three years of the test period, January 1972–December 1984. The return on an arbitrage portfolio for month t, $R_{At} = R_{Lt} - R_{Wt}$, where R_{Lt} and R_{Wt} are the returns on the bottom and top quintiles of firms, ranked by the earnings performance measure. R_{mt} = the return on the CRSP equally weighted index for month t. R_{ft} = the risk-free (1-month Treasury bill) rate at the beginning of month t. Year 1 is months 1–12 after the ranking period. Year 2 is months 13–24 after the ranking period. Year 3 is months 25–36 after the ranking period. The earnings performance measure is defined as the change in earnings (before extraordinary items and discontinued operations) from the previous year to the current year, divided by the standard deviation of annual earnings changes over the five previous years. Smallest and largest firms are the bottom and top quintiles ranked by market value of equity at the portfolio formation date.[a]

$$R_{At} = \alpha_A + \beta_A(R_{mt} - R_{ft}) + \epsilon_{At}$$

	α	β	$\bar{R}^2$	NOBS
Panel A—($P_1 - P_5$): All Losers Minus All Winners				
Year 1	−0.003	0.261*	0.24	132
	(−1.04)	(6.49)		
Year 2	0.006*	0.227*	0.23	132
	(2.73)	(6.38)		
Year 3	0.005*	0.184*	0.20	132
	(2.80)	(5.89)		
Panel B—($P_{11} - P_{51}$): Smallest Losers Minus Smallest Winnners				
Year 1	−0.010*	0.222*	0.11	132
	(−3.03)	(4.21)		
Year 2	0.002	0.192*	0.07	132
	(0.42)	(3.25)		
Year 3	0.002	0.193*	0.08	132
	(0.65)	(3.46)		
Panel C—($P_{11} - P_{15}$): Smallest Losers Minus Largest Losers				
Year 1	−0.002	0.510*	0.36	132
	(−0.50)	(8.61)		
Year 2	0.007**	0.450*	0.28	132
	(1.86)	(7.15)		
Year 3	0.009*	0.388*	0.21	132
	(2.23)	(5.96)		
Panel D—($P_{15} - P_{55}$): Largest Losers Minus Largest Winners				
Year 1	0.001	0.107*	0.02	132
	(0.46)	(2.09)		
Year 2	0.004	0.132*	0.05	132
	(1.34)	(2.87)		
Year 3	0.002	0.126*	0.04	132
	(0.73)	(2.66)		
Panel E—($P_{51} - P_{55}$): Smallest Winners Minus Largest Winners				
Year 1	0.010*	0.395*	0.20	132
	(2.24)	(5.73)		
Year 2	0.010*	0.390*	0.20	132
	(2.32)	(5.86)		
Year 3	0.009*	0.321*	0.13	132
	(1.99)	(4.47)		

Table III—*Continued*

	$R_{At} = \alpha_A + \beta_A(R_{mt} - R_{ft}) + \epsilon_{At}$			
	α	β	$\bar{R}^2$	*NOBS*
Panel F—($P_{15} - P_{51}$): Largest Losers Minus Smallest Winners				
Year 1	−0.008*	−0.288*	0.12	132
	(−1.95)	(−4.30)		
Year 2	−0.006	−0.258*	0.10	132
	(−1.43)	(−3.97)		
Year 3	−0.007**	−0.195*	0.05	132
	(−1.63)	(−2.93)		

[a] t-statistics are in parentheses. *NOBS* = number of observations used to estimate the regression.

* Coefficient is significantly different from zero at $\alpha = 0.05$ level (two-tailed test).

** Coefficient is significantly different from zero at $\alpha = 0.10$ level (two-tailed test).

account for the superior performance documented by us and by DeBondt and Thaler.

III. Examining Risk Differences Between Extreme Earner Portfolios

As DeBondt and Thaler (1987) describe, standard procedures, based on prior period returns, to test for beta differences between winner and loser firms are inappropriate for this study due to equity beta shifts that accompany extreme changes in leverage. Thus, to test for risk differences, we follow their procedure of regressing the return on an arbitrage portfolio against the market risk premium, i.e.;

$$R_{At} = \alpha_A + \beta_A(R_{mt} - R_{ft}) + \epsilon_{At}, \quad (6)$$

where $R_{At} = R_{Lt} - R_{Wt}$, where R_{Lt} and R_{Wt} are the returns on the loser and winner portfolios (ranked by *PERF*) for month t, R_{ft} = the risk-free (one-month Treasury bill) rate set at the beginning of month t, and all other variables are as before.[15]

The intercept, α_A, from this regression is the Jensen performance index; the slope coefficient, β_A, is an estimate of the difference in CAPM betas between the two portfolios over the test period. Since each portfolio formation period overlaps 24 months with the two adjacent periods, we estimate (6) by running three separate regressions for months 1–12, 13–24, and 25–36 of the test period, respectively. Within each regression, the observations are nonoverlapping, and this enables us to determine which of the three years of the test period drive(s) the results.

The results of this analysis, shown in Panel A of Table III, indicate that, while the poorest earner portfolio is significantly riskier than the best earner portfolio, the poorest earners still outperform the best earners at annual rates of about 7.4% and 6.2% ($1.006^{12} - 1 = 0.074$ and $1.005^{12} - 1 = 0.062$), controlling for risk differences, in years 2 and 3 of the test period, respectively. Thus, while risk can

[15] Treasury bill rates are taken from Ibbotson and Sinquefield (1987).

account for some of the superior performance of the poor earner portfolio, most of the return discrepancy is still left unexplained since the cumulative return differences (not shown in the table) are 9.4% and 8.5% in years 2 and 3.

Chan (1988) suggests that, due to potential time-varying risk of the arbitrage strategy, running equation (6) over the full sample period may produce a biased estimate of the strategy's abnormal performance. When we follow Chan's procedure of running a separate regression for each year (each time the portfolios are updated), most of the return discrepancy over the three-year test period is eliminated, but our finding of abnormal performance in years 2 and 3 is not entirely erased.[16]

IV. Results for Size-Matched and *PERF*-Matched Pairs

A. Test Procedures

To ascertain whether size differences or stock market overreaction to earnings are driving the return differences, we compare matched pairs of firms. First, we compare firms that are alike in size but are characterized by extreme differences in earnings performance. Then, we compare firms that have similar earnings performance but dissimilar sizes. If stock market overreaction is the cause, we expect significant return differences between portfolios matched by size but characterized by disparate earnings performance, and insignificant return differences between portfolios matched by earnings performance but contrasting in size. If differential size is the cause, we expect the opposite.

To compare the subsequent return performances of matched pairs, we proceed as follows. At the beginning of each portfolio formation period (January 1 after the extreme earnings year), the firms in the sample are reranked by size (inde-

[16] Our results with Chan's technique are:

	Year 1	Year 2	Year 3
$\bar{\alpha}$	−0.0048	0.0043	0.0045
u	−2.42	1.97	2.36

where

$$\bar{\alpha} = \frac{\sum_{i=1}^{11} \alpha_i}{11}$$

and

$$u = \frac{1}{\sqrt{11}} \sum_{i=1}^{11} t_i \sqrt{(T_i - 3)/(T_i - 1)},$$

where α_i is the Jensen performance index, $t_i = t(\hat{\alpha}_i)$, and T_i = the number of observations (12) in each test period. According to Chan, u approaches a standard normal distribution as the number of test periods gets large. When averaged over the three years, the annual excess return is about 1.6%, which is similar to what Chan finds (1988, Table IV).

pendently of their previous performance ranking) and again formed equally into five portfolios. Each firm in the sample, therefore, is characterized by a performance/size combination, and there are 25 such combinations (portfolios). Each portfolio is indexed by $P_{i,j}$, where the i index refers to earnings performance and the j index refers to size. As before, $i = 1$ refers to the poorest earnings performers and $i = 5$ to the best. Likewise, $j = 1$ refers to the smallest firms and $j = 5$ to the largest firms.

Table IV shows statistics for the average number and size of the firms in each portfolio, as well as annual t-statistics to test the null hypotheses that small (large) poor performers are equal in size to small (large) good performers. Although the comparison of P_{11} vs. P_{51} in Panel B shows that our size control is not perfect, the size difference is immaterial, as shown in Panel A (30868 vs. 37867). Furthermore, if there is any bias from our imperfect size control, it favors the finding of significant positive abnormal performance in the P_{11}-P_{51} comparison since small losers are slightly smaller than small winners. In spite of this, as shown below, we find no evidence of overreaction to earnings in these portfolios, either in the Jensen performance tests or in the 36-month cumulative excess returns.

It is interesting to note that, in the extreme size portfolios ($j = 1$ and $j = 5$), there is a large discrepancy in the number of poor earners versus good earners. This, of course, is not surprising since size is measured at the end of the extreme performance year. As noted above, poor earners experienced large market value declines relative to good earners over this year.

B. Jensen Performance Measures for Matched Portfolios

Panels B–F of Table III examine the Jensen performance measures and test period risk differences for five groups of size-matched and *PERF*-matched arbitrage portfolios. Panels B and C compare the smallest poor earner portfolio (P_{11}) to the smallest good earner portfolio (P_{51}) and to the largest poor earner portfolio (P_{15}), respectively. Panels D and E compare the largest best earner portfolio (P_{55}) to the largest worst earner portfolio (P_{15}) and to the smallest best earner portfolio (P_{51}), respectively. Finally, Panel F compares portfolios that are contrasted by both earnings performance and size. Comparing large losers and small winners (P_{15} and P_{51}) pits the size and overreaction effects directly against one another since the two phenomena make opposite predictions as to which portfolio outperforms the other.

The results in Panels B–F of Table III are consistent with the size phenomenon and inconsistent with the overreaction to earnings phenomenon. For the size-matched portfolios in Panels B and D, losers do not significantly outperform winners in any of the three years. For the *PERF*-matched portfolios in Panels C and E, small firms show a significant and superior performance over large firms in five out of six regressions (all except year 1 of Panel C). Additionally, the results in Panel F favor the size phenomenon over the overreaction phenomenon. More specifically, all of the point estimates of α are negative, and they are significant at the $\alpha = 0.10$ level in years 1 and 3. Furthermore, the straight

Table IV
Comparison of Size-Matched Portfolios
Panel A

Average number (Number) and average market value of equity in 000's (Size) over the 1971–1981 sample period of the firms in the P_{1j} and P_{5j} portfolios, where $P_{i,j}$ ($i = 1$ and 5; $j = 1,5$) denotes the performance, size designation of the portfolio. $i = 1$ and 5 are losers and winners, respectively. $j = 1$ are the smallest firms; $j = 5$ are the largest. The firms are ranked into quintiles independently by size and earnings performance each year. Losers and winners are the bottom and top quintiles ranked by earnings performance. The earnings performance measure is defined as the change in earnings (before extraordinary items and discontinued operations) from the previous year to the current year, divided by the standard deviation of annual earnings changes over the five previous years.[a]

	P_{11}	P_{51}	P_{12}	P_{52}	P_{13}	P_{53}	P_{14}	P_{54}	P_{15}	P_{55}
Number	41.27	9.91	30.55	21.73	22.45	27.91	21.45	30.91	15.64	41.82
Size at End of Extreme Performance Year	30868	37867	104911	110009	265079	266454	606302	613931	3067029	3706430

Panel B

t-Statistics (t) to test the null hypothesis that the mean size of small losers equals the mean size of small winners and the null hypothesis that the mean size of large losers equals the mean size of large winners, 1971–1981.

Year	P_{11} vs. P_{51} Small Losers Versus Small Winners t	P_{15} vs. P_{55} Large Losers Versus Large Winners t
1971	0.32	−2.38*
1972	−1.01	−1.10
1973	−0.66	−2.69*
1974	−2.08*	0.10
1975	−0.37	−0.36
1976	−1.84	−1.35
1977	−0.56	−0.69
1978	−2.98*	−0.33
1979	−2.16*	1.89
1980	−1.22	−0.63
1981	−2.73*	0.41

[a] The t-tests are performed on the log of market value of equity, because market value of equity is (approximately) lognormally distributed.

* Statistically significant at $\alpha = 0.05$ level.

differences in 36-month cumulative excess returns for the five groups of size-matched and *PERF*-matched portfolios are (*t*-statistics are in parentheses):

P_{11}-P_{51}	P_{11}-P_{15}	P_{15}-P_{55}	P_{51}-P_{55}	P_{15}-P_{51}
−0.021	0.293	0.124	0.437	−0.314
(−0.36)	(3.2)	(1.1)	(4.9)	(−3.3)

For the size-matched portfolios (P_{11}-P_{51} and P_{15}-P_{55}), there is no evidence of differential performance between losers and winners. For the *PERF*-matched portfolios (P_{11}-P_{15} and P_{51}-P_{55}) small firms significantly outperform large firms. Additionally, small winners outperform large losers (P_{15}-P_{51}). Thus, the size- and *PERF*-matched portfolio results provide little support for the overreaction to earnings hypothesis; rather, they show that the overreaction to earnings phenomenon appears to be a manifestation of the size phenomenon.[17]

V. Conclusions

The results presented here fail to support the overreaction to earnings hypothesis. Although the poorest earnings performers outperform the best earnings performers by a statistically significant 16.6 percent over the 36 months subsequent to the extreme earnings year, this result is due primarily to differences in size between the two groups. Poor earners tend to be smaller firms than good earners. When poor earners are matched with good earners of equal size, there is little difference in return behavior. When poor (or good) earners of disparate sizes are compared, small firms outperform large firms, and smaller winners outperform larger losers. Thus, the statistically significant differences between the returns of extreme prior period performers appear to be the result not of investor overreaction to earnings but of the size effect. Our conclusions contrast with those of DeBondt and Thaler (1987, p. 579), who maintain, "The winner-loser effect is not primarily a size effect."

[17] As an additional test of the effects of size and overreaction to earnings on returns, we conducted pooled, cross-sectional time-series regressions of portfolio risk premia against beta, *PERF*, and log (size). The results of these tests supported those of the tests reported in the paper and are available from the author upon request.

REFERENCES

Amihud, Yakov and Haim Mendelson, 1986, Asset pricing and the bid-ask spread, *Journal of Financial Economics* 17, 223–249.

Ball, Ray and Philip Brown, 1968, An empirical evaluation of accounting income numbers, *Journal of Accounting Research* 6, 159–178.

Brooks, Leroy D. and Dale A. Buckmaster, 1976, Further evidence of the time series properties of accounting income, *Journal of Finance* 31, 1359–1373.

——— and Dale A. Buckmaster, 1980, First-difference signals and accounting income time-series properties, *Journal of Business Finance and Accounting* 437–454.

Chan, K. C., 1988, On the contrarian investment strategy, *Journal of Business* 147–163.

DeBondt, Werner F. M. and Richard H. Thaler, 1985, Does the stock market overreact? *Journal of Finance* 40, 793–805.

——— and Richard H. Thaler, 1987, Further evidence on investor overreaction and stock market seasonality, *Journal of Finance* 42, 557–580.

Fama, Eugene F. and Kenneth R. French, 1986, Common factors in the serial correlation of stock returns, Working Paper, Graduate School of Business, University of Chicago.

Foster, George, 1986, *Financial Statement Analysis*, 2nd ed. (Prentice-Hall).

Griffin, Paul A., 1982, *Usefulness to Investors and Creditors of Information Provided by Financial Reporting: A Review of Empirical Research* (Financial Accounting Standards Board).

Ibbotson, Roger and Rex Sinquefield, 1987, *Stocks, Bonds, Bills, and Inflation: 1987 Yearbook* (Ibbotson Associates).

Kahneman, Daniel and Amos Tversky, 1982, Intuitive prediction: Biases and corrective procedures, in Daniel Kahneman, P. Slovic, and Amos Tversky, eds.: *Judgment Under Uncertainty: Heuristics and Biases* (Cambridge University Press, London).

McElreath, Robert B., Jr. and C. Donald Wiggins, 1984, Using the COMPUSTAT tapes in financial research: Problems and solutions, *Financial Analysts Journal* 40, 71–76.

Pastena, Victor and William Ruland, 1986, The merger bankruptcy alternative, *Accounting Review* 51, 288–301.

Schwert, G. William, 1983, Size and stock returns, and other empirical regularities, *Journal of Financial Economics* 12, 3–12.