Review of Quantitative Finance and Accounting, 9 (1997): 17–34

Does Post-Earnings-Announcement Drift in Stock Prices Reflect A Market Inefficiency? A Stochastic Dominance Approach

VICTOR L. BERNARD
University of Michigan

H. NEJAT SEYHUN
University of Michigan

Abstract. This paper uses a stochastic dominance approach to test for market efficiency following earnings announcements. We find that the stocks that recently announced good earnings news stochastically dominate those that recently announced bad news. The results cast serious doubt on any belief that asset pricing model misspecifications might explain post-earnings-announcement drift.

Key words: Market efficiency, Earnings drift, Stochastic dominance

I. Introduction

More than two dozen studies have documented that after earnings numbers are announced, stock prices (net of market-wide movements) continue to drift in the direction of the earnings surprise for several months (Ball [1992], Bernard [1993]). Such evidence is consistent with a failure of the market to impound earnings information completely at the time it first becomes public. However, given how visible and widely followed are earnings reports, attributions of such evidence to market inefficiency have frequently been greeted with skepticism. An alternative possibility is that drift in the estimated abnormal returns represents a premium for some unidentified risk.

Several recent studies have focused on whether post-earnings-announcement drift in stock prices is best explained as a market inefficiency or a risk premium. Ball [1992, p. 342] argues, largely on the basis of evidence in Bernard and Thomas [1990], that the phenomenon "seems most likely due to either substantial information-processing costs or market inefficiency." However, Ball adds that it is unclear what substantial information processing costs would prevent elimination of the anomaly, and that the implications of the evidence for market efficiency are inherently limited by "our substantially incomplete knowledge of security pricing in a competitive market" (p. 342). In other words, firm conclusions are rendered difficult by the same joint hypothesis problem first pointed out by Fama [1970]: that one cannot test market efficiency without simultaneously testing some model of expected returns. It is impossible to assure that evidence apparently at odds with market efficiency isn't actually an indication of shortcomings in the hypothesized asset pricing model and its characterization of risk.

This paper deals with the joint hypothesis problem by using a stochastic dominance approach. Like any test of market efficiency, this approach requires *some* assumptions about asset pricing. However, the assumptions are extremely mild. For example, if one is willing to assume that investors prefer more wealth to less, then first-order stochastic dominance of one security over another implies an arbitrage opportunity.[1] Second-order stochastic dominance also implies an arbitrage opportunity, so long as investors are assumed risk averse. Given how weak are these assumptions, a demonstration of stochastic dominance constitutes compelling evidence of market inefficiency.

The ability to test market efficiency while imposing only the mildest of assumptions about asset pricing comes, of course, at some potential cost. The difficulty is that stochastic dominance demands so much of the data that even gross market inefficiencies may fail to produce stochastically dominant trading strategies. Indeed, some might doubt that inefficiencies stark enough to pass a stochastic dominance test could survive in active markets. Such doubts may explain why stochastic dominance tests have been so rarely used in the literature. However, if there are instances where stochastic dominance *can* be demonstrated, then such evidence should eliminate any reasonable concerns that the related market anomaly could be attributed to a failure to control for risk.

The tests in this study provide striking evidence of stochastic dominance of stocks that have recently announced good earnings news over those that recently announced bad news. More specifically, over the quarter following the earnings announcement, the portfolio of stocks with scaled unexpected earnings (SUE) in the highest decile dominates the portfolio in the lowest decile by *first-order* stochastic dominance. The estimated probability of obtaining such a result by chance is less than one-half of one percent. Even after allowing for 3 percent round-trip transactions costs, the highest SUE decile portfolio dominates the lowest by second-order stochastic dominance.

More surprisingly, stochastic dominance often holds even between portfolios for which the SUE differences are less extreme. Among ten SUE decile portfolios, there are 45 possible pairs to be compared; in 35 of these 45 (nonindependent) cases, the portfolio with the higher SUE dominates the other by either first-order or second-order stochastic dominance. For the 15 (nonindependent) comparisons of SUE portfolios at least five deciles apart, the higher SUE decile portfolio dominates the lower by first-order or second-order stochastic dominance in every case. For this group of 15 comparisons, first-, second-, or third-order stochastic dominance continues to hold in 14 cases, even after allowing for a 3 percent round-trip transactions cost. The likelihood of observing such a result by chance is remote.

The results of this paper cast serious doubt on any belief that asset pricing model misspecifications might explain post-earnings-announcement drift. An understanding of this anomaly appears to require either some model of inefficient markets, or identification of some cost (other than transactions costs) that impede the impounding of public information in prices.

The remainder of the paper is organized as follows. In section II, we provide a brief overview of the phenomenon of post-earnings-announcement drift, and report estimates of such drift within our sample. Section III explains how the stochastic dominance criterion can be used to discriminate between alternative explanations for the drift. The test results are presented in section IV, and conclusions appear in section V.

II. Post-earnings-announcement drift

A long series of studies, from Ball and Brown [1968] to Bernard, Thomas, and Abarbanell [1993], have documented the existence of post-earnings announcement drift. Estimated abnormal returns have been shown to vary with unexpected earnings, for six to nine months after earnings become public. The phenomenon is robust; a combination of studies documents its existence in every single year examined, from 1965 through 1991 (Bernard [1993], Bernard, et al [1993]). The phenomenon is not explainable as the product of a failure to adjust for well-known risk factors, such as beta or factors from the APT, nor is it the product of a number of potential research design flaws, such as beta or factors from the APT, nor is it the product of a number of potential research design flaws, such as survivor bias, database restatement and other hindsight biases, or measurement error in CRSP returns (see Bernard and Thomas [1989, 1990] and Ball [1992]).

A tabular description of the phenomenon appears in Table 1. The data underlying the

Table 1. Announcement-Period Reaction and Post-Earnings-Announcement Drift for Standardized Unexpected Earnings (SUE) Decile Portfolios

			Post-announcement abnormal return:	
SUE Decile	Number of firm-quarters	Two-day announcement period abnormal return	through first subsequent earnings announcement	through second subsequent earnings announcement
1	10,544	−2.05%	−3.36%	−4.99%
2	10,654	−1.55	−2.61	−3.46
3	10,660	−0.99	−1.64	−2.06
4	10,494	−0.37	−0.58	−0.76
5	10,565	0.19	0.16	0.38
6	10,868	0.77	0.68	0.83
7	10,956	1.08	1.14	1.43
8	11,118	1.34	1.37	1.78
9	11,304	1.55	1.90	2.28
10	11,119	1.77	2.21	2.70
Long in 10; Short in 1		3.82%	5.57%	7.69%

Standardized unexpected earnings (SUE) is equal to the seasonally differenced quarterly earnings, scaled by its historical estimated standard deviation, using up to 80 observations.
Estimated abnormal returns are equal to raw returns, less the contemporaneous return on a size-matched decile portfolio.
The two-day announcement period abnormal return includes the size-adjusted return for the day prior to and the day of the earnings announcement as indicated by Compustat.

table are approximately 110,000 quarterly observations for NYSE/AMEX firms for 1971–1991.[2] Each observation is classified into a decile, based on a crude measure of unexpected earnings used in many prior studies: the change in quarterly earnings relative to the comparable quarter of the prior year, and scaled by the historical standard deviation of that change.[3] (We refer to this measure as standardized unexpected earnings, or SUE.) For each decile, Table 1 presents information about stock price behavior for a 2-day announcement period, the quarter following the announcement (up to and including the date of the announcement of the following quarter), and the two quarters following the announcement. The estimated abnormal stock returns are equal to raw returns, less the return on a size-matched decile portfolio.[4]

We examine size-adjusted returns for two reasons. First, our predictions regarding arbitrage opportunities strictly apply only to abnormal returns net of market movements. If we were to examine total returns, inclusive of the non-predictable market returns, our tests would be less powerful. Second, there is evidence that size related returns in January are already subject to stochastic dominance (Seyhun [1993]). January returns in small firms dominate the January returns in larger firms by first-order stochastic dominance. By using size-adjusted returns, we eliminate the possibility that our findings are somehow caused by the size effect. However, we also replicate our tests using both raw returns as well as market-adjusted returns. These tests produce mostly similar results and they are discussed separately.

Table 1 shows that the 2-day size-adjusted announcement returns increase monotonically with SUE. This simply confirms that earnings announcements convey some new information. The key columns are those that show the *post*-announcement abnormal returns *also* increase monotonically with SUE. A long position in SUE decile 10, in combination with a short position in SUE decile 1, produces an estimated abnormal return of 5.57 percent over the quarter subsequent to the earnings announcement. The estimated abnormal return continues to increase during the second subsequent quarter, reaching 7.69 percent by the time the second earnings announcement has occurred.

Bernard and Thomas [1989, 1990] show that about one-fourth of the post-announcement drift is concentrated in the three days surrounding subsequent earnings announcements. This is what one would expect in market with naive earnings expectations. Specifically, if the market failed to recognize that a change in quarterly earnings (relative to the prior year) tends to be followed by changes in the same direction for the next two or three quarters, it would be surprised when (on average) such changes are announced for the subsequent quarters. In fact, Bernard and Thomas [1990] show that reactions to at least the subsequent four announcements are predictable, based on current SUEs, with signs and magnitudes that are consistent with the market underestimating the autocorrelation in quarterly earnings.[5] Bernard and Thomas [1990] argue (and Ball [1992] concurs), that this portion of post-announcement drift concentrated around subsequent earnings announcements is unlikely to be explained by a failure to adjust fully for risk.[6,7]

In this study, we focus on the longer-term (3- to 6-month) drift, where a risk adjustment explanation is more plausible. It is this longer-term drift that is more relevant for practical purposes; only over the longer interval does the drift plausibly exceed normal transactions costs. A number of funds take positions over this longer interval as at least one element in their trading strategies.

III. Stochastic dominance tests

We now turn to stochastic dominance tests as a vehicle for assessing the likelihood that unidentified risk factors could explain post-earnings-announcement drift. The stochastic dominance approach provides a simple method of choice among risky alternatives (Whitmore and Findlay (1978)). There are three major types of stochastic dominance: first-order (FSD), second-order (SSD), and third-order (TSD). Under FSD, dominance indicates an arbitrage opportunity so long as investors are assumed to be non-satiated expected utility maximizers. SSD demands less of the data to demonstrate dominance, but assumes that investors are both non-satiated and risk averse. TSD is least demanding of the data, but assumes investors are non-satiated and risk averse, with decreasing absolute risk aversion.

To demonstrate FSD, one compares the cumulative probability functions of the payoffs to alternative assets. An asset X with a cumulative probability distribution F_1 would dominate an asset Y with cumulative probability distribution G_1 by FSD if and only if $F_1(x) < G_1(x)$ for all possible outcomes, x.

An asset X would dominate asset Y by SSD, if and only if $F_2(x) < G_2(x)$, for all possible x, where F_2 and G_2 denote the areas under F_1 and G_1 respectively. Thus, the cumulative probability functions F_1 and G_1 may overlap, so long as the area beneath G_1 is always greater for values less than any given payoff.

Finally, asset X would dominate asset Y by TSD if and only if $\mu_F > \mu_G$ and $F_3(x) < G_3(x)$ for all possible x, where μ_X and μ_Y are the expected returns to assets X and Y, and F_3 and G_3 denote the areas under F_2 and G_2 respectively.[8]

Stochastic dominance results imply hierarchy. FSD implies SSD, which in turn implies TSD. Hence, a finding that only TSD exists in turn implies that SSD or FSD do not exist.

To implement a stochastic dominance approach, the realized returns to various SUE portfolios are used to infer the unobservable population distributions. Particularly with respect to extreme values in the population, any given sample may not provide a good guide for frequency of population returns. For example, the absence within a particular sample of an extreme negative return on a high SUE portfolio, even though such extreme observations are present in the (unobservable) population, could lead one to incorrectly conclude that such a portfolio dominates one with low SUEs. While this possibility (sometimes labeled the "peso problem") always exists, it is rendered less likely in this study by our use of a large sample that spans two decades.

IV. Empirical evidence on stochastic dominance

A. Main results

Figure 1 compares the estimated cumulative probability functions for the post-announcement size-adjusted returns to four SUE decile portfolios: decile 1 (lowest SUE), decile 4, decile 7, and decile 10. The returns are measured from the day after the earnings

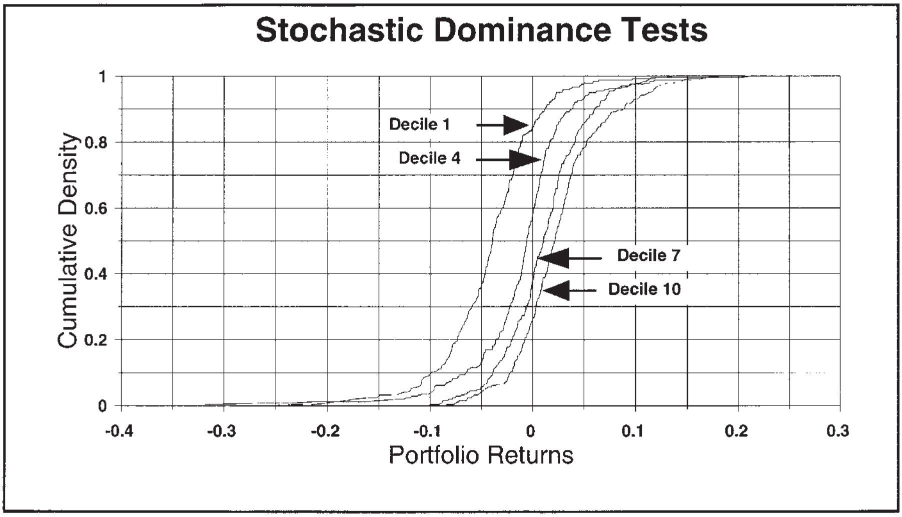

Figure 1. Cumulative density function (CDF) of the abnormal post-earnings-announcement drift, calculated from the day after the earnings announcement through the subsequent earnings announcement. CDFs are shown for Standardized Unexpected Earnings (SUE) deciles 1 (lowest), 4, 7, 10 (highest). The decile portfolios are formed every month from 1971 through 1991, thus giving rise to 242 observations for each CDF.

announcement through the day of the first subsequent announcement. The returns for all firm-quarters within a given SUE decile that share a given month as the fiscal quarter end constitute a portfolio, and are aggregated. Since we have data for a period that spans 242 months, each SUE decile portfolio produces returns for quarterly intervals with 242 different starting times. These 242 (partially overlapping) quarterly returns constitute the empirical distribution for a given SUE portfolio.

Figure 1 shows that the probability of a loss varies substantially across SUE portfolios 1, 4, 7, and 10. The probabilities are 84 percent, 57 percent, 37 percent, and 26 percent, respectively. The probability of a loss in excess of 5 percent also varies substantially: 36 percent, 13 percent, 5 percent, and 4 percent, respectively. Yet, such large differences are not necessarily sufficient to guarantee stochastic dominance; dominance tests take into account differences in such probabilities at *every* level of return.

In Figure 1, the cumulative probability function SUE decile 10 *always* lies to the right of that for SUE decile 1. Put differently, the probability of achieving any given return is always higher for SUE decile 10, regardless of the level of that return. The higher probabilities are difficult to attribute to chance, based on the non-parametric Kolmogorov-Smirnoff test.[9]

Any investor who prefers a higher return to a lower one would prefer SUE decile 10, regardless of the more specific characteristics of his/her utility function; SUE portfolio 10 dominates SUE portfolio 1 by first-order stochastic dominance. Given how mild are these assumptions about investors' utility functions, this represents compelling evidence of a market inefficiency.

First-order stochastic dominance does not hold for the remaining comparisons (10 vs 7; 10 vs. 4; 7 vs. 4; 7 vs. 1; 4 vs. 1). However, in four of five of these comparisons, FSD fails to hold only because of a small degree of overlap in the distribution functions for extreme high returns, and SSD does hold. That is, the lower SUE portfolios produce high returns within certain small ranges more frequently than do the higher SUE portfolios. Since the higher SUE portfolios dominate over all other ranges of the return distribution, only an investor with an extreme preference for high returns within the region of overlap would prefer the lower SUE portfolio. Any other investor, including *all* investors who are risk averse, would strictly prefer the higher SUE portfolio. Formal statistical tests of these propositions are provided in Section D of the paper.

In the comparison of SUE decile portfolios 4 and 1, no stochastic dominance is established. The reason is that the higher SUE portfolio produces two returns that are more negative (−32 percent and −27 percent) than the most negative return from the lower SUE portfolio (−24 percent). Even though the distribution function for SUE portfolio 4 lies to the right of that for SUE portfolio 1 throughout the remainder of the return distribution, the two large negative returns are sufficient to destroy all forms of stochastic dominance. The reason is that investors with an extreme disutility for the large negative return might not prefer portfolio 4, even though it produces higher returns with a greater frequency.

Table 2 Panel A summarizes how frequently stochastic dominance holds for pairs of portfolios from all of the various SUE deciles. FSD holds not only for the decile 10 over 1; it also holds in ten other comparisons. Note that where no stochastic dominance holds, it is for comparisons of portfolios with relatively similar levels of SUE. Where SUEs are separated by five or more deciles, stochastic dominance (either FSD or SSD) holds in 15 of 15 comparisons. Subject to further investigation of statistical significance and the impact of transactions costs, such evidence is difficult to explain except as the product of a market inefficiency. High SUE stocks appear to be underpriced relative to low SUE stocks.

Table 2 Panel B presents comparable results, in the case where post-earnings-announcement returns are measured over the two quarters following the announcement. One might expect movement to a longer window would enhance the likelihood of establishing stochastic dominance, because the longer window captures more of the drift. However, the drift during the second quarter is not as large as that in the first quarter (refer to Table 1), and the longer interval introduces more noise. The results indicate that over this longer window, stochastic dominance holds for fewer comparisons. In Table 2 Panel B, SUE portfolio 10 no longer dominates SUE portfolio 1. In the comparisons where SUEs differ by at least five deciles, SSD or FSD holds in 10 of 15 comparisons.

B. Sensitivity tests

Table 3 presents supplemental results based on levels of portfolio aggregation different from those used in the main tests. In the tests above, portfolios were formed 242 times, by grouping firms with fiscal quarter ends in the same month. In Table 3 Panel A, we form

Table 2. Stochastic Dominance of SUE Decile Portfolios over Lower SUE Decile Portfolios; Portfolios formed Monthly

Panel A: Post-Earnings-Announcement Drift Measured over One Quarter

Potentially dominating SUE decile portfolio:	SUE decile portfolio subject to domination:								
	1	2	3	4	5	6	7	8	9
2	SSD								
3	SSD	—							
4	—	—	—						
5	SSD	FSD	SSD	SSD					
6	SSD	—	—	FSD	—				
7	SSD	FSD	FSD	SSD	SSD	SSD			
8	SSD	FSD	SSD	SSD	—	SSD	—		
9	SSD	FSD	SSD	FSD	SSD	SSD	—	SSD	
10	FSD	FSD	SSD	FSD	SSD	SSD	SSD	FSD	SSD

Panel B: Post-Earnings-Announcement Drift Measured over Two Quarters

Potentially dominating SUE decile portfolio:	SUE decile portfolio subject to domination:								
	1	2	3	4	5	6	7	8	9
2	—								
3	SSD	SSD							
4	—	—	—						
5	SSD	SSD	SSD	SSD					
6	SSD	FSD	FSD	FSD	—				
7	—	—	—	—	—	—			
8	SSD	SSD	SSD	FSD	—	—	—		
9	SSD	SSD	SSD	FSD	SSD	SSD	SSD	SSD	
10	—	SSD	—	SSD	—	—	TSD	—	—

SUE decile portfolios can dominate other SUE decile portfolios by first-order stochastic dominance (FSD), second-order stochastic dominance (SSD), or third-order stochastic dominance (TSD). Where no stochastic dominance is indicated, a dash (—) appears.

Stochastic dominance tests are based on comparison of empirical distribution of size-adjusted returns of the two SUE portfolios, from the day after the earnings announcement through the first (in Panel A) or second (in Panel B) subsequent announcement. SUE portfolios are constructed for fiscal quarters ended in each of 242 months for which data are available during the 1970–1991 period, and thus the empirical distributions are based on 242 observations.

Standardized unexpected earnings (SUE) is equal to the seasonally differenced quarterly earnings, scaled by its historical estimated standard deviation, using up to 80 observations.

portfolios by grouping all firms whose fiscal quarters end within the same calendar *quarter*. Thus, there are only 82 portfolios (one per quarter). This approach would be relevant to the investor with a longer investment horizon who buys and sells the shares of all firms announcing earnings over a 3-month period. It also avoids the poor level of diversification that can occur when grouping by month, for those months (e.g., February,

Table 3. Supplemental Stochastic Dominance Tests: Portfolios formed Quarterly; Individual Stocks

Panel A: Dominance of SUE Decile Portfolios over Other SUE Portfolios, formed Quarterly

Potentially dominating SUE decile stocks:	SUE decile stocks subject to domination:								
	1	2	3	4	5	6	7	8	9
2	SSD								
3	SSD	FSD							
4	SSD	FSD	SSD						
5	SSD	—	—	—					
6	SSD	FSD	SSD	SSD	SSD				
7	SSD	FSD	—	—	SSD	—			
8	SSD	FSD	SSD	SSD	SSD	SSD	SSD		
9	SSD	FSD	SSD	SSD	SSD	FSD	SSD	—	
10	SSD	FSD	SSD	SSD	SSD	FSD	SSD	SSD	—

Panel B: Dominance of Individual Stocks over Other Stocks in Lower SUE Deciles

Potentially dominating : SUE decile stocks	SUE decile stocks subject to domination:								
	1	2	3	4	5	6	7	8	9
2	SSD								
3	SSD	—							
4	SSD	—	—						
5	SSD	—	—	—					
6	SSD	SSD	SSD	SSD	SSD				
7	SSD	—	—	—	—	—			
8	SSD	SSD	SSD	SSD	SSD	—	SSD		
9	SSD	SSD	SSD	SSD	SSD	—	SSD	—	
10	SSD	SSD	SSD	SSD	SSD	—	SSD	—	—

Portfolios (in Panel A) or individual stocks (in Panel B) can dominate others in lower SUE deciles by first-order stochastic dominance (FSD), second-order stochastic dominance (SSD), or third-order stochastic dominance (TSD). Where no stochastic dominance is indicated, a dash (—) appears.

Stochastic dominance tests are based on comparison of empirical distribution of size-adjusted returns of portfolios (in Panel A) or individual stocks (in Panel B) from two SUE deciles, from the day after the earnings announcement through the first subsequent announcement. Each SUE decile contains approximately 11,500 firm-quarters, and thus the empirical distributions in Panel B are based on this many observations. In Panel A, firms are grouped into calendar quarters, according to their fiscal quarter end. The are 84 quarterly observations underlying the empirical distributions examined in Panel A.

Standardized unexpected earnings (SUE) is equal to the seasonally differenced quarterly earnings, scaled by its historical estimated standard deviation, using up to 80 observations.

May, August, and November) in which few firms have fiscal quarter ends. The potential disadvantage of the approach arises in the form of statistical inefficiency; there are fewer observations with which to estimate a cumulative density function.

Table 3 Panel A shows that when we form portfolios on a quarterly basis, the stochastic dominance results generally grow stronger. FSD holds in nine cells, and either SSD or FSD holds in 37 of 45 comparisons, including all comparisons where the SUE deciles differ by five or more. The strengthening of the results is due to the avoidance of small, poorly-diversified portfolios associated with months with few fiscal-quarter ends.

Table 3 Panel B replicates the procedures applied to portfolio returns, but compares distributions of *individual security* returns for stocks in different SUE deciles. These results "stack the deck" against a finding of stochastic dominance, in the sense that individual securities expose the investor to risks that would be readily diversifiable within an SUE portfolio. Nevertheless, SSD holds for many of the comparisons, including the comparison of decile 10 versus 1, and in 14 of 15 comparisons where the SUEs are at least five deciles apart. This is a striking result; it suggests that the differential degree of mispricing of SUE portfolios is so robust that stochastic dominance holds even in the face of exposure to individual security risk.

As mentioned earlier, we replicated our tests using raw returns as well as market-adjusted returns. In addition, we used fixed holding periods of exactly 63 and 126 trading days, instead of the subsequent earnings announcement dates as the ending date. The 63 trading days correspond to a 3-month holding period while the 126 days corresponds to 6-month holding period. For the 3-month raw returns, we found 29 instances of stochastic dominance out of 45 pairs. Moreover, decile 1 was dominated by all higher deciles by first- or second-order stochastic dominance. For the 6-month raw returns, there were 38 instances of stochastic dominance out of 45 pairs. When portfolios were separated by at least three deciles, there were 28 instances of domination out of 28 pairs. Hence, raw returns exhibit pretty similar patterns as the size-adjusted returns.

Next, we examined the market-adjusted returns. For market returns we used the value-weighted portfolio of NYSE, AMEX, and NASDAQ stocks. To compute the market-adjusted returns we simply subtracted the holding period return to the market portfolio from the holding period return to the individual stock. Market adjusted returns gave somewhat stronger results. For the 3-month horizon, there were 40 instances of stochastic dominance out of 45 pairs. For the 6-month horizon, there were 39 instances of stochastic dominance out of 45 pairs.

A final note concerns stability of the results. We have implicitly assumed that the distribution of size-adjusted returns within a given SUE decile is stable through time. To test that assumption, we conduct tests like those in Table 2 Panel A (where portfolios were formed monthly), separately for the 1970s and 1980s. The results for each decade are similar, and are similar to those for the overall period. FSD is found in 11 of 45 comparisons during the 1970s, and for 15 of 45 comparisons in the 1980s. Either SSD or FSD is found in 34 of 45 comparisons during the 1970s, and in 33 of 45 comparisons during the 1980s. Comparable stability tests were based on portfolios formed quarterly (as in Table 3 Panel A), and again the results were similar over time.

C. Tests that account for transactions costs

The existence of portfolios that stochastically dominate others before consideration of transactions costs indicates a market inefficiency in the sense that stock prices appear not to reflect fully the available earnings information. However, an interesting question is whether the inefficiency represents a profitable trading opportunity for investors with various levels of transactions costs. Here, we consider the impact on the tests of round-trip transactions costs of 1 percent, 2 percent, and 3 percent. These amounts easily accommodate commission fees, which are often less than .15 percent; they also accommodate the mean effective bid-ask spread of approximately .44 percent (Peterson and Fialkowski [1994]). When 3 percent costs are considered, they exceed commissions and spreads for all but the smallest stocks.[10] However, large market orders could exert price pressure, creating a cost not captured by the spread.

Table 4 presents the results. We assume that a shift from one decile portfolio to another involves transactions costs of 1 percent (Panel A), 2 percent (Panel B), or 3 percent (Panel C). The introduction of 1 percent transactions costs causes a change from FSD to SSD in several cells, and from SSD to no stochastic dominance in several others. Nevertheless, we still find stochastic dominance in most cells and FSD in the most extreme comparison: SUE portfolios 10 and 1. Of course, the number of cells where stochastic dominance can still be demonstrated declines as the transactions costs increase. However, even for transactions costs of 3 percent, SSD continues to hold for 13 of the 15 comparisons where the SUEs are at least five deciles apart, and TSD holds in one of the two remaining comparisons. SUE portfolio 10 continues to dominate SUE portfolio 1 by SSD even after accounting for a 3 percent transactions cost. Thus, subject to consideration of statistical significance, the results suggest market inefficiency, even if it were defined so as to require profitable trading opportunities *after* consideration of transactions costs.

We also conducted an alternative test as a check on the influence of transactions costs. The test was motivated by the concern that, while the above tests assume a constant level of transactions costs across firms, we recognize that bid-ask spreads and any costs due to price pressure tend to be greater for small firms. In our alternative test, we repeated the above analysis while restricting attention to only larger firms (those with market value of equity in excess of $100 million), and while allowing for transactions costs of 1 percent. We continue to find either first- or second-order stochastic dominance in 14 of the 15 cells where the SUEs are at least five deciles apart.

D. Significance tests

Since population cumulative probability functions can be estimated only with sampling error, there is a possibility that our results are due to chance. To assess the likelihood of this possibility, we conduct a set of randomization tests. The results appear in Table 5.

The randomization procedure works as follows. We randomly assign (with replacement) firm-quarters to "pseudo-SUE" deciles. Since the assignment is random, the expected abnormal returns should be the same for each decile, and one should stochastically dominate the other only by chance. We calculate the one-quarter-long post-announcement size-adjusted returns for "pseudo-SUE" decile portfolios (in Panel A) or individual securities (in Panel B). For a given pseudo-SUE decile, there are 242 portfolio returns (one for each quarterly period beginning in a different month of the test period); there are 11500 individual security returns. We then compare the distributions of these returns across SUE

Table 4. Stochastic Dominance of SUE Decile Portfolios over Lower SUE Decile Portfolios, after Transactions Costs

Panel A: 1 Percent Transactions Cost; Drift Measured over One Quarter

Potentially dominating SUE decile portfolios:	SUE decile portfolio subject to domination:								
	1	2	3	4	5	6	7	8	9
2	SSD								
3	SSD	—							
4	—	—	—						
5	SSD	SSD	TSD	—					
6	SSD	—	—	SSD	—				
7	SSD	FSD	SSD	SSD	SSD	—			
8	SSD	SSD	SSD	SSD	—	—	—		
9	SSD	SSD	SSD	SSD	SSD	SSD	—	—	
10	FSD	FSD	SSD	FSD	SSD	SSD	SSD	—	—

Panel B: 2 Percent Transactions Cost; Drift Measured over One Quarter

Potentially dominating SUE decile portfolio:	SUE decile portfolio subject to domination:								
	1	2	3	4	5	6	7	8	9
2	—								
3	SSD	—							
4	—	—	—						
5	SSD	TSD	—	—					
6	SSD	—	—	—	—				
7	SSD	FSD	SSD	SSD	—	—			
8	SSD	SSD	SSD	SSD	—	—	—		
9	SSD	SSD	SSD	SSD	SSD	—	—	—	
10	SSD	FSD	SSD	FSD	SSD	—	—	—	—

Panel C: 3 Percent Transactions Cost; Drift Measured over One Quarter

Potentially dominating SUE decile portfolio:	SUE decile portfolio subject to domination:								
	1	2	3	4	5	6	7	8	9
2	—								
3	—	—							
4	—	—	—						
5	SSD	—	—	—					
6	SSD	—	—	—	—				
7	SSD	FSD	—	—	—	—			
8	SSD	SSD	SSD	—	—	—	—		
9	SSD	SSD	SSD	TSD	—	—	—	—	
10	SSD	SSD	SSD	SSD	—	—	—	—	—

See footnotes to Table 2 for explanation.

Table 5. Randomization estimates of Probability of Finding Stochastic Dominance By Chance

Panel A: Stochastic Dominance Tests Based on Abnormal Returns to *Portfolio*

Number of simultaneous comparisons of return distributions	Number of observations underlying each return distribution	Frequency of first-order stochastic dominance (FSD)	Frequency of second-order stochastic dominance (SSD)	Frequency of third-order stochastic dominance (TSD)
Portfolio from one "pseudo-SUE" decile is tested for domination by portfolio from *one* other decile	242	.004	.202	.285
Portfolio from one "pseudo-SUE" decile is tested for domination by portfolios from *each of two* other deciles	242	.004	.098	.147
Portfolio from one "pseudo-SUE" decile is tested for domination by portfolios from *each of three* other deciles	242	.000	.044	.067
Portfolio from one "pseudo-SUE" decile is tested for domination by portfolios from *each of four* other deciles	242	.000	.031	.055

Panel B: Stochastic Dominance Tests Based on Abnormal Returns to *Individual Stocks*

Number of simultaneous comparisons of return distributions	Number of observations underlying each return distribution	Frequency of first-order stochastic dominance (FSD)	Frequency of second-order stochastic dominance (SSD)	Frequency of third-order stochastic dominance (TSD)
Individual stocks in one "pseudo-SUE decile" are tested for domination by stocks in *one* other decile	11,500	.000	.121	.196
Individual stocks in one "pseudo-SUE decile" are tested for domination by stocks in *each of two* other deciles	11,500	.000	.039	.061

Tests are conducted by assigning firm-quarters to "pseudo-SUE" portfolios at random. Since the assignments are random, the null hypothesis of no difference between the returns distributions of portfolios (in Panel A) or individual stocks (in Panel B) should hold, except for sampling error. Stochastic dominance tests like those conducted in prior tables are applied, and the presence of FSD, SSD, or TSD is recorded. The process is repeated 1000 times.

deciles, using the same procedures applied previously. We check for FSD, SSD, and TSD for each pair of pseudo-SUE deciles. We repeat the procedure 1000 times.

The first row of Table 5 Panel A indicates that FSD obtains by chance only 0.4 percent of the time. Thus, any given finding of FSD is statistically significant at conventional levels. In contrast, SSD holds 20 percent of the time by chance. Therefore, our previous demonstration of SSD for any *single* pair of SUE portfolios cannot be considered a statistically significant result at the standard .05 level.

The second, third, and fourth rows of Table 5 Panel A present estimates of the probability that a given SUE portfolio would be dominated by chance by *each* of two, three, or four other SUE portfolios. The estimated probabilities decline from 9.8 percent, to 4.4 percent, to 3.1 percent as the number of simultaneous comparisons increase from two, to three, to four. The final indicated probability—3.1 percent—would be relevant in assessing, for example, the significance of the finding in Table 2 Panel A that portfolio 1 is dominated by *each* of portfolios 7 through 10. Thus, when viewed in isolation, that finding appears statistically significant at conventional levels.

Ideally, we would report the significance level associated with (for example) finding SSD in 35 of 45 cases in Table 2 Panel A, or in all 15 of the comparisons where the SUEs are separated by at least five deciles. Even modern computing power is insufficient to render randomization of all 45 tests economically feasible. Nevertheless, it is apparent that the results in Table 2 Panel A are almost certainly not attributable to chance. To see this, focus on the 15 comparisons with SUEs at least five deciles apart. In the first column, the probability of decile 1 being dominated by each of *five* deciles (6 through 10) must be less than the 3 percent indicated in Table 5 for SSD in *four* comparisons—and that much can be said even before taking into account that one of the comparisons in the first column produces FSD. In the second column, the probability of finding of FSD by chance in all four relevant comparisons is estimated in Table 5 to be 0.0 percent. The probability of finding SSD in all three comparisons in the third column is estimated to be 4.4 percent. In the fourth column, FSD occurs in both relevant comparisons—a finding that should occur by chance only 0.4 percent of the time. The SSD in the single relevant cell of the fifth column should occur by chance 20 percent of the time. Given the dependence in the results across columns, it is difficult to know how to aggregrate these probability estimates, but clearly the probability of observing the conjunction of these results by chance is remote.

To conclude, in those 15 comparisons where power should be greatest, SSD or FSD holds for every comparison, a result that is almost certainly not attributable to chance. The results in Table 4, where transactions costs are considered, are not as strong, but an analysis like that applied above to Table 2 Panel A reveals that the results in Table 4 are also unlikely to be attributable to chance.

The probability of observing stochastic dominance by chance is lower when individual securities are the unit of observation, as in Table 2 Panel B, because there are more observations underlying each return distribution. These probabilities are estimated in Table 5 Panel B. Although stochastic dominance does not hold as frequently in the individual security data, it still appears that the results are difficult to attribute to chance.

V. Conclusions

Although post-earnings-announcement drift has been recognized for more than two decades, financial economists still debate whether the anomaly reflects a market inefficiency or a failure to adjust abnormal returns fully for risk. The results of this paper should eliminate any reasonable concern that risk adjustment problems explain the phenomenon. That is, one need not rely on any specific model of asset pricing in order to demonstrate the existence of arbitrage opportunity based on publicly available information about earnings. So long as we make the mildest of assumptions—that investors prefer more wealth to less—then extreme good news stocks dominate extreme bad news stocks. If we add one other mild assumption—that investors are risk averse—then for SUE decile portfolios at least five deciles apart, the portfolio with the higher SUE always dominates the other. Even after allowing for 3 percent transactions costs, dominance continues to hold in nearly all of those comparisons.

Bernard and Thomas [1989,1990] presented evidence suggesting that post-earnings announcement drift is not explainable as a product of a failure to adjust returns fully for risk. Nevertheless, the evidence is not conclusive and doubts remain (Ball [1992]). *Ex ante*, it was unclear whether stochastic dominance tests could help resolve the debate, because such tests are so demanding of the data that even gross inefficiencies could fail to produce stochastically dominant trading strategies. However, given that stochastic dominance has been demonstrated here, the results constitute compelling evidence of market inefficiency.

The above findings suggest that attempts to explain post-earnings-announcement drift would be most fruitfully focused on reasons why market prices in competitive markets might not reflect all available information. The reasons are far from clear; most existing models of mispricing, such as DeLong, Shleifer, Summers, and Waldman [1990], predict *random* pricing errors, not the systematic, predictable errors investigated here. Bhushan [forthcoming] posits that a combination of costly trading and some naive investors who make systematic valuation errors could explain the phenomenon. However, this argument begs the question of why a block of investors would make common, systematic errors in the first place.

The findings in this paper also serve to illustrate the usefulness of stochastic dominance tests for assessing the likelihood that a given stock market anomaly is attributable to inefficiency. Even though stochastic dominance tests are free of all but the mildest of assumptions about asset pricing, they have not appeared frequently in the literature. One possible explanation is that they are viewed as so demanding of the data that dominance would be difficult to demonstrate, even when market inefficiencies do exist. Alternatively, some might consider it implausible that inefficiencies stark enough to pass a stochastic dominance test could survive in active markets. The results in this paper indicate that it *is* possible to identify broad groups of stocks that stochastically dominate others. In so doing, the results suggest that stochastic dominance tests would be a useful tool in a variety of applications involving tests of market efficiency.

Appendix A: Implementation of stochastic dominance

The underlying data consist of a finite set of portfolio returns. Consequently, probability distributions are defined in terms of discrete observations. The ordered relative sample frequency $f(x_i)$ for portfolio X with K_1 observations is defined as,

$$f(x_i)\begin{cases} = 1/K_1, \text{ if } x_i \in X \\ = 0 \qquad \text{otherwise,}\end{cases} \tag{A.1}$$

where x_i denotes the simple arithmetic returns to decile portfolios. The relative sample frequencies $g(x_i)$ for portfolio Y is defined analogously with K_2 observations.

Let $N = K_1 + K_2$

i) Portfolio X dominates portfolio Y by first-order stochastic dominance (FSD) if and only if $F_1(x_n) \leq G_1(x_n)$ for all $n \leq N$, with strict inequality for at least one $n \leq N$, where

$$F_1(x_n) = \sum_{i=1}^{n} f(x_i) \qquad n = 1 \ldots N \tag{A.2}$$

and $G_1(x_n)$ is defined analogously.

ii) Portfolio X dominates portfolio Y by second-order stochastic dominance (SSD) if and only if $F_2(x_n) \leq G_2(x_n)$ for all $n \leq N$ with strict inequality for at least one $n \leq N$ where

$$F_2(x_n) = \sum_{i=2}^{n} F_1(x_{i-1})(x_i - x_{i-1}) \qquad n = 2 \ldots N \tag{A.3}$$

and $F_2(x_1) = 0$. $G_2(x_n)$ is defined analogously.

iii) Portfolio X dominates portfolio Y by third-order stochastic dominance (TSD) if and only if $F_3(x_n) \leq G_3(x_n)$ for all $n \leq N$ with strict inequality for at least one $n \leq N$ and $F_2(x_N) \leq G_2(x_N)$ where

$$F_3(x_n) = 1/2 \sum_{i=2} [F_2(x_i) + F_2(x_i - 1)]\,[x_i - x_{i-1}] \qquad n = 2 \ldots N \tag{A.4}$$

and $F_3(x_1) = 0$. $G_3(x_n)$ is defined analogously.

Acknowledgments

Professor Victor Bernard, Price Waterhouse Professor of Accounting at the University of Michigan passed away on November 14, 1995 at the age of 42. He is missed by family, friends and colleagues. Seyhun is a Professor of Finance, Michigan Business School. Jacob Thomas and Jim Wahlen graciously provided a portion of the data used in this study.

Notes

1. The stochastic dominance referred to here and elsewhere in the paper is expressed in terms of per dollar returns. If instead dominance were defined in terms of dollar payoffs without holding constant the amounts invested in the assets being compared, additional conditions would be required to establish that first-order stochastic dominance implies an arbitrage opportunity (Jarrow [1986]). In any case, however, arbitrage requires that claims on the compared assets must trade.
2. The sample is obtained from firms included in the daily CRSP files that also appear on any edition of the quarterly Compustat files from 1982 through 1987 or 1991. To be included, firms were required to have at least 10 consecutive quarters of earnings (before extraordinary items) available on Compustat, and the price data required to calculate the needed returns.
3. Other measures of unexpected earnings generate similar results (Bernard and Thomas [1989, 1990]; Freeman and Tse [1989]). Since the SUE decile assignments could not be made until all firms announced earnings for a given quarter, there is some hindsight bias introduced here, as first explained by Holthausen [1983]. However, the bias turns out to be trivial (Bernard and Thomas [1989]).
4. The daily abnormal returns are summed over time to produce the estimates reported here, thus implicitly assuming daily rebalancing. However, when returns are compounded, the post-earnings-announcement drift is as large if not larger (Bernard and Thomas [1989]).
5. The autocorrelation in quarterly earnings changes (relative to the comparable quarter of the prior year) tends to be positive but declining over the first three lags, negative at the fourth lag, and slightly negative thereafter. Thus, a market that ignored these autocorrelations would exhibit reactions to subsequent announcements that are correlated with the initial quarterly earnings change in the same pattern. That, in fact, is what Bernard and Thomas [1990] document in the data.
6. One reason is that the risk shifts necessary to explain the pattern of predictable market reactions (see prior footnote) would be so complex as to challenge plausibility. Another reason is that the estimated abnormal returns around subsequent earnings announcements are large per unit time, and therefore explainable only by risk premiums that are also large. For example, the risk premium around the first subsequent announcement for a long position in SUE decile 10, combined with a short position in SUE decile 1, would have to be over 100 percent on an annualized basis, before compounding.
7. Ball and Bartov [1993] speculate that a portion of the short-interval, concentrated drift could possibly reflect some systematic "bid-ask bounce." If so, that portion of the drift would disappear over longer return intervals. Bernard and Thomas [1990] and Skinner [1993] report some evidence casting doubt on this possibility.
8. See Appendix A for implementation of the stochastic dominance conditions.
9. The Kolmogorov-Smirnoff (K-S) test uses on the maximum vertical distance between two cumulative density functions to test the hypothesis that they may be generated by the same distribution. The maximum distance between the empirical cumulative densities for portfolio 10 and 1 is 63.6 percent. The critical level for the Kolmogorov-Smirnoff test at the 1 percent level is 14.8 percent. (See Neave and Worthington [1988]). Hence, the K-S tests rejects the hypotheses that cumulative densities for portfolios 1 and 10 come from the same density function.
10. Stoll and Whaley [1983] estimate posted spreads for NYSE firms to be 0.69 percent for the largest market value decile, and 2.93 percent for the smallest firm decile. For market orders that execute immediately, these costs can be reduced by trading within the posted spread. Peterson and Fialkowski [1994] estimate that the mean *effective* spread is .20 percent for the largest firm decile, rises steadily to 2.2 percent in the next-to-smallest decile, and is 6.1 percent in the smallest decile.

References

Ball, R., "The Earnings-Price Anomaly." *Journal of Accounting and Economics* pp. 319–346, (June/September 1992).

Ball, R., and E. Bartov, "The Earnings Event-Time Seasonal and the Calendar-Time Seasonal in Stock Returns: Naive Use of Earnings Information or Announcement Timing Effect?" Working paper, University of Rochester (1993).

Ball, R., and P. Brown, "An Empirical Evaluation of Accounting Income Numbers." *Journal of Accounting Research* pp. 159–178, (Autumn 1968).

Bernard, V., "Stock Price Reactions to Earnings Announcements: A Summary of Recent Anomalous Evidence and Possible Explanations." In R. Thaler (ed.), *Advances in Behavioral Finance*. New York: Russell Sage Foundation, pp. 303–340, (1993).

Bernard, V., and J. Thomas, "Post-Earnings-Announcement Drift: Delayed Price Response or Risk Premium?" *Journal of Accounting Research* pp. 1–36, (Supplement 1989).

Bernard, V., and J. Thomas, "Evidence that Stock Prices Do Not Fully Reflect the Implications of Current Earnings for Future Earnings." *Journal of Accounting and Economics* pp. 305–341, (December 1990).

Bernard, V. J. Thomas, and J. Abarbanell, "How Sophisticated is the Market in Interpreting Earnings News?" *Journal of Applied Corporate Finance* pp. 54–62, (Summer 1993).

Bhushan, R., "An Informational Efficiency Perspective on the Post-Earnings-Announcement Drift." *Journal of Accounting and Economics* (forthcoming).

DeLong, B., A. Shleifer, L. Summers, and R. Waldman, "Noise Trader Risk in Financial Markets." *Journal of Political Economy* pp. 703–738, (1990).

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

Freeman, R., and S. Tse, "The Multi-period Information Content of Earnings Announcements: Rational Delayed Reactions to Earnings News." *Journal of Accounting Research* pp. 49–79, (Supplement 1989).

Holthausen, R., "Abnormal Returns Following Quarterly Earnings Announcements." *Proceedings of the Seminar on the Analysis of Security Prices* University of Chicago, pp. 37–59, (May 1983).

Jarrow, R., "The Relation Between Arbitrage and First Order Stochastic Dominance." *Journal of Finance* 41, 915–921, (1986).

Neave, H. R., and P. L. Worthington, *Distribution-Free Tests.* London: Unwin-Hyman, 1988.

Peterson, M., and D. Fialkowski, "Posted versus Effective Spreads: Good prices or Bad Quotes?" *Journal of Financial Economics* pp. 269–293, (June 1994).

Seyhun, N., "Can Omitted Risk Factors Explain the January Effect? A Stochastic Dominance Approach." *Journal of Financial and Quantitative Analysis* pp. 195–212, (June 1993).

Skinner, D. "Bid-Ask Spreads Around Earnings Announcements: Evidence from the NASDAQ National Market System," Working paper, University of Michigan (1993).

Stoll, H. R., and R. E. Whaley, "Transactions Costs and the Small Firm Effect." *Journal of Financial Economics* pp. 57–79, (June 1983).

Whitmore, G. A., and M. C. Findlay, *Stochastic Dominance*. Lexington, MA: Lexington Books, 1978.