Earnings Management?
Sample Selection Bias, Averaging, and Scaling Lead to Erroneous Inferences

Cindy Durtschi
Utah State University

and

Peter Easton
Center for Accounting Research and Education
University of Notre Dame*

December 2007

We thank Donal Byard and workshop participants at Baruch College for helpful comments on an earlier draft of this paper and students in the 2007 Limperg Institute "Advanced Capital Markets" class organized by Jan Bouwens at Tilburg University for helpful discussions during the early stages of this work.

*Corresponding author
305A Mendoza College of Business
University of Notre Dame
Notre Dame, IN 46556
peaston@nd.edu

Earnings Management?
Sample Selection Bias, Averaging, and Scaling Lead to Erroneous Inferences

Abstract

A vast literature following Hayn [1995] and Burgstahler and Dichev [1997] attributes the so-called "discontinuity" in earnings distributions around zero to earnings management. Durtschi and Easton [2005] make the point that the shape of these distributions is not *ipso facto* evidence of earnings management. Jacob and Jorgensen [2007] question this conclusion based on evidence from an alternate set of analyses, which consist of comparing the distribution of fiscal year (t) earnings with the distribution of a weighted average of six quarters of earnings ending in the third quarter of year t. We show that the Jacob and Jorgensen [2007] result is an almost inevitable consequence of their choice of benchmark; the average number of observations in three "as-if" years. This, and other evidence, supports our alternate conclusion that, rather than being evidence of earnings management, more plausible explanations for the shape of the Jacob and Jorgensen [2007] distributions are sample selection bias, scaling, and the effects of averaging. Distributions that are not affected by these research-design flaws do not exhibit patterns that suggest earnings are being managed to avoid losses.

1. Introduction

A vast literature following Hayn [1995] and Burgstahler and Dichev [1997] attributes the so-called "discontinuity" in earnings distributions around zero to earnings management.[1] Durtschi and Easton [2005] show that, although earnings management is one possible explanation, the evidence suggests that these discontinuities are driven by sample selection bias and scaling. Jacob and Jorgensen [2007] introduce an alternate methodology and conclude that their findings suggest that "while scaling and associated selection biases might contribute to the observed discontinuities, they are not primarily responsible for these discontinuities." We show that this conclusion is not supported by their analyses. Hence we reiterate the point in Durtschi and Easton [2005] that the shapes of frequency distributions of earnings metrics are not *ipso facto* evidence of earnings management, and assert that before one can draw conclusions regarding the presence/absence of earnings management, evidence beyond the shapes of these particular distributions must be bought to bear.

The Jacob and Jorgensen [2007] – hereafter JJ -- methodology consists of comparing the distribution of fiscal year (t) earnings with a benchmark distribution, which is the average of the distributions of three "as-if" years of earnings. Each of the three "as-if" years span four consecutive quarters: (1) ending with quarter one of fiscal year t, (2) ending with quarter two of fiscal year t, and (3) ending with quarter three of fiscal year t. In other words, the Jacob and Jorgensen [2007] benchmark is a weighted

[1] See, for example, Beatty, Ke, and Petroni, [1999], Beaver, McNichols, and Nelson [2004], Burgstahler and Dichev [1997], Brown and Caylor [2004], Collins, Pincus, and Xie [1999], Coulton, Taylor and Taylor [2005], Dechow, Richardson and Tuna [2003], Degeorge, Patel, and Zeckhauser [1999], Easton [1999], Holland and Ramsay [2004], Jacob and Jorgensen [2007], Kang [1999], Leone and Van Horn [2003], Phillips, Pincus, Rego, and Wan [2004], Kerstein and Rai [2007], Rego and Frank [2006], and Revsine, Collins, and Johnson [2005].

average of quarterly earnings over the six quarters ending in the third quarter of year t. JJ argue that earnings measured over these alternate years are less likely to suffer from the effects of managerial income manipulation than earnings of the fiscal year; they claim that differences between the distribution of fiscal year earnings and the distribution of the average of the three "as-if" years of earnings is evidence of earnings management.

We show that JJ's conclusions are not supported by their analyses for three reasons: (1) the JJ benchmark is not meaningful due to a) the characteristics of quarterly data and the fact that Compustat quarterly data are often restated, and b) the fact that averaging changes the shape of a distribution, (2) the JJ method, which relies on the availability of four continuous sets of quarterly data (four quarters per set) exacerbates sample selection problems,[2] and (3) deflation by beginning-of-year market capitalization contributes to the final shape of the (deflated) earnings distributions.[3]

Specifically, we argue that the weighted average earnings benchmark used by JJ is inappropriate because averaging across fiscal years does not make economic sense if, indeed, fiscal years are chosen for sound economic reasons and each year the firm faces different economic factors. We also note that the use of quarterly Compustat data is inappropriate because Compustat reports restated earnings, so that the data JJ analyze differs from the earnings data that was ostensibly managed and announced to the market. In addition, we argue that GAAP for interim financial statements differs from GAAP for

[2] The JJ analysis of earnings per share is based on Compustat Quarterly data item 27, which is calculated by Compustat and provided for each quarter that had three continuous prior quarters of income data and a moving average of shares used to calculate earnings per share. Thus the analyses based on a comparison of three years of "as-if" earnings and fiscal year earnings, requires seven consecutive quarters of data. We will explain this point further in a later section of the paper.

[3] Scaling also effects the conclusions from Kerstein and Rai [2007] published contemporaneously with JJ. We will elaborate on this point.

the annual report.[4] Finally we show that averaging smoothes the distribution of raw data by combining the idiosyncratic performance in one year with the idiosyncratic performance in the next year, and that averaging also draws observations toward the mean: therefore, differences between the distributions of the averaged "as-if" annual earnings and earnings of the fiscal year are likely due to the effects of averaging rather than due to earnings management.

To clarify and demonstrate the effects of averaging on the shape of frequency distributions, we introduce an analogy based on scores from the Australian Football League. Australian Rules Football has four quarters per game just as the fiscal year is divided into four quarters. We show that, by averaging team-scores across consecutive games, we produce the same effect as JJ observe when averaging net income (earnings per share) across consecutive fiscal years.

JJ point to the fact that there is a discontinuity in the distribution of their sample of fiscal year earnings, but no discontinuity in a distribution of averaged "as-if" earnings as evidence of earnings management. We show that this discontinuity in the net income distribution, which JJ create by severely censoring the data and deleting more small loss observations than small profit observations, is averaged away in the JJ benchmark. In other words, a discontinuity in net income, which is evidently *not* due to earnings management (rather it is the result of sample selection procedures), is eliminated when net income is averaged over three 'as-if' years. This further calls into question the validity of the JJ benchmarking procedure.

[4] We use the acronym GAAP to include Generally Accepted Accounting Principles and Generally Accepted Accounting Practices.

The remainder of our paper unfolds as follows. First, we discuss the JJ benchmark and address the question: is the distribution of an average based on data that spans two fiscal years a meaningful benchmark for comparing fiscal year data? Second, we discuss our sample. Third, we show how the original frequency distributions for net income (annual and quarterly), which show no evidence of an irregularity at zero are changed by the Burgstahler and Dichev [1997] – hereafter BD -- and JJ methodologies via sample selection, averaging, and deflation into distributions that show an irregularity at zero. It follows that the shapes of the BD and JJ distributions cannot be interpreted as evidence of earnings management. Fourth, we show that JJ's distribution of earnings per share is the result of sample selection, and their results are an artifact of a faulty benchmark. We end with a brief discussion of the JJ analysis of changes in deflated net income, earnings per share, and changes in earnings per share.

2. A Weighted-Average is not a Meaningful Benchmark

The methodological innovation in JJ is the comparison of the distribution of fiscal year t income and a benchmark distribution that is the average of three distributions of net income for "as-if" years; the first "as if" year ends with quarter one of year t, the second ends with quarter two of year t, and the third ends with quarter three of year t. This innovation is described in Figure 1. Figure 1 shows how averaging over three "as-if" years effectively assigns a weight of 1/12 to earnings of quarter two of year t-1, 2/12 to earnings of quarter three of year t-1, 3/12 to earnings of quarter four of year t-1, 3/12 to earnings of quarter one of year t, 2/12 to quarter two of year t and 1/12 to quarter 3 of year t.

We argue in this section that the distribution of earnings of fiscal year t cannot be reasonably compared to the weighted average of earnings from the last three quarters of year t-1 and the first three quarters of year t for two reasons. First, the start-date and end-date of a fiscal year are likely chosen for sound economic reasons making a comparison to annual earnings for a year other than the fiscal year questionable. Second, using a weighted average of earnings that span two fiscal years will mitigate the effects of idiosyncrasies in the individual years, thus affecting the shape of the earnings distribution.[5]

2.1. The Fiscal Year is Chosen for Sound Economic Reasons

Unlike the JJ weighted average "as-if" earnings, fiscal year earnings are for a period that is chosen for sound economic reasons. Managers, investors, creditors, labor unions etc., tend to focus on fiscal period earnings, rather than earnings of "as-if" years that end at dates other than the fiscal-year end. For example, compensation packages are generally adjusted annually based fiscal year performance; in addition, investment, marketing, production, and inventory decisions and contracts with suppliers are based on a fiscal year. Because of this, the fiscal year, unlike the other annual periods, has a

[5] Although Figure 1 captures the key aspects of the JJ method, the weighting in the calculation of the distribution of the average as-if earnings is more complicated. More details follow. The fiscal year distribution is computed as the sum of the four quarters of net income in a given fiscal year. This sum is then deflated by beginning-of-year market capitalization and observations are classified into bins that are 0.005 wide. The number of observations in each fiscal year bin is then divided by the total number of fiscal year observations and the resulting percentage is compared to the percentage of observations in the benchmark bin of the same size. To create the benchmark bin, JJ first create three "as-if" years of earnings. Notice that calculation of the as-if years creates an averaging effect; but the JJ method also creates other effects due to averaging. Each observation of "as-if" earnings is deflated by the beginning-of-as-if year market capitalization. JJ classify these deflated as-if year observations into bins that are 0.005 wide. The number of observations in each bin for the three "as-if" years is divided by the total number of observations for the respective as-if year to determine the percentage of observations per bin. These percentages are averaged across the three "as-if" distributions, and that average is used as the benchmark against which to compare the percentage of observations in the same bin in the fiscal year earnings distribution.

natural start-date and end-date, which is consciously chosen by management, but may (likely will) have nothing to do with earnings management.

As an illustration of the existence of natural year start-dates and end-dates, 60 percent of retail department and variety stores have a fiscal-year end of late-January after the completion of the busy holiday season when their inventory is depleted. They make investment decisions (particular regarding inventory), re-stock their shelves, and begin a new cycle where the outcome of their investment decisions, advertising strategy, etc. will be realized over the forthcoming fiscal year culminating in the next holiday season when the success or failure of their investment decisions becomes known. By contrast, only 17 percent of retail grocery stores have a January fiscal year end, as they do not have the same degree of seasonality in their sales. Furthermore, firms are rarely satisfied with merely repeating the success or failure of the previous year. Firms plan toward, work toward, and invest in improvement, learning from the prior year's experience. In addition, the information, economic, and competitive environments differ from year to year.

All of these reasons suggest that a weighted average of earnings over two fiscal years is not a meaningful benchmark to use as a comparison to the distribution of earnings of a single fiscal year.

Other issues that are inherent in using quarterly numbers from Compustat lead us to further question the use of a weighted average of quarterly observations spanning two fiscal years as a benchmark for evaluating the distribution of fiscal year earnings. First, under the integral theory of interim reporting, quarterly statements are considered an installment to the annual report. While interim period financial reports are reviewed by

auditors as they audit the annual report, they are compiled using a different set of guidelines; thus, recognition of revenue and expenses may be affected by judgments that are not allowed when issuing annual results. For example, expenses that would normally be charged to operations in one period for annual accounting purposes can be deferred and expensed over several interim periods based on allocation using sales volume, production levels, or some other basis. Cost of goods sold, generally the largest expense in the quarterly income statement, may be computed based on estimated gross profit. Revenues, on the other hand, must be recognized and reported in the period in which they are earned and cannot be deferred to other periods to, for example, present a more stable revenue stream. This means that revenue from seasonal businesses cannot be manipulated to eliminate seasonal trends (APB 28; Baker, et. al. 2007). In addition, numbers in interim reports are often restated at the annual audit. In short, the distribution of the weighted average of quarterly income taken across years is likely not an appropriate benchmark for evaluating the distribution of fiscal year income.

2.2. Averaging Changes the Shape of the Distribution

Averaging net income across fiscal years has two effects on the distribution of averaged data; averaging smoothes the irregularities in the distribution and averaging draws observations toward the mean. The first point, that averaging smoothes a distribution, is straightforward; averaging across several periods tends to offset idiosyncrasies from one period against idiosyncrasies of another period.[6] In order to use

[6] The devils advocate may argue that earnings management may be one of these idiosyncrasies. We agree; but we hasten to add (1) earnings management is just one of many factors that may differ across observations, and (2) researchers should not assume the existence of earnings management if they have not ruled out other, possibly more plausible, explanations.

"as-if" years that combine quarters from one year with quarters from a prior year as the benchmark for understanding the distribution of fiscal year net income, one must assume that the business and economic environment of one fiscal year is essentially the same as that of the next, that there is no growth, that the competitive environment is the same, that the economic environment is the same, that innovations are evenly spread throughout the years, that supply, distribution and sales issues are evenly spread over time, that various contracts for compensation, contracts with suppliers etc. do not change between fiscal years, and that the "installments" estimated for one year apply equally as well to the following year.

While some may argue that fiscal years are an artifact, with no inherent meaning, it is hard to imagine in a complex world that any twelve months are precisely the same as the prior twelve months for any firm. It follows that calculating a weighted average performance measure as in JJ, where the idiosyncrasies from one year can offset idiosyncrasies of the other year, creates a smoothing effect. We use an analogy of team-scores from the Australian Football League (AFL) to demonstrate this point.

2.3. The AFL Analogy

Similar to a business reporting cycle, each football match has four quarters. If we assume the score for each quarter of a football match is analogous to quarterly earnings of the firm, the score for the match (i.e., the sum of the scores for each of the four quarters) is then analogous to the annual net earnings.

As in a firm at the beginning of the fiscal year, the football coach prepares the team with a strategy for each match, he has an opportunity to modify this strategy during

the match in light of the competition encountered, and he has the goal of maximizing the match score. Similarly, corporate management prepares the firm (the production facility, the choice of products or services, the sources of inputs, etc.) for the forthcoming business year; there are opportunities to modify the business strategy throughout the year in light of economic events, and the goal is to maximize profits. Just as the football match has a natural beginning and a natural end, so does the business (fiscal) year and that end is a time to tally up and make judgments on the performance of the team/firm, and to think about strategy for the next match/year.

JJ add earnings from quarters of the previous year to the earnings of the current year to provide a measure of income for a period that is equal to one year of consecutive quarters (see Figure 1). We follow the exact same pattern to create a rolling set of "as-if" match-scores which include four consecutive quarters (i.e., for the four consecutive quarters that end with quarter one of the current match, with quarter two of the current match, and with quarter three of the current match). We then average the three "as-if" matches in the same way that JJ average their three "as-if" years and use the average to benchmark expected scores for the current match.

When seen in the football context, one might say that we are not adding apples and apples – the competition, the makeup of the team, the strategy, may differ from match to match. We posit the question; does adding net income across fiscal years make any more sense? Does not the makeup of the firm, the strategy of the firm, the economic/competitive environment of the firm, etc. change from fiscal year to fiscal year? In combining, either, years or matches, idiosyncratic superior performance in the previous year/match may be averaged with idiosyncratic inferior performance in the

current year/match. This averaging may (likely will) result in a change in the shape of the distribution of scores/earnings.

Whether one believes this analogy is sufficiently sound is, in some ways, irrelevant as the point of this exercise is to show the inevitable effects of averaging on the shape of frequency distributions; the distribution is smoothed and observations are drawn toward the mean.

We show that while scores for AFL football matches are not managed as is alleged with earnings, we see similar effects when we compare actual match scores with the average of three "as-if" match scores as JJ show when comparing fiscal-year earnings with the average of three "as-if" year earnings.

Our data comes from 22 rounds of matches among the 16 teams that have comprised the Australian Football League (AFL) over the years 1997 to 2006.[7] In other words we analyze data for 220 matches for each of 16 teams (a total of 3,420 matches). The same 16 teams have comprised the AFL for these 10 years.

Scores are made in two different ways in Australian Rules Football: *goals*, when the ball is kicked between the two goal posts are awarded six points, and *behinds*, when the ball is kicked between the goal post and a behind post (which is placed outside the goal post) or if the ball hits a goal post, or if an opposition player sends the ball between the goal posts by touching it with any part of the body, are awarded one point. The winner is the team who has the highest total score by the end of the match. If the scores are even at the end of play, then the match is a draw.

[7] 1997 represents a natural year to start the analysis because the teams comprising the AFL have been the same from that year until today.

Figure 2, Panel A demonstrates the effect of averaging scores across football matches. The dark line, which represents the average score of three "as-if" matches, is noticeably smoother than the light line which shows the actual distribution of match scores. This smoothing effect of averaging is almost inevitable, outstanding events/scores in one match will be averaged with outstanding scores in another.

The second effect of averaging is to draw observations toward the mean. This effect is also demonstrated with the AFL scores. Rather than looking at raw scores, we put scores in six point bins (the amount of a goal). The results are shown in Figure 2, Panel B. The dark line, which plots the average of three "as-if" matches, is more peaked than the light line, which plots the regular match total.

2.4. Differences between the Benchmark Distribution (Averaged Data) and the Distribution of Raw Data are not Evidence of Earnings Management

We now move from the football analogy to net income as used by JJ and show that the differences between the benchmark (the distribution of averaged, deflated net income) and the distribution of deflated, fiscal year net income may not be used as evidence of earnings management. We begin by discussing our sample selection criteria and then we show that the distribution of all available observations of annual fiscal period net income available on Compustat in our sample period does not (in seeming contradiction to BD and JJ) exhibit an irregularity around zero net income. Next we apply the JJ (and/or the BD) sample selection criteria to this distribution and show that it then exhibits a distinct irregularity. In other words, there is a discontinuity in this distribution that is due to sample selection bias. We then show that the JJ weighted average of these earnings does not exhibit this discontinuity. Again this demonstrates

that the JJ weighted average smoothes the earnings distribution and therefore is not an appropriate benchmark for evaluating the distribution of fiscal year earnings.

3. Sample Selection, Description of Variables, and Evidence of Sample Selection Bias

For each set of analyses, we obtain the required data from the Compustat files spanning 1977 to 2006. This period spans periods used by BD (1976 to 1994), JJ (1981 to 2001), and Durtschi and Easton [2005] – hereafter DE -- (1983 to 2003) as well as the more recently available years. We replicate the relevant findings in BD, JJ and DE to ensure the results hold in our sample. To be consistent with all three papers, we eliminate regulated firms as in BD.

3.1. Net Income Deflated by Beginning-of-Year Market Capitalization

The JJ analysis uses a benchmark of the weighted average of quarterly data from two consecutive years, computed as described in Figure 1, to provide a distribution that is compared with fiscal year earnings. We first discuss the distribution of fiscal year net income shown in JJ and show that the shape of this distribution is (contrary to their claim) an artifact of their sample selection procedure. Then we compare their distribution, which exhibits a discontinuity that is not due to earnings management, with the JJ benchmark weighted average net income.

Figure 3 replicates DE Figure 3 (with our larger sample period) and shows the frequency distribution of all observations of reported un-deflated annual net income (annual Compustat data item 172) for 1977 to 2006. We partition firms into $100,000 intervals of net income as in DE, JJ, and Dechow, Richardson, and Tuna [2003]. As in

DE there is no discontinuity in the distribution around zero.[8] In other words, absent the severe sample censoring due to the criteria imposed by JJ, there is no discontinuity.

3.2. Sample Selection Bias

In Figure 4, we focus on the numerator of the JJ fiscal year analysis, which is the sum of net income (Quarterly Compustat data item 69) for the four quarters of a fiscal year. Absent differences in the sample, the distribution of this net income number will be the same as the distribution of annual net income shown in Figure 3. However, the JJ sample selection criteria severely censor the data; an observation only appears in the JJ sample if it has four consecutive quarters of data in a fiscal year on the quarterly Compustat file. This restriction causes 16 percent of the total sample to be eliminated. Even more striking (see Figure 4, Panel A) is that these lost observations are not spread equally over the distribution; instead, they are concentrated in the smaller net income numbers and more negative income observations are deleted than positive net income; 51 percent of the smallest negative net income numbers (those between zero and -$100,000) are eliminated by this sample selection criterion, while only 42 percent of the smallest positive net income numbers are eliminated. In addition, requiring four consecutive quarters of data has induced a discontinuity into the distribution *which likely has nothing to do with earnings management*.[9]

---

[8] JJ mention (p.13) that the peak of the distribution for un-scaled net income is at zero; however, for the annual net income data item (annual Compustat data item 172), there are only 185 observations with net income of zero (as compared with 3,696 observations in the bin -0.1, and 5,220 observations in the 0.1 bin). JJ are silent on whether they include net income of zero in the positive or the negative net income bins. We follow DE and remove the small number of observations with net income of zero, classifying them as neither positive nor negative.

[9] Is it possible that Compustat has four quarters of consecutive data only for firms that have managed earnings (i.e., non-managers do not have four quarters of consecutive data)? This seems unlikely. If this is

Requiring four quarters of consecutive data is not the only selection criterion imposed by JJ. JJ deflate net income by beginning-of-year market capitalization to obtain a distribution that is similar to BD, Figure 1. When the additional data requirement of the availability of beginning-of-year market capitalization is included, even more observations are lost.

Figure 4, Panel B shows the effect of the combination of the two JJ sample selection criteria; requiring four quarters of data and beginning-of-year market capitalization. The discontinuity has become more evident as the sample selection bias further effects the distribution. Now, 73 percent of the observations in the smallest negative net income bin are deleted while 65 percent of the observations in the smallest positive net income bin are deleted.

We have shown in this section that sample selection bias explains the discontinuity in the distribution of fiscal year net income. In the next section we compare this distribution (where there is a discontinuity created without earnings management) with the average of three "as-if" years of earnings obtained via the same (JJ) sample selection criteria.

3.3. Averaging Net Income Across Years

To highlight the smoothing effect of averaging the net income data used by JJ, Figure 5 shows the effect of averaging in the numerator of the variable (deflated net

---

the assertion, more evidence must be provided before we can conclude that the shape of the distribution is due to earnings management rather than sample selection bias.

income) analyzed by JJ.[10] Panel A shows the distribution of fiscal year net income (Quarterly Compustat data item 69), after imposing the JJ selection criteria (the light line).[11] The dark line is the weighted average of fiscal year net income for years t and t-1. This illustrates the effect of averaging; the plot of average of two fiscal years of net income is smoother and the discontinuity created by JJ's sample selection criteria (not by earnings management) is no longer observed. Figure 5, Panel B shows the difference between fiscal year net income in year t and the average of two consecutive fiscal years of income. As in the football analogy, there are significant deviations from the benchmark.

Figure 5, Panel C illustrates the way that averaging draws the distribution toward the mean. We replicate the JJ deflated net income analysis by first recreating the distribution of deflated fiscal year net income (JJ Figure 1a) and, on the same figure, plotting the weighted average across two years of net income. Our analysis differs from JJ in only one respect. In determining the weighted average net income, we deflate each "as-if" net income number by the same denominator as the fiscal year (JJ deflate each "as-if" net income by its own beginning-of-as-if-year market capitalization). We make this comparison because we note that it is possible that some of the effect observed by JJ may be due to the fact that they average across ratios where both the denominator (beginning-of-as-if-year market capitalization) as well as the numerator (net income) change.

[10] The focus of most of JJ's analysis is on the distribution of the ratio of net income to beginning-of-year market capitalization. We analyze the numerator of this ratio first and then we consider the effects of scaling by market capitalization.

[11] JJ's fiscal year net income is the sum of four quarters of quarterly Compustat data item 69.

As in the football analogy, the distribution of the weighted average net income (the line) is more peaked than the distribution of fiscal year, non-averaged net income (the histogram). Panel D shows the differences between these two distributions. The graph in Panel D is noticeably different from JJ's Figure 2, where the deflator differs across 'as-if" years.

In sum, the distribution of the weighted average across two years of net income is smoother and drawn toward the mean. It follows that this distribution of the weighted average does not provide a reliable "benchmark" against which to compare non-averaged (fiscal year) net income; thus, such a comparison cannot be used to demonstrate the possibility of earnings management.

So far we have shown that 1) sample selection bias causes a discontinuity in the distribution of the net income numbers used in the JJ analysis, and 2) using the average of three distributions as a benchmark is unreliable. We will now complete the analysis of the reasons why the JJ distribution exhibits an irregularity at zero by analyzing the effects of deflating net income by beginning-of-year market capitalization.

4. The Effects of Scaling

We begin by re-iterating the point made by DE that it is extremely unlikely that firms manage earnings relative to beginning-of-year prices. For this reason we need to be concerned if deflating by beginning-of-year prices distorts the shape of the distribution.

The fact that Figure 5, Panel D differs from JJ's Figure 2 based only on a difference in the denominator leads to our last step -- creating the BD and JJ figures (which they offer as evidence of earnings management) from the actual distribution of net

income. To begin we look back at Figure 3. This is the actual distribution of fiscal year net income for the complete sample of observations. This distribution shows no discontinuity to the left of zero. Next, after applying the JJ sample selection criteria, we observe a discontinuity in the distribution (see Figure 4, Panel B). However, this distribution does not have a "mini" peak immediately to the right of zero as in JJ Figure 1a and BD Figure 1 (replicated in our Figure 6). In order to show how scaling leads to this mini peak, we systematically analyze the net income observations beginning with those in the vicinity of zero, progressively moving in our analysis to the tails of the net income distribution. But first, we examine differences in the denominator, market capitalization, to the left and to the right of zero.

4.1. Scalers Differ Between Fiscal Year Loss Observations and "as-if" Year Loss Observations

DE report that stock prices are less for each dollar amount of loss than for the same dollar amount of profit. In a similar vein, Figure 7, Panel A shows, for each $100,000 net income interval, the 25th-percentile, median, and 75th-percentile of distribution of beginning-of-year market capitalization (i.e., price – Quarterly Compustat data item 14 – times number of shares outstanding – Quarterly Compustat data item 61) for all the observations that are included in the JJ analyses. Figure 7, Panel A is similar to Figure 6 in DE, which shows the distribution of prices for each cent of earnings per share.

Figure 7, Panel A clearly demonstrates that, like prices, the market capitalization of firms that have a net loss tends to be smaller than for firms reporting the same dollar amount of profit. For example, firms/observations with annual net income in the smallest negative bin, -0.1 (between zero and -$100,000) have a median market capitalization of

$4.46 million while observations in the smallest positive net income bin (0.1) have a median market capitalization of $5.28 million. As an illustration of the effect of deflation, -0.1/4.46 = -0.0224 while +0.1/5.28 = 0.0189; this demonstrates the point made by DE that scaling will cause observations with positive net income to be drawn toward zero while observations with negative net income will be "pushed" away from zero.

Figure 7, Panel B separates the beginning-of-year market capitalization for the "as-if" years ending in quarters one through three from the beginning-of-*fiscal*-year market capitalization. This figure captures the spirit of DE footnote 32 where it is noted that prices may be different at the beginning of the fiscal year than they are at the beginning of the three JJ constructed years. Figure 7, Panel B focuses on market capitalization around zero net income, rather than distribution of stock prices around zero earnings per share (as in DE) and shows that there is a difference in the denominators to the left of zero. If firms report a negative net income, their market capitalization tends to be larger at the beginning of the "as-if" years than it is in the fiscal year. The effect is that when JJ compare deflated fiscal year income to the average of the three deflated "as-if" year's income, deflation will pull fiscal year observations even further away from zero than it will pull the "as-if" year observations.[12] This will contribute to the JJ conclusion that there are too few observations of small negative net earnings in the fiscal year. By contrast, there are no such distinct differences to the right of zero in the positive net income deflators.

---

[12] For example, net income for an observation falling in the first negative net income interval (-0.1) in the "as if" years would be divided by a market capitalization of 4.77, and be pulled to the 5th bin to the left of zero. The same net income number in the fiscal year would be divided by 3.19 market capitalization and end up in the 7th bin to the left of zero. Net Income for an observation in the smallest positive bin (0.1) would be divided by 5.21 if it was in the "as if" years, and 5.45 in the fiscal year. In either case the positive net income would lie in the 4th bin to the right of zero in either the BD or JJ distribution of deflated net income.

4.2. Detailed Analysis of the Effects of Scaling

Figure 8 illustrates the effect of scaling. In this figure, we begin with the annual fiscal-year net income sample that meets the JJ sample selection criteria (see Figure 4, Panel B). We first consider the effect of deflation on net income observations in the vicinity of zero and progressively move outward in the net income distribution. We show the effect described by DE as well as the tendency for observations in the negative tail of the net income distribution to remain toward the tail of the distribution while the observations in the positive tail of the net income distribution are drawn toward zero creating the dip in the deflated net income distribution immediately to the left of zero and the peak immediately to the right.

Figure 8, Panel A is the frequency distribution of deflated net income for all observations with net income between -$100,000 and $100,000. The interval width in this distribution is the same as in JJ and BD. Forty-three percent (794/1842) of the observations with positive net income are in the deflated-net-income bin (0.005) immediately to the right of zero, while only 29.6 percent (299/1008) of the observations with negative net income are in the deflated-net-income bin (-0.005) immediately to the left of zero.[13]

In Figure 8, Panel B, we repeat these analyses but for net income in the range -$1,000,000 to $1,000,000. The observations from Panel A (net income between -$100,000 and $100,000) are shown as the light portion of the frequency distribution in Panel B; the new observations added to the histogram in the darker shade. Twenty-five percent of the 10,968 observations with positive net income are in one of the four bins

[13] This analysis is a replication of the analysis shown in DE Figure 5, Panel C. We extend this analysis to include all net income intervals.

immediately to the right of zero, while only 18 percent of the observations with negative net income are in the four bins immediately to the left of zero. Again, we see that deflation tends to draw observations of positive net income toward zero, while moving observations of negative net income away from zero.

In Figure 8, Panel C, all observations in Panel B are now in a light shade, while the new observations which come from net income between -$5,000,000 and $5,000,000 are added in the darker shade. Here, when we add observations between +/- one million and +/- five million dollars, we see the first hint of the discontinuity (indentation) from the JJ fiscal year graph in the bin immediately to the left of zero. In addition, the mini-peak of the distribution immediately to the right of zero begins to become evident as observations from the positive tail continue to be drawn toward zero.

In Figure 8, Panel D, all observations included in Panel C are shown in a light shade and the additional net income observations in the range -$10,000,000 to $10,000,000 are added in the darker shade. The dip in the interval immediately to the left of zero is now more distinct, while observations continue to pile up in the interval immediately to the right of zero; very few large negative net income numbers are drawn to the interval immediately to the left of zero.

Figure 8, Panels E and F complete the picture as the net income observations between -$20 million and $20 million are added in Panel E, then all remaining observations (those where net income is less than -$20 million or greater than $20 million) are added in Panel F. What becomes apparent from this figure is that, after net loss exceeds -$5 million, very few loss observations have a market capitalization that is high enough to put them in deflated-net-income bins close to zero, while an abundance of

observations with net income greater than $5 million have a market capitalization that is high enough to draw them toward zero.[14]

To summarize these analyses; we show that observations of very high net income are drawn by deflation to the smallest positive BD and JJ interval; this tendency is *much* less evident for observations of large losses. This effect of deflation leads to the appearance that firms tend to manage earnings to avoid a loss. But the results suggest that scaling, which draws large positive income observations toward zero, causes the kink rather than management of earnings to avoid a loss.

5. Analysis of Earnings per Share

JJ also examine the distribution of earnings per share in a manner that is very similar to their analysis of deflated net income. However, rather than using reported earnings per share, they use Compustat quarterly data item 27, which is described by Compustat as basic earnings per share excluding extraordinary items, twelve months moving. According to the Compustat manual this item represents basic earnings per share applicable to the last 12-month period and

> "…is calculated by adding four quarters of available for common and dividing by the 12-months moving shares figure. At year end, this figure will be the same as the

[14] This point is also particularly pertinent to the analyses in Kerstein and Rai [2007] who examine movements among the four BD/JJ intervals around zero. Movements (to the right, to the left, or no movement at all) constitute their coded dependent variable in a logistic regression. Kerstein and Rai [2007] attempt to control for the scale effect identified in DE and elaborated on herein by adding log of market capitalization as another independent variable; this may help somewhat but the effects of scale are likely to remain because the relation between the magnitude of net income and market capitalization, which affects the classification of the data, likely will not be captured by simply adding log of market capitalization as another explanatory variable. Perhaps the best way to overcome the scale issue is to repeat the study for the four or five earnings per share intervals around zero; as DE point out, firms do not manage price-deflated earnings and Kerstein and Rai's [2007] hypotheses are based on a discussion of earnings – not deflated earnings. It would be interesting to see the results of a replication of Kerstein and Rai [2007] based on earnings per share.

> annual earnings figure reported to shareholders, in interim quarters this figure will equal (within four cents deviation) the sum of four quarters of quarters of Earnings per Share (basic) – Excluding extraordinary items."

The Compustat manual also points out that quarterly earnings per share is carried on a restated basis, such that, when a firm reports for a new quarter and at the same time reports different data than was originally reported for the corresponding quarter of the prior year, the data for that corresponding quarter from the prior year is changed. Compustat further point out that this handling tends to affect data item 27 for prior quarters "in that from one to three quarters of a four-quarter total may contain restated data, whereas the other quarters may not." The manual notes a similar problem when a company acquires another and includes either, six, nine or twelve months of an acquisition in a particular quarter; this will tend to over or understate certain observations of data item 27.

Finally, Compustat points out that there will be sample selection issues when using quarterly data item 27 as there will be gaps in the earnings series of some companies, generally because there are periods in which quarterly earnings are not reported, such as firms that report on a semi-annual or annual basis rather than quarterly, also where there are mergers, or a firm changes fiscal years.

While JJ's use of quarterly data item 27 enables them to continue the spirit of their comparison of net income with a weighted average across two years, it is apparent from the outset that this variable is fraught with difficulties that would create problems when drawing conclusions regarding a nuance such as earnings management around zero; the fact that these data are re-stated and there is sample selection bias are two of the greatest concerns.

In addition, because, as JJ note, there are a significantly larger number of cases where quarterly data item 27 is missing for the first three quarters of the fiscal year than where it is missing for the fourth quarter, they restrict their sample to firm-years where this data item is available for all four quarters of the year. The result is that many fourth quarter observations, those that can be compared to a reported annual number, are deleted from the sample. Since it is more likely that profitable firms, which tend to be larger, will have data for all four quarters, it is possible that, by restricting their analysis to firms that have four quarters of data, JJ have eliminated many more loss firms and, by sample selection, created a distribution in which there is an irregularity at zero, which may be erroneously interpreted as evidence of firms managing earnings to avoid losses. We illustrate the effects of sample selection bias in the next section.

5.1. Sample Selection Bias in Data Item 27; Earnings per Share, Moving Average

5.1.1. JJ Require Seven Quarters of Data for Inclusion in Their Sample

It is difficult to measure the complete effect of sample selection bias for quarterly Compustat data item 27 because of the way each data item 27 is created. Figure 9 describes the creation of data item 27, which sheds some light on this. Notice in Figure 9, that each observation of quarterly data item 27 is created by "adding four quarters of available for common and dividing by the 12-months moving shares figure." This means that, to meet the JJ sample selection criteria, where an observation only appears if there are complete observations for quarterly data item 27 for all quarters in a fiscal year, there must be seven consecutive quarters of income figures available for any particular firm. So, for example, if there was no income data for quarter three in the prior year the observation would be deleted from the JJ analysis.

5.1.2. Comparison with Data from the Annual Files

We can objectively determine which observations of quarterly Compustat data item 27 for quarter four (which is fiscal year earnings) are missing from the Quarterly Compustat file because there is a comparable data item on the Annual Compustat file (data item 58). Figure 10, Panel A shows the frequency distribution of annual earnings per share (data item 58); Panel B, shows those observations that survive to be in the JJ sample. In Panel B, 30 percent of the smallest negative earning per share (-$0.01) are missing and 29 percent of the smallest positive earnings per share ($0.01) are missing. In addition, when the JJ selection criteria are used (i.e., there must be four consecutive quarters of quarterly data item 27) 23 percent of the existing quarter four observations of quarterly data item 27 equal to -$0.01 earnings per share are eliminated, and fewer (20 percent) of existing quarter four observations of quarterly data item 27 equal to $0.01 earnings per share are eliminated. This asymmetric deletion of observations could make it appear that there are fewer than expected observations of earnings per share of -$0.01 and more than expected observations of $0.01.

It is impossible to do a similar analysis of observations that are missing from the "benchmark" JJ weighted average because there are no comparable annual earnings per share data on the Annual Compustat file. What we do know is that there are considerably fewer observations of quarterly data item 27 for quarters one through three than for quarter four. An examination of the items which do exist, but then are not used in the benchmark calculation because they are not part of a string of four consecutive quarters, provides little useful information as there is no reported number to use as a comparison.

5.1.3. Summary

In summary, there are two things that suggest that JJ's sample selection criteria seriously affect the shape of their observed distribution of earnings per share. First these data are restated, and given the caveats made by Compustat, that it may be restated in some quarters of a four quarter period and not others, the data are unreliable to use in a study of managed earnings; second, the JJ sample selection criteria remove a larger proportion of firms reporting zero and small losses than firms reporting small profits.

5.2. Averaging of Earnings per Share

Possibly the greatest difficulty with using quarterly Compustat data item 27 comes not in quarter four, where we, and Compustat, can at least compare it to annual earnings per share reported to shareholders, but in the interim quarters, which JJ average together and use as a benchmark to determine the expected shape of the distribution of annual earnings per share. Recall that Compustat states that, in interim quarters, quarterly data item 27 will equal (*within four cents deviation*) the sum of four quarters of basic earnings per share, excluding extraordinary items and that one to three quarters of a four quarter total may include restated data whereas the others may not. It is these imprecise numbers that JJ average together to create the benchmark against which to compare fiscal year quarterly data item 27. Because of the many problems with data item 27, detailed analysis it is not meaningful. That said, averaging has the same effect on quarterly data item 27 as it does on the JJ net income numbers: averaging smoothes the distribution and draws observations toward the mean.

Figure 11 shows the effect of averaging the observations that survive the JJ sample selection criteria and, therefore, can be included in the "benchmark" calculation. The dark line is the distribution of the JJ average of quarters one through three, the light line is distribution of data item 27 for quarter four. Once again, the averaged line is more smooth, and drawn toward the mean as compared to the fiscal-year line. The differences between the dark and light line, which are due to averaging, are erroneously used as evidence of earnings management in JJ.

6. Analysis of Change in Net Income

JJ do not tabulate nor plot their evidence regarding changes in net income. They briefly state that "similar patterns appear in the histogram of earnings changes." We find no discontinuities in the distribution of un-deflated change in net income even after using the JJ sample selection criteria; the discontinuity only appears after deflation. We illustrate this point in Figure 12 where we repeat our analysis of deflation on net income, but here for changes in net income. Once again, we take increasingly larger intervals of changes in net income, deflate and plot them.

Figure 12, Panel A shows the distribution of deflated change in net income for observations of change in net income between -$50,000 and $50,000. Seventy-seven percent of the negative observations are in the four change bins closest to zero, while 78 percent of the positive observations, once deflated, are in the four bins closest to zero; there is no discontinuity. In Panel B we add changes in net income in the range +/- $100,000 and so on until all changes have been deflated and added to the distribution. It is interesting to note that only after the change in net income is greater than $1,000,000

does the distribution stop being symmetrical. The slight discontinuity to the left of zero does not appear until changes in net income greater than $5,000,000 are added to the distribution.

7. Analysis of Changes in Earnings per Share

JJ do not tabulate results for their analysis of change in earnings per share but they note that they perform a "similar analysis" on changes in primary earnings per share. They conclude that there are significantly fewer than expected decreases of one cent in fiscal year earnings per share from the previous year. Considering the issues raised previously about the use of quarterly data item 27, which requires four quarters of data, the sample selection issues, and the imprecision inherent in this data item, we question whether any conclusions can be drawn for changes in quarterly data item 27.

In our analysis (un-tabulated) we find no discontinuity in the distribution of changes in quarterly data item 27 and we note that even using the JJ sample selection criteria, there is no discontinuity in the distribution of change in this variable. Therefore, we posit that the JJ results are due to averaging. To demonstrate this point, we first note that, to create the "difference" figure, JJ require seven quarters of quarterly data item 27 (i.e., 11 consecutive quarters of earnings per share data).

We add just one more quarter in our analyses so that we can compare change in fiscal year earnings per share (from t-1 to t) to the average of that change and the change from fiscal years t-2 to t-1. The results of these analyses are shown in Figure 13. The averaging effects that were observed in the analyses of earnings are more pronounced in these analyses of earnings changes. The difference to which JJ refer when stating that the

chances of observing the sequence of expected frequencies as being 1 in 131,072 is, in effect, the difference between the two lines in this figure (the dark line representing the distribution of the *average* change in annual earnings per share and the light line representing the change in annual earnings per share). In other words, we have demonstrated that the JJ effect may be seen when we take average changes in fiscal-year earnings per share. This result suggests that the JJ comparison to the changes in "as-if" year earnings per share reflects the effects of averaging; their result should not be interpreted as evidence of earnings management.

8. Conclusions

In this paper we have examined conclusions regarding earnings management that JJ draw from comparisons of distributions of earnings and earnings changes. We show that the JJ results are driven first by creating a faulty benchmark, which not only makes little economic sense, but, which is unreliable because of the characteristics of quarterly data on which they base their analyses. We note that quarterly data are often, and inconsistently, restated; thus using such data to find a nuance such as earnings management to avoid a loss is unreliable, at best. In addition, and perhaps most importantly, we have shown how averaging changes the shape of a distribution by smoothing and drawing observations toward the mean; therefore the JJ benchmark, which is a weighted average, will show deviations from expectations that are not due to earnings management, rather they are an artifact of comparing a smoothed (averaged) distribution to a non-smoothed distribution.

Next we show that the JJ method, which relies on the availability of four continuous sets of quarterly data (four quarters per set), exacerbates sample selection problems. We show how their sample selection criteria eliminate more small loss observations than observations of small profits. We reiterate the point made in DE that, if any sample selection criterion leads to the deletion of more observations with small losses than observations of small profits, a comparison across zero or small losses to small profits cannot be used as evidence of earnings management.

Finally, we demonstrate how the choice of deflator contributes to the final shape of the (deflated) earnings distributions. We show that market capitalization for observations of small losses tends to be less than for observations of small profits; we show how this difference in market capitalization tends to draw observations of small losses away from zero while profit observations tend to be drawn toward zero. We show that there are many firms with very large positive net income that are moved by deflation to the bins immediately to the right of zero. We note that it does not seem reasonable to make the implicit assumption (in the extant literature) that these firms, which have net income of $1,000,000, $5,000,000 or $20,000,000, end up in these bins because they are managing their income (deflated by market capitalization) to avoid a loss.

We also show that market capitalization (the deflator) for loss firms differs between the JJ benchmark years and the fiscal year further contributing to the results JJ report regarding the differences between their deflated benchmark earnings and deflated fiscal year earnings.

We conclude, as do DE, that the observed shapes of earnings distributions around zero is not *ipso facto* evidence of earnings management. In this paper we have provided

more plausible explanations for the shape of these distributions; sample selection bias, scaling, and averaging. The distributions that are not affected by these research-design flaws do not exhibit patterns that suggest earnings are being managed to avoid losses.

REFERENCES

Accounting Principles Board opinion number 28, *Interim Financial Reporting* (1973).

Baber, W., and S.Kang. "The Impact of Split-adjusting and Rounding on Analysts' Forecast Error Calculations." *Accounting Horizons 16,4* (December 2002): 277-289.

Baker, R., V. Lemke, T. King, *Advanced Financial Accounting* Beatty, 7th Edition, New York, New York. 2007.

Beatty, A.; B. Ke; and K. Petroni. "Earnings Management to Avoid Earnings Declines Across Publicly and Privately Held Banks." *The Accounting Review* 77 (July 2002): 547-570.

Beaver, W.; M. McNichols; and K. Nelson. "An Alternative Interpretation of the Discontinuity in Earnings Distribution." Working paper, Stanford University, 2004.

Brown, L., and M. Caylor. "A Temporal Analysis of Thresholds: Propensities and Valuation Consequences." *14th Annual Conference of Financial Economics and Accounting*, February 18, 2004.

Burgstahler, D., and I. Dichev. "Earnings Management to Avoid Earnings Decreases and Losses." *Journal of Accounting & Economics* 24 (1997): 99-126.

Collins, D; M. Pincus; and H. Xie. "Equity Valuation and Negative Earnings: The Role of Book Value of Equity." *The Accounting Review* 74 (January 1999): 29-61.

Coulton, J.; S. Taylor; and S. Taylor. "Is Benchmark Beating by Australian Firms Evidence of Earnings Management?" *Accounting and Finance*, Dec 2005, Vol 45, Issue 4, p 553-576.

Dechow, P.; S. Richardson; and I. Tuna. "Why are Earnings Kinky? An Examination of the Earnings Management Explanation." *Review of Accounting Studies* 8 (2003): 355-384.

Degeorge, F.; J. Patel; and R. Zeckhauser. "Earnings Management to Exceed Thresholds." *The Journal of Business* 72 (1999): 1-33.

Durtschi, C. and P. Easton. "Earnings Management? The shapes of the frequency distributions of earnings metrics are not evidence *ipso facto.*" *Journal of Accounting Research* 2005: 557-592.

Easton, P. "Security Returns and the Value Relevance of Accounting Data." *Accounting Horizons* 13 (December 1999): 399-412.

Holland, D., and A. Ramsay. "Do Australian Companies Manage Earnings to Meet Simple Earnings Benchmarks." *Accounting and Finance,* Mar 2003, Vol 43, Issue 1, p 21-40.

Ittner, C., D. Larcker; and M. Rajan. "The Choice of Performance Measures in Annual Bonus Contracts." *The Accounting Review* 72 (April 1997): 231-255.

Jacob. J., and B. Jorgensen. "Earnings Management and Accounting Income Aggregation." *Journal of Accounting and Economics* 43 (2007): 369-390.

Kerstein, J., and A. Rai. "Intra-year Shifts in the Earnings Distribution and their Implications for Earnings Management." *Journal of Accounting and Economics 44* (2007): 399-419.

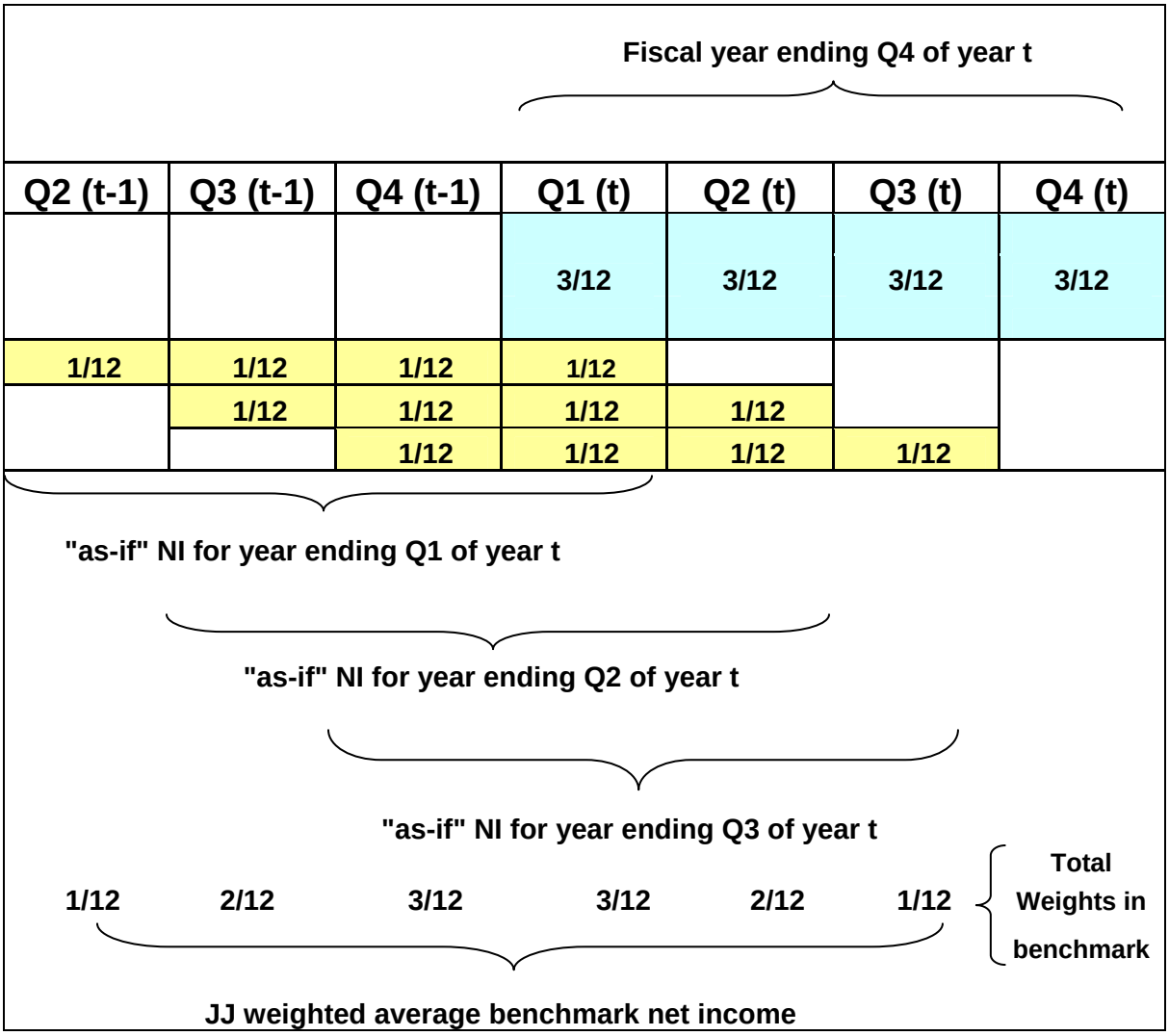
Leone, A., and R. Van Horn. "Earnings Management in Not-for-Profit Institutions: Evidence from Hospitals." Simon School of Business Working Paper FR-99-03, University of Rochester, 1999.

Phillips, J.; M. Pincus; S. Rego; and H. Wan. "Decomposing Changes in Deferred Tax Assets and Liabilities to Isolate Earnings Management Activities." *Journal of the American Taxation Association*, 2004 Supplement, Vol 26, p43-66.

Rego, S., and M. Frank. "Do Managers use the Valuation Allowance Account to Manage Earnings around Certain Earnings Targets?" Darden Business School Working Paper No. 03-09, January, 2006.

Revsine, L.; D. Collins; and W. Johnson. *Financial Reporting and Analysis*. 3rd Edition. Prentice Hall. NJ. 2005.

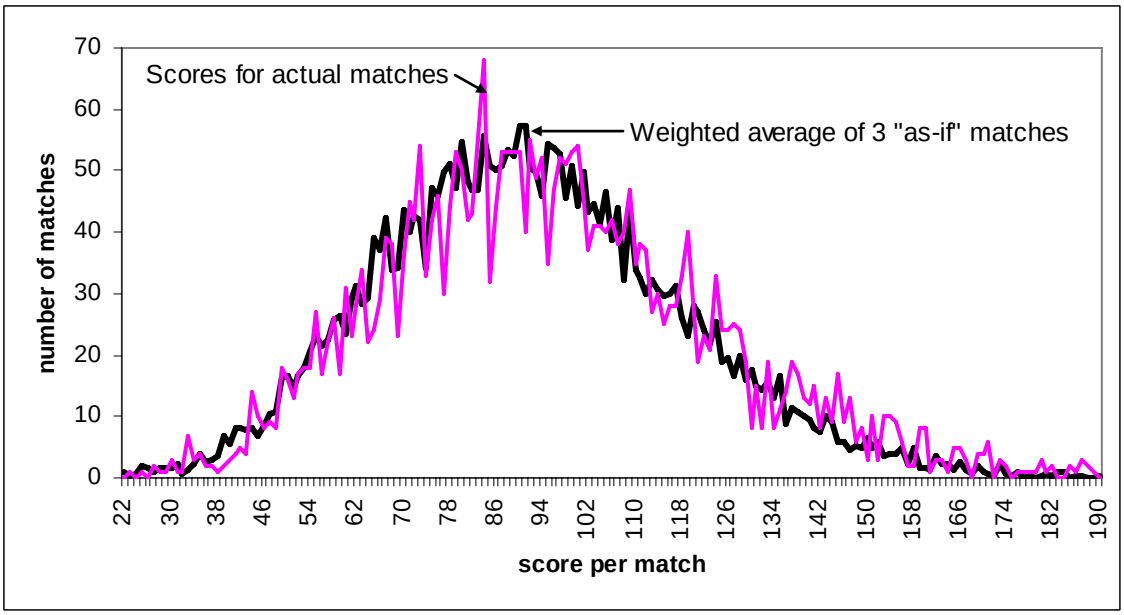
Figure 1: **Depiction of the JJ Method for Constructing the Weighted Average Net Income Benchmark**



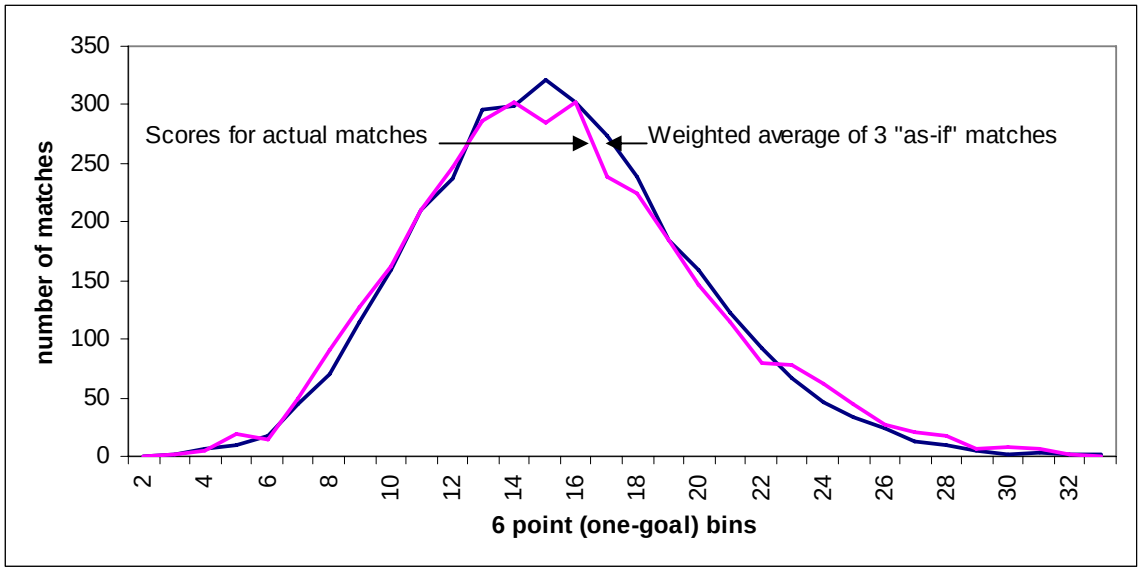
Fiscal year net income is the sum of quarterly net income for quarters one through four of a fiscal year (Quarterly Compustat data item 69), divided by beginning of year market capitalization: price (Quarterly Compustat data item 14) multiplied by number of shares outstanding (Quarterly Compustat data item 61) This fiscal year is compared to the average of net income calculated for three "as-if" years. These threes "as-if" years of net income are the sum of four consecutive quarters of Quarterly Compustat data item 69) ending in quarter 1, quarter 2, and quarter 3 of year t. The income of each "as-if" year is deflated by beginning-of-year market capitalization for that year. In other words, JJ assign a weight of 1/12 to earnings of quarter two of year t-1, 2/12 to earnings of quarter three of year t-1, 3/12 to earnings of quarter four of year t-1, 3/12 to earnings of quarter one of year t, 2/12 to quarter two of year t and 1/12 to quarter 3 of year t. The sample selection process requires seven quarters of continuous quarterly net income data.

Figure 2: The Effect of Averaging on Scores from Australian Football League Games

Panel A



Panel B



Panel A shows the frequency distribution of the scores for 22 matches between the 16 teams in the Australian Football League from 1997 to 2006: a total of for 3,420 matches are in the sample. The dark line is the average of three "as-if" matches where the average is calculated in the same way as in JJ. Each of the "as- if" matches consists of four consecutive quarters ending (1) in quarter one of the match in week t, (2) in the second quarter of the match in week t, and (3) in the third quarter of the match in week t. The light line is the frequency distribution of actual match scores. Panel B shows the same scores as in Panel A, but in 6 point bins. The dark line is the average of three "as-if" matches, the light line is the actual match scores in 6-point bins.

Figure 3: **Frequency Distribution of Net Income**

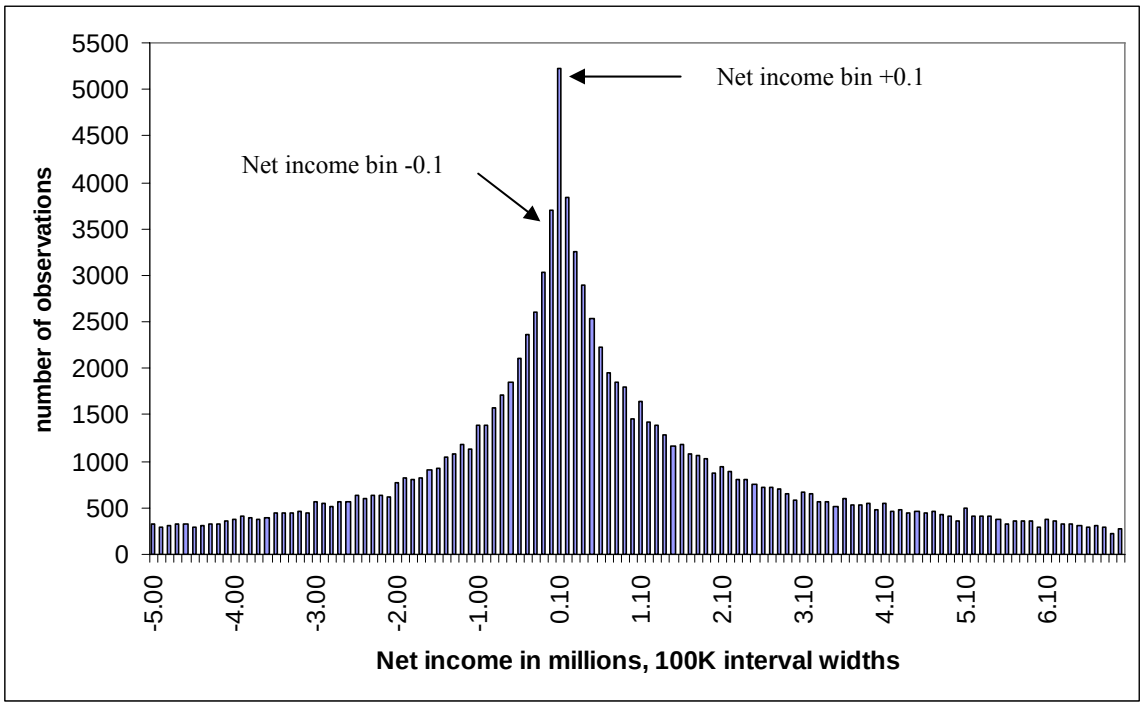
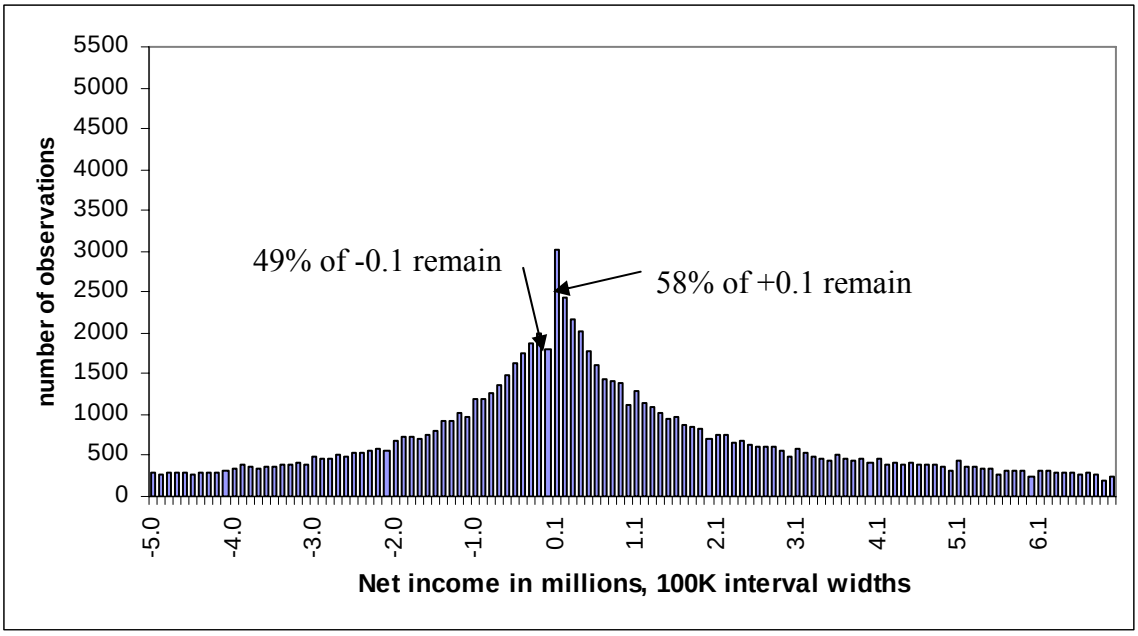


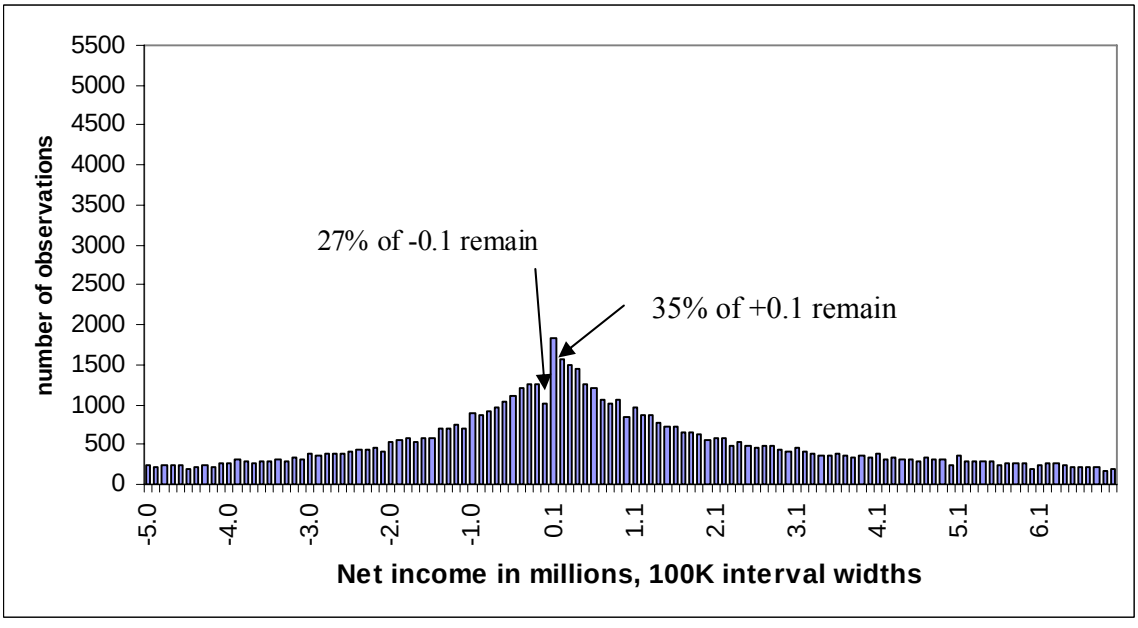
Figure 3 is the frequency distribution of annual net income (Annual Compustat data item 172) for the 108,299 observations between -$5,000,000 and $7,000,000. 29,998 observations are < -$5,000,000, and 29,558 observations are > $7,000,000.

Figure 4: **The Effect of Sample Selection on the Frequency Distribution of Annual Net Income (calculated as in JJ)**

Panel A.



Panel B.

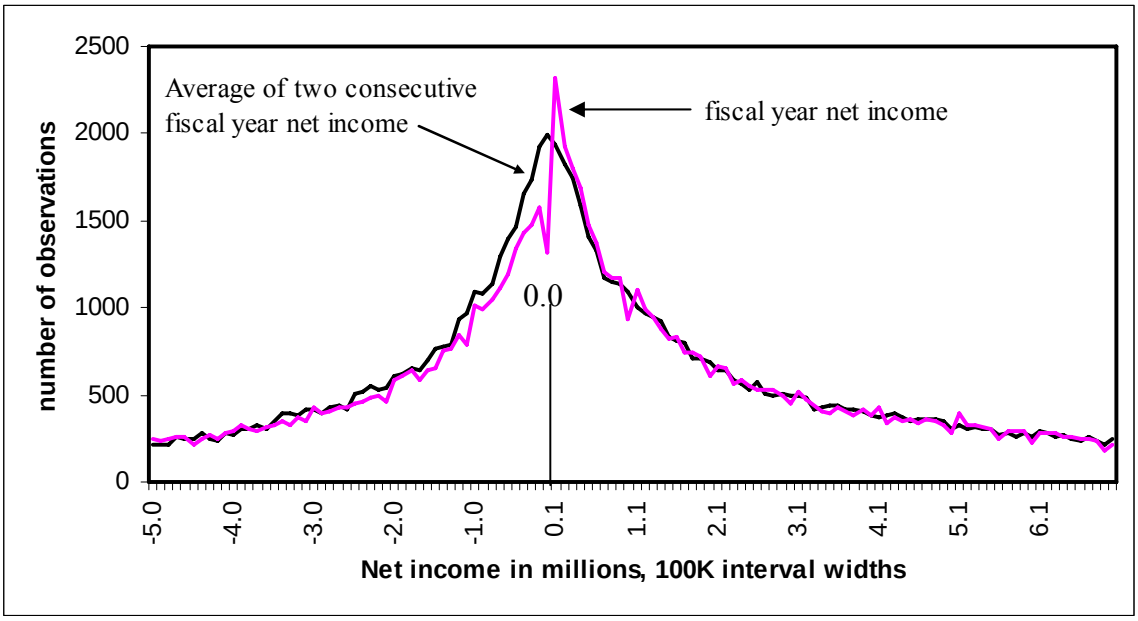


Panel A and Panel B both have the same scale as Figure 3. Panel A, is the frequency distribution of fiscal year net income computed as in JJ as the sum of four quarters of data (Quarterly Compustat data item 69) after removing the observations that do not meet the first JJ sample selection criteria of having four consecutive quarters of data for that fiscal year; 134,495 of the 161,700 observations that meet this sample selection criteria are shown. Data are put in $100,000 bins as in DE and Dechow et al. Panel B is fiscal year

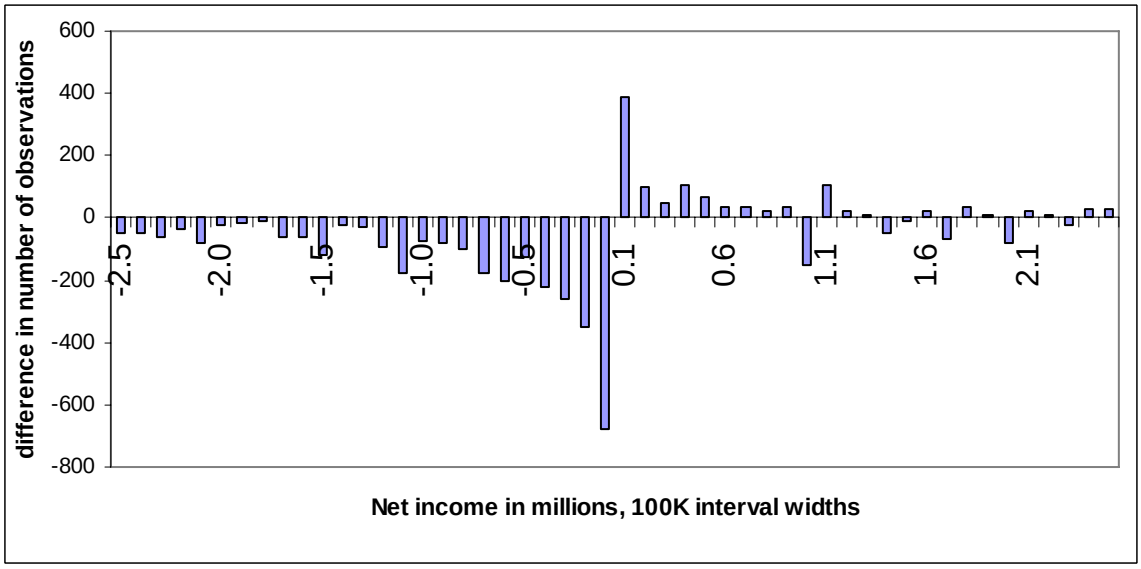
net income computed as in JJ as the sum of four quarters of data (Compustat data item 69) after removing those observations that do not meet both JJ sample selection criteria: four consecutive quarters of data and beginning-of-year market capitalization (Quarterly Compustat data item 14 times Quarterly Compustat data item 61).

Figure 5: The Effects of Averaging on the Distribution of Net Income

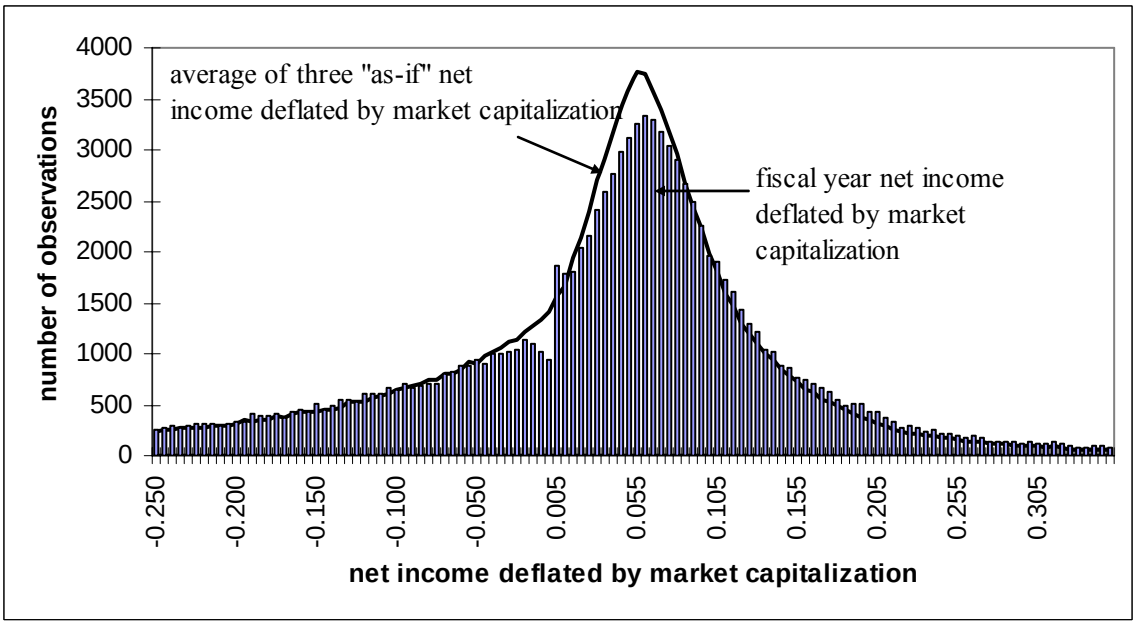
Panel A



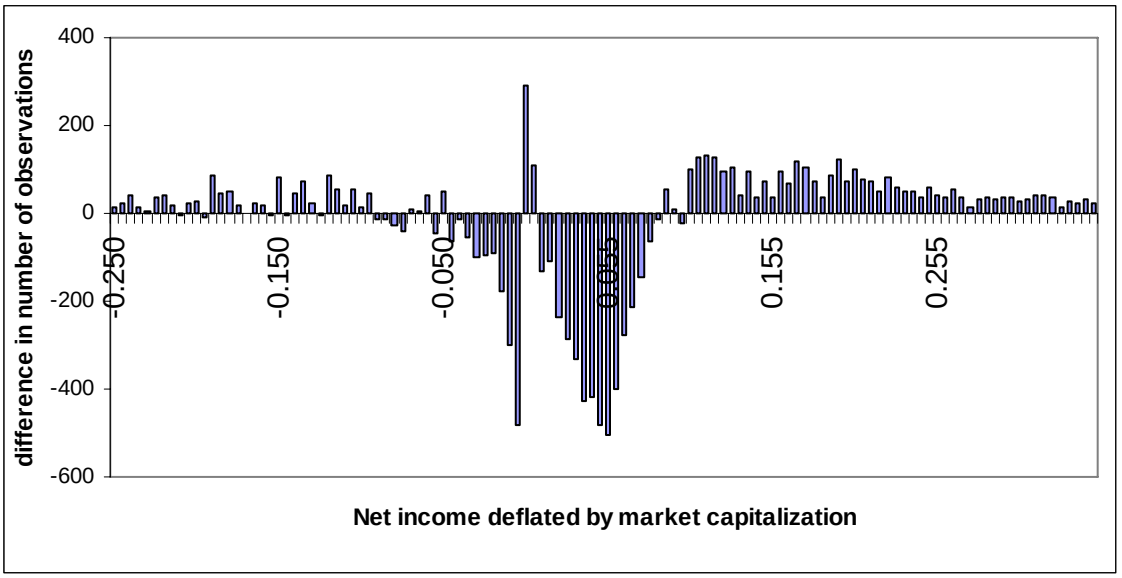
Panel B



Panel C



Panel D



Panel A compares the distribution of fiscal year net income (the dark line) with the average of fiscal year net income (the light line) across two consecutive years. Results are in $100,000 bins ranging from -$5,000,000 to $7,000,000. Net income is the sum of Quarterly Compustat data item 69 over the four quarters of the fiscal year. Each observation has eight consecutive quarters of data; 74,540 of the 144,027 observations with requisite data are shown. Panel B plots the difference between the frequency distributions for each $100,000 net income bin. Panel C is a replication of JJ Figure 1 with the exception that the denominator for all four years of net income is beginning-of-fiscal year market capitalization rather than a different denominator for each "as-if" year; The numerator of the variable plotted as the histogram is the sum of the four quarters of Quarterly Compustat data item 69 for the fiscal year. The numerator of the variable plotted as the line is the weighted average of three "as-if" years of net income ending in quarter one, quarter two, and quarter three. The denominator for both the fiscal year and all the weighted average is beginning-of-year market capitalization for the fiscal year (i.e., price – Quarterly Compustat data item 14 times number of shares outstanding – quarterly Compustat data item 61). Shown are 106,381 of the

126,436 observations available. Observations are put in bins of width 0.005, as in BD, DE and JJ. Panel D shows the difference between the number of observations plotted in the histogram and the number plotted via the line for each 0.005 interval.

Figure 6: Replication of JJ, Figure 1a

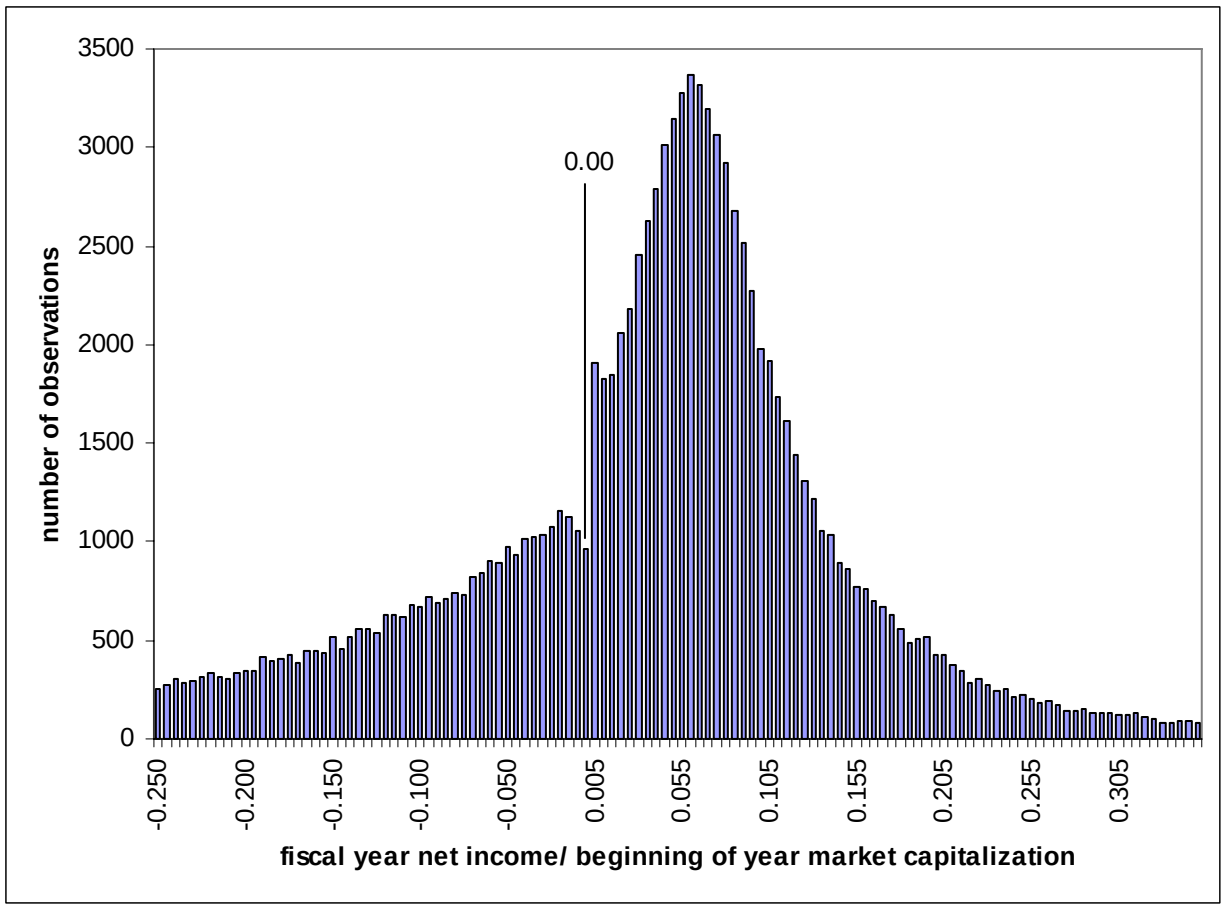
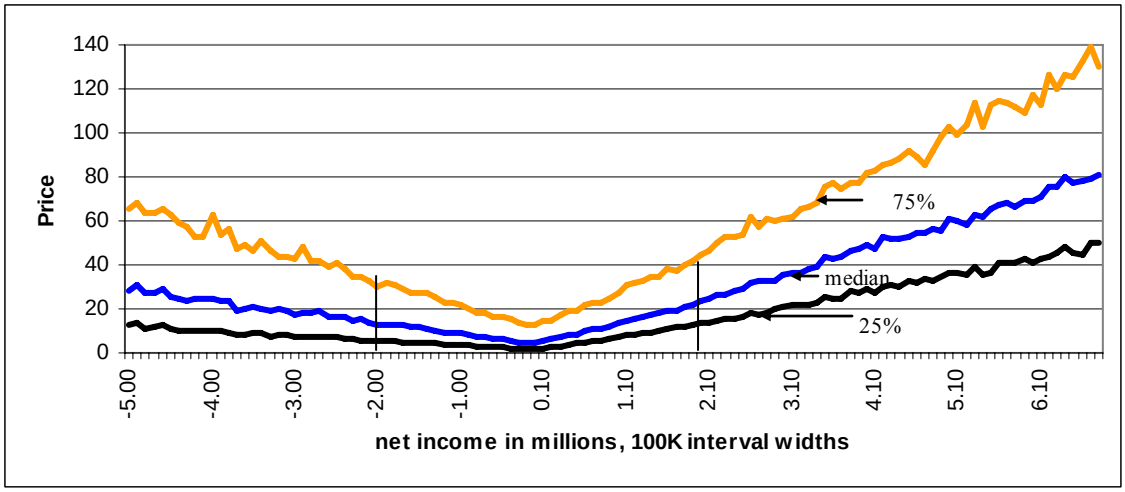


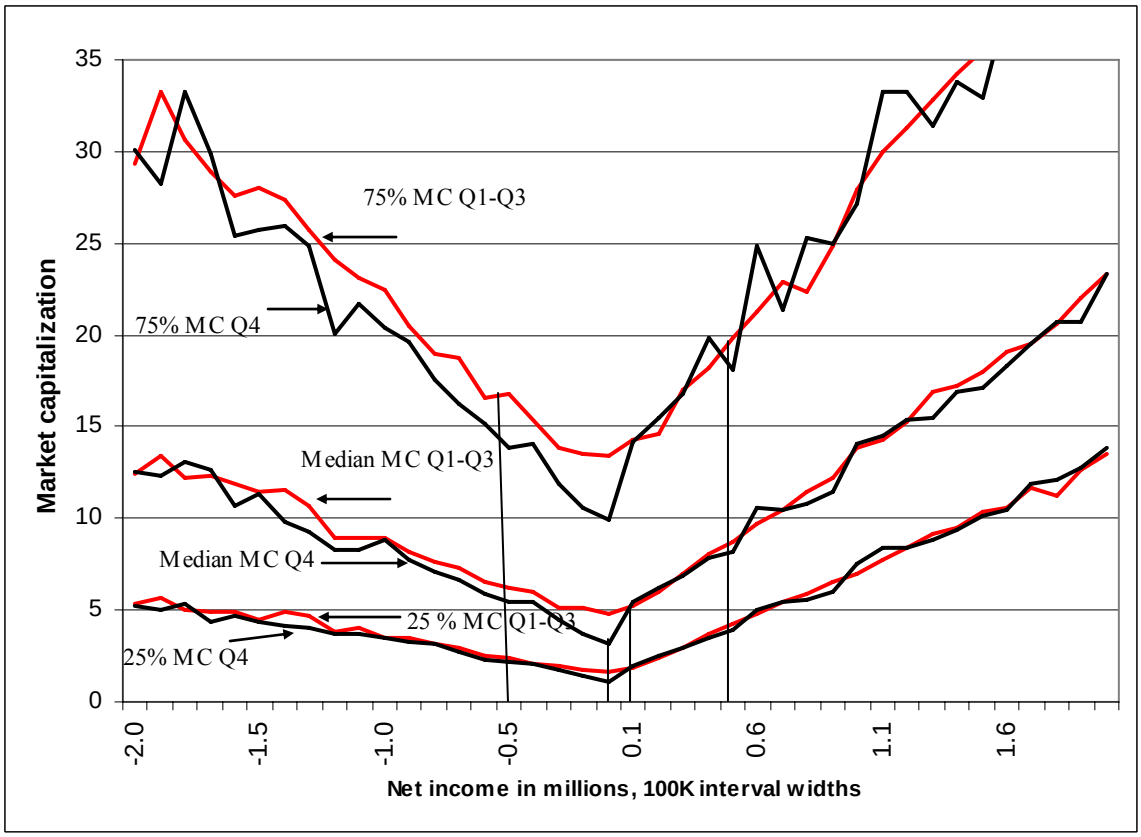
Figure 6 is the frequency distribution of deflated fiscal year net income for all available observations from 1977 to 2006. The numerator is fiscal year net income computed as the sum of the four quarters of net income for the fiscal year (Quarterly Compustat data item 69). The denominator is beginning-of-fiscal year market capitalization (i.e., price – Quarterly Compustat data item 14 times number of shares outstanding – Quarterly Compustat data item 61). Observations are put in bins for each 0.005 interval ranging from -0.250 to 0.35. Shown are 107,912 observations of the 128,477 available observations.

Figure 7: **The Distribution of the Deflator (beginning-of-period market capitalization) Used by JJ**

Panel A



Panel B

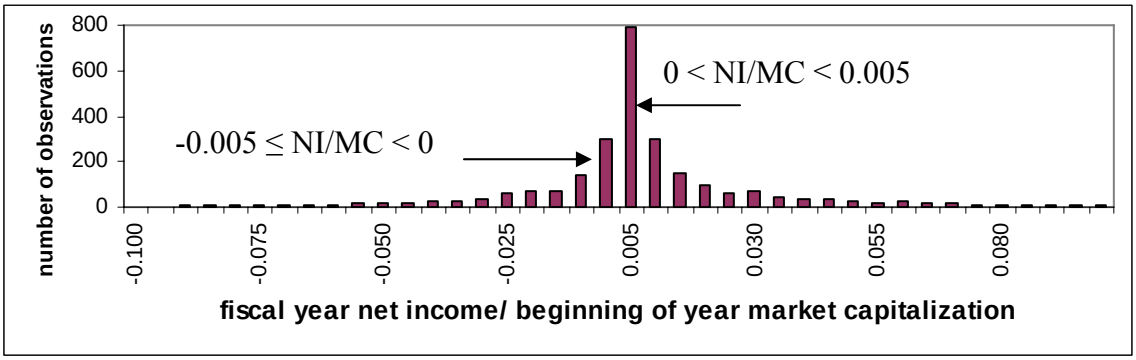


Panel A shows the distribution (25th percentile, median and 75th percentile) of market capitalization (i.e., price – Quarterly Compustat data item 14 times number of shares outstanding – Quarterly Compustat data item 61) for all net income observations between -$5,000,000 and $7,000,000 that meet the JJ sample selection criteria. Observations are placed in $100,000 net income bins. Panel B focuses on observations

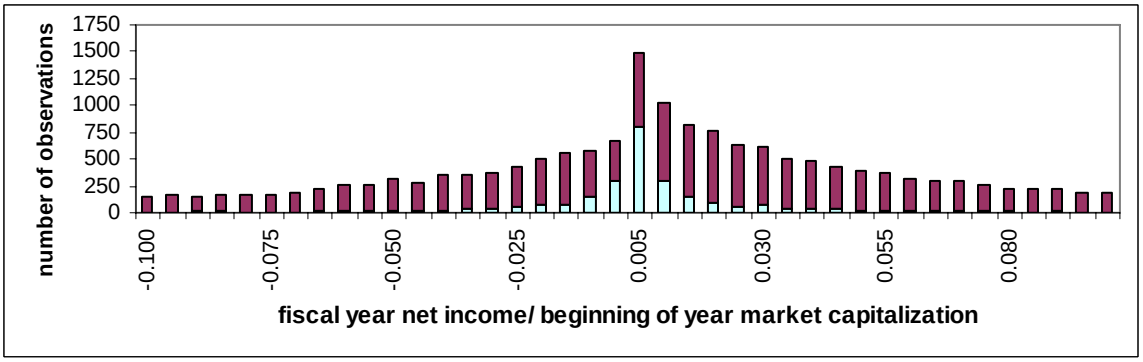
between -$2,000,000 and $2,000,000 and plots both separates beginning-of-fiscal-year market capitalization and the average of the beginning-of-year market capitalization for the three "as-if" years, which are used as the JJ benchmark. The light lines represent the average of beginning-of-year market capitalization for the three "as-if" years ending in quarters one, two, and three of year t. The dark lines represent the beginning-of-fiscal-year market capitalization.

Figure 8: **Effect of Deflation on the Distribution of (Deflated) Net Income**

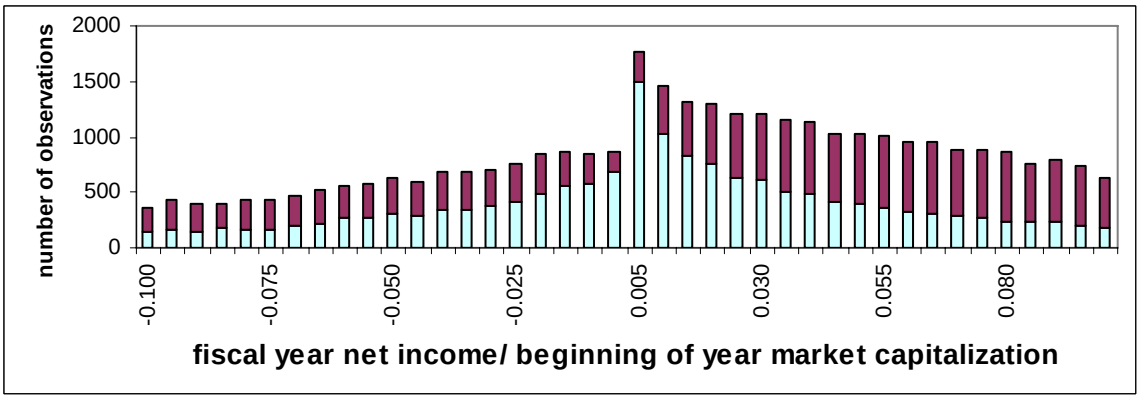
Panel A



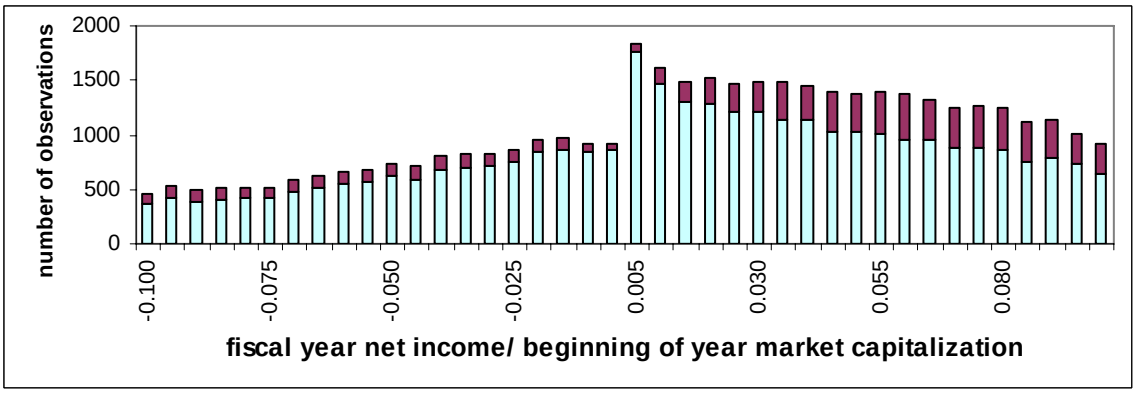
Panel B



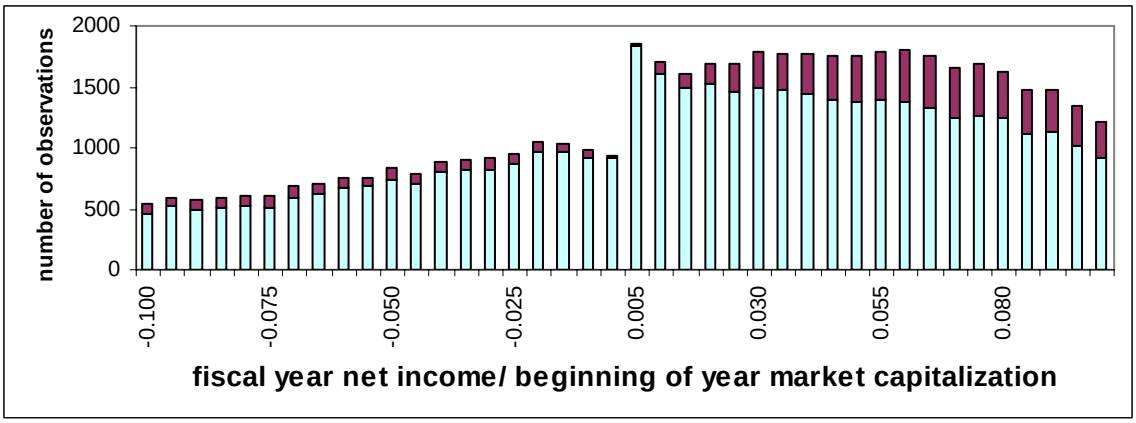
Panel C



Panel D



Panel E



Panel F

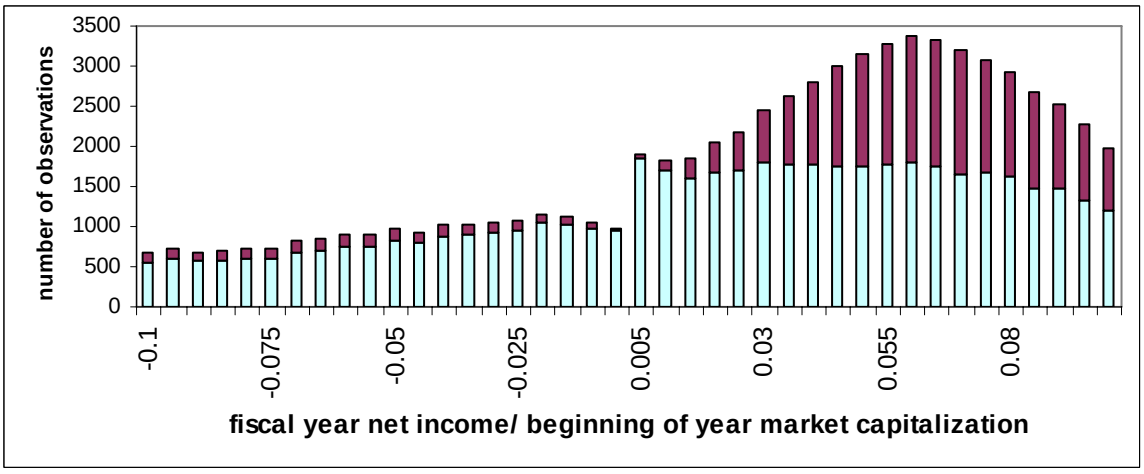


Figure 8 builds the BD and JJ frequency distribution of fiscal year deflated net income. These panels focus on the 40 BD/JJ bins around zero that include deflated net income ranging from -0.1 to 0.1. Each bin is 0.005 wide as in BD and JJ. Net income is the sum of four quarters of quarterly net income (Quarterly Compustat data item 69) for the fiscal year. The denominator is beginning-of-year market capitalization (i.e., price – Quarterly Compustat data item 14 times number of shares outstanding – Quarterly Compustat data item 61). Panel A shows the frequency distribution for all observations with net income between -$100,000 and $100,000, after these observations have been deflated by beginning-of-year market capitalization. Panel B, shows the observations from Panel A in light color, with the net income observations from -$1,000,000 to $1,000,000 (after deflation) added in the darker color. Panel C, includes observations from -$1,000,000 to $1,000,000 in the light shade and adds observations between -$5,000,000 and $5,000,000 (after deflation) in the darker shade. In Panel D, all observations included in Panel C are in the light shade and the additional observations from -$10,000,000 to $+10,000,000 are deflated and added in the darker shade. Panel E deflates observations between -$20M and $20M and adds them to the top of all observations between from -$10,000,000 to $10,000,000. Panel F deflates all remaining observations (those where net income is less than -$20 M and greater than $20M) and adds them in dark to all the prior observations.

Figure 9 – A Description of Compustat Quarterly Data Item 27

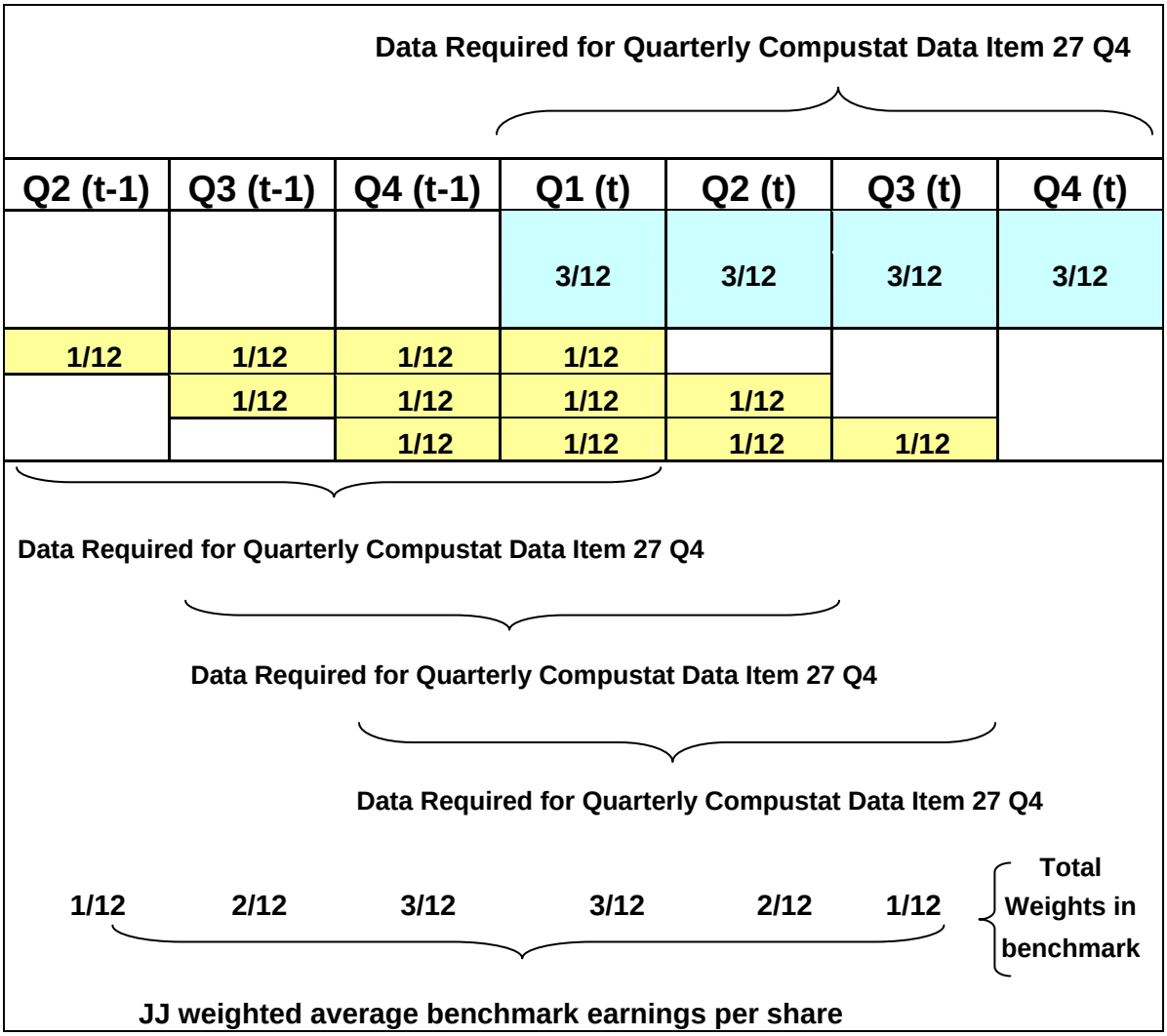
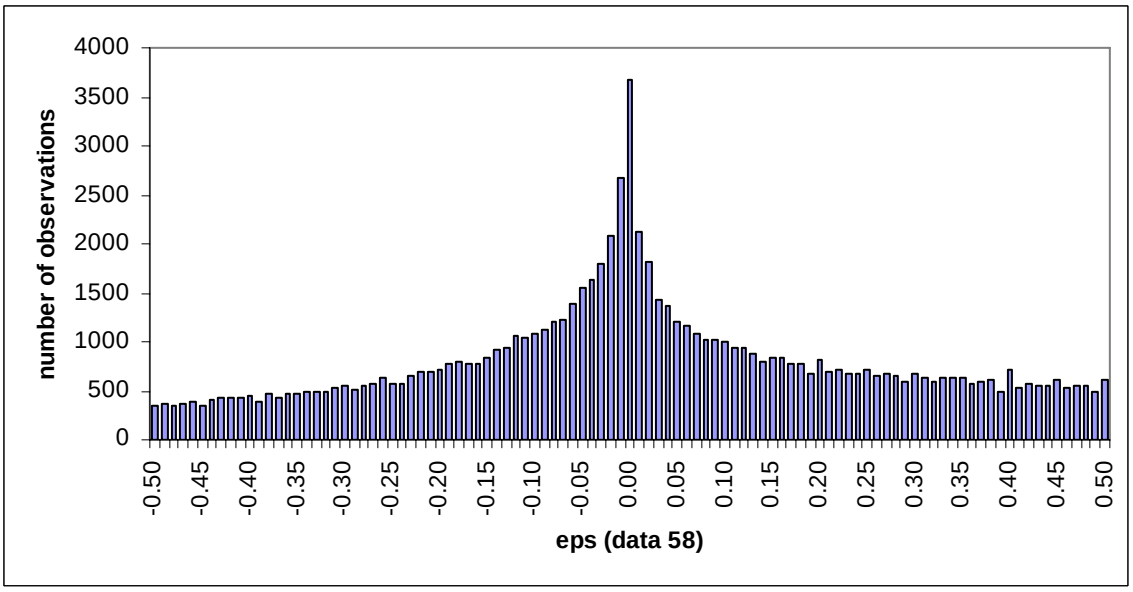


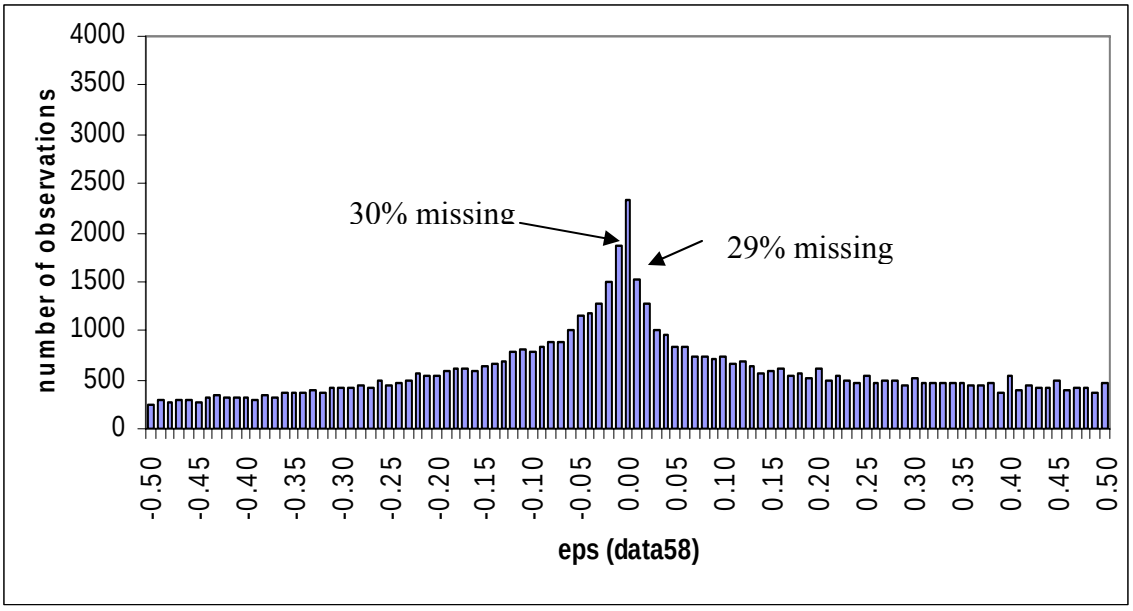
Figure 9 describes the Compustat computation of quarterly data item 27. Compustat adds "four quarters of available for common and divides by the 12-months moving shares figure." Data 27 for quarter four "will be the same as the annual earnings figure reported to shareholders, in interim quarters this figure will equal (within four cents deviation) the sum of four quarters of Earnings per Share (basic) – Excluding extraordinary items."

Figure 10: The Effect of Sample Selection on the Frequency Distribution of Earnings per Share

Panel A



Panel B



Panel A is the frequency distribution of annual earnings per share without extraordinary items (Compustat annual data item 58). Observations are shown between -$0.50 and $0.50, and 117,405 of the 185,401 observations available from the Annual Compustat file between 1977 and 2006 are displayed in the figure. Intervals have one-cent width. Panel B is the frequency distribution of earnings per share without extraordinary items (Compustat data item 58) for firm-years that meet the JJ sample selection criteria (i.e., four consecutive quarters of data item 27 are available on the Quarterly Compustat file).

Figure 11: **The Effect of Averaging on the Shape of the Distribution of Earnings per Share**

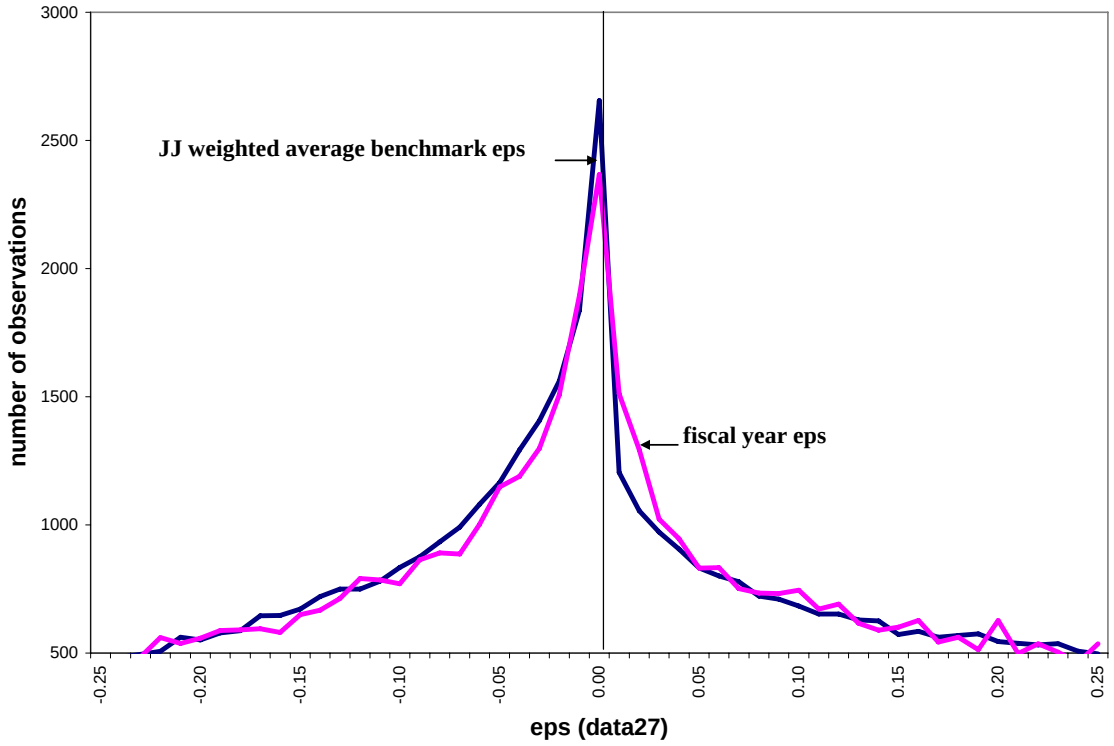
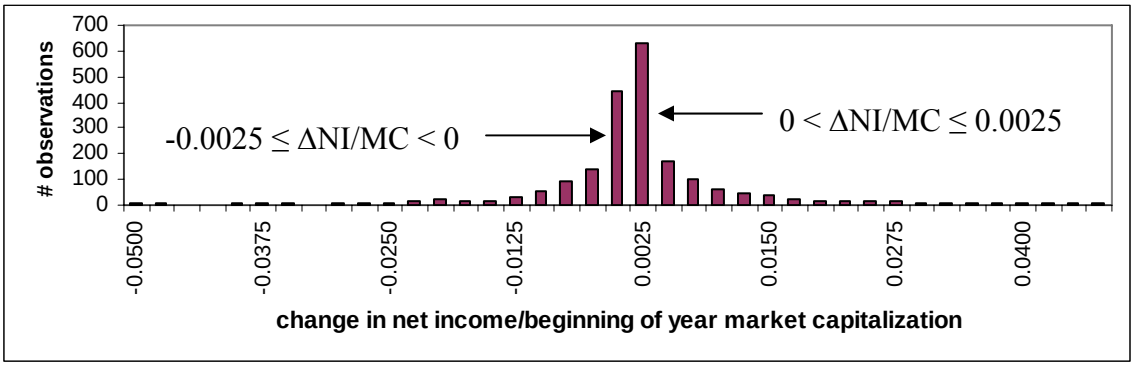


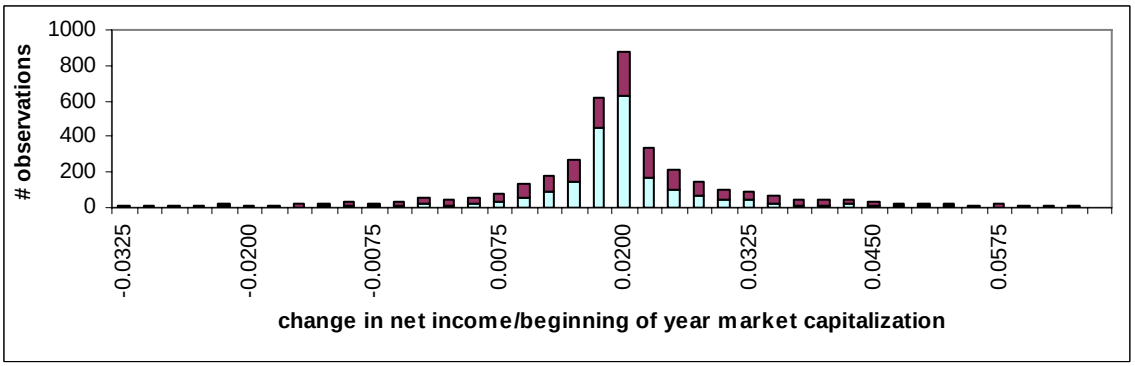
Figure 11 shows the frequency distribution of earnings per share without extraordinary items – 12 month moving (Quarterly Compustat data item 27) for observations that meet the JJ sample selection criteria (i.e., have four consecutive quarters of data item 27 in a fiscal year). The dark line shows the distribution of the average of three data 27 items: quarter one, quarter two, and quarter three of a fiscal year and represents 61,501 of the 144,642 observations that meet the sample selection criteria. The light line shows the frequency distribution of the fiscal year (i.e., quarter four) data item 27 and represents 60,970 of the 144,642 observations that meet the JJ sample selection criteria.

Figure 12: Effect of Deflation on the Distribution of Change in Net Income

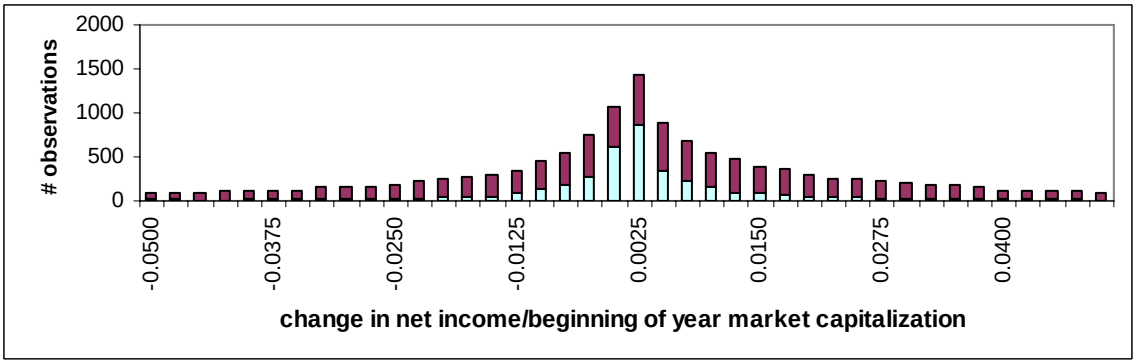
Panel A



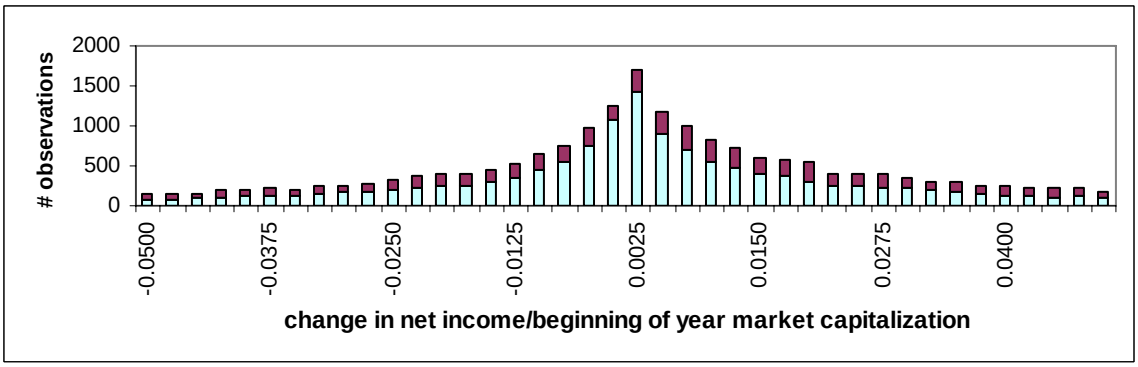
Panel B



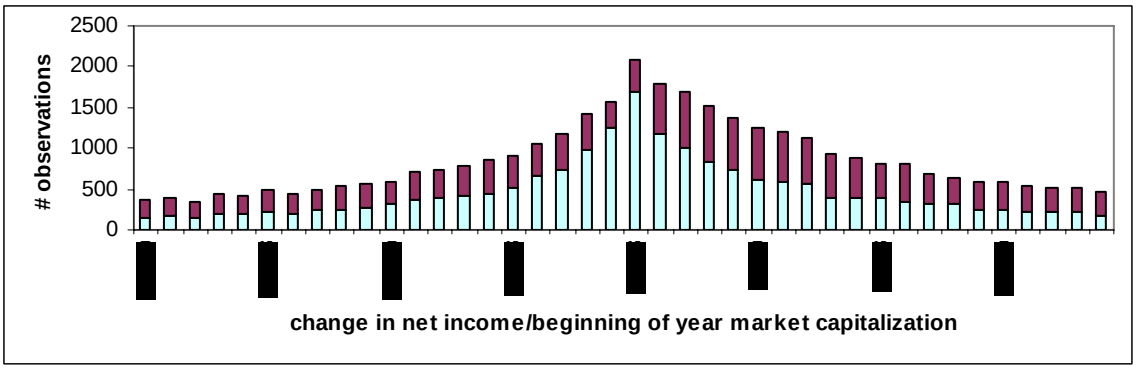
Panel C



Panel D



Panel E



Panel F

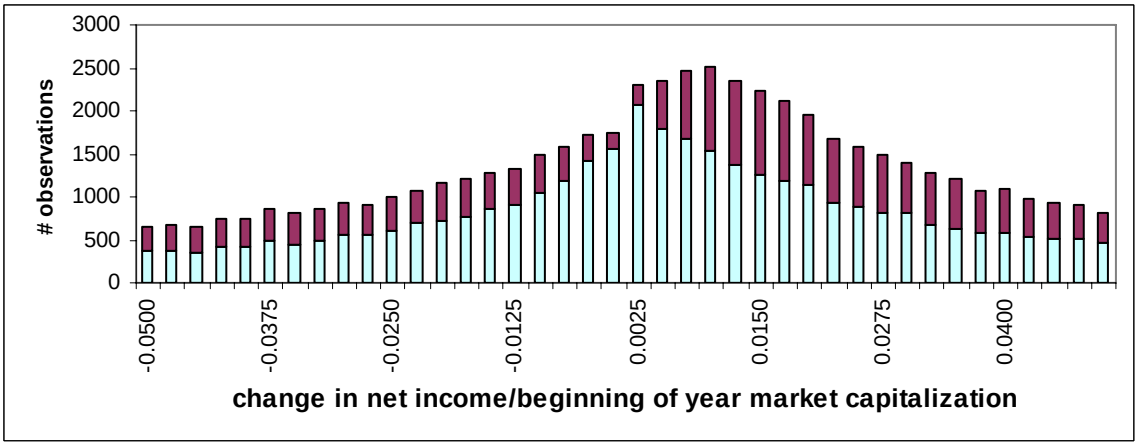


Figure 12 builds the BD and JJ frequency distribution of fiscal year deflated change in net income. These panels focus on the 40 BD/JJ bins around zero that include deflated change net income ranging from -0.0500 to 0.500. Each bin is 0.0025 wide as in BD and JJ. Net income is the sum of four quarters of net income (Quarterly Compustat data item 69). Panel A deflates changes in net income in the range -$50,000 to $50,000. Panel B deflates changes in net income in the range +/-$100,000 and adds the additional observations in the dark shade to the observations from Panel A (in the light shade). Panel C deflates changes in net income from +/- $500,000 and adds them (in the dark shade) to the distribution in Panel B (in the light shade). Panel D deflates changes in net income from +/- $1,000,000 and adds them in the dark shade to the observations shown in Panel C. Panel E deflates changes in net income from +/- $5,000,000 and adds them to the observations shown in Panel D. Panel E deflates all remaining changes in net income, those greater than $5,000,000 and less than $-5,000,000 and adds them (in dark shade) to all the observations shown in Panel D.

Figure 13: The Effect of Averaging on the Distribution of Changes in Earnings per Share

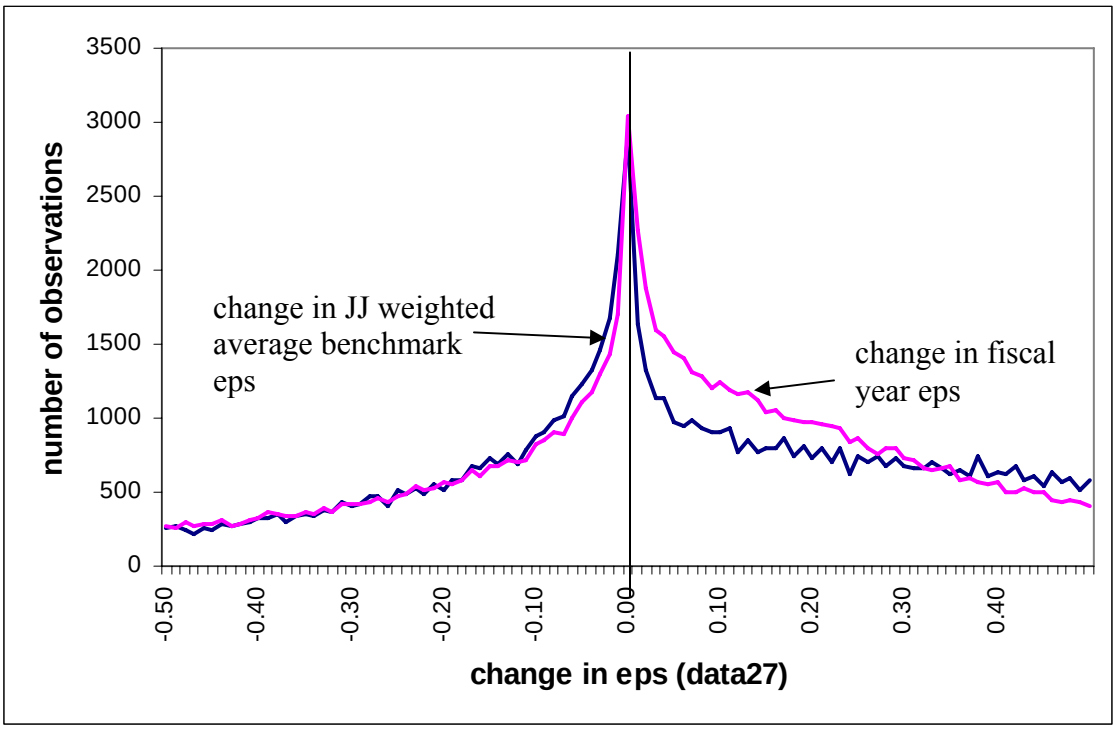


Figure 13 shows the effect of averaging on the distribution of extraordinary items – 12 month moving average (Compustat quarterly data item 27). The figure compares the weighted average of change in earnings across three "as-if" years ending quarter one of year t, quarter two of year t, and quarter three of year t.