Journal of Accounting and Economics 60 (2015) 168–186

Contents lists available at ScienceDirect

Journal of Accounting and Economics

journal homepage: www.elsevier.com/locate/jae

Do scaling and selection explain earnings discontinuities? ☆

David Burgstahler [a], Elizabeth Chuk [b,*]

[a] Julius A. Roller Professor of Accounting, University of Washington/Seattle, USA
[b] University of Southern California, USA

ARTICLE INFO

Available online 16 September 2014

JEL classification:
G14
M40

Keywords:
Earnings management
Discontinuities

ABSTRACT

Earnings distributions commonly exhibit statistically significant discontinuities at zero earnings, which are widely interpreted as evidence of earnings management to avoid a loss. In contrast, Durtschi and Easton (2005, 2009, hereafter DE) assert that discontinuities are instead explained by some combination of prior researchers' choice(s) of scaling and sample selection as well as a scaling-related effect due to a systematic relation between the sign of earnings and market prices. Resolution of the conflicting interpretations of discontinuities is important because (1) it affects how investors, regulators, and scholars view earnings management and (2) it demonstrates the importance of a close linkage between theory and research design choices. We point out that DE provide no evidence that scaling or selection create discontinuities, but only evidence showing that changes in scaling or selection eliminate discontinuities. We demonstrate why the research designs used by DE eliminate discontinuities and why alternative designs using the same data yield statistically significant discontinuities that cannot be attributed to either scaling or selection.

1. Introduction

There is extensive evidence of discontinuities in distributions of reported earnings at prominent earnings benchmarks, where distributions comprise fewer observations immediately below the benchmark and more observations immediately above the benchmark than are expected if the distribution is smooth.[1] For example, Burgstahler and Dichev (1997, hereafter BD) show that distributions of scaled earnings exhibit discontinuities at zero. In addition, BD document that the strength of discontinuities varies with the costs and benefits of meeting benchmarks. This evidence is widely interpreted as consistent with the theory that managers take both real and accounting actions to avoid losses. This interpretation is further supported

☆ We appreciate comments and suggestions from Bob Bowen, Ilia Dichev, Weili Ge, Dave Guenther, Frank Hodge, Bjorn Jorgensen, Bill Kinney, Allison Koester, Dave Maber, Rick Mergenthaler, Greg Miller, Shiva Rajgopal, Tatiana Sandino, Terry Shevlin, D. Shores, Bob Trezevant, Jerry Zimmerman (editor), an anonymous reviewer, and workshop participants at the 2010 UBC, Oregon, Washington Accounting Conference, Boston College, Ohio State University, Southern Methodist University, University of Colorado, University of Illinois, University of Michigan, University of Southern California, University of Texas, National University of Singapore, and Singapore Management University.
* Corresponding author.

[1] The smoothness assumption is discussed in more detail in Section 2 and in Burgstahler and Chuk (2014). Benchmarks for which discontinuities have been documented include the profit/loss benchmark (Hayn, 1995; Burgstahler and Dichev, 1997; Degeorge et al., 1999; Leuz et al., 2003; Daske et al., 2006); prior-year earnings (Burgstahler and Dichev, 1997; Degeorge et al., 1999; Beatty et al., 2002; Daske et al., 2006; Donelson et al., 2013); and analyst forecasts (Degeorge et al., 1999; Burgstahler and Eames, 2006; Daske et al., 2006; Donelson et al., 2013). This paper focuses on discontinuities at the profit/loss benchmark.

http://dx.doi.org/10.1016/j.jacceco.2014.08.002
0165-4101/

by survey evidence in Dichev et al. (2013) indicating that 99.4% of CFOs believe at least some firms manage earnings and also by survey evidence in Graham et al. (2005) indicating that managers are willing to incur real costs in order to meet benchmarks.

Durtschi and Easton (2005, hereafter DE1, 2009, hereafter DE2) assert that inferences attributing discontinuities to actions taken by managers are "erroneous." These papers provide examples where discontinuities are eliminated when the research design is changed and assert that discontinuities in earnings distributions are driven by "deflation, sample selection, and a difference between the characteristics of profit and loss observations." These assertions have led many to question whether discontinuities at zero earnings represent compelling empirical evidence of earnings management, or instead are due to the artifactual explanations advocated in DE1 and DE2 (hereafter referred to collectively as DE).

The purpose of this paper is to distinguish between two competing interpretations of the empirical results reported in DE. The first interpretation is that discontinuities do not exist in earnings distributions, but the scaling and sample selection choices, widely used in the literature, induce the discontinuities. The second interpretation is that discontinuities exist in earnings distributions but the alternative scaling and selection choices implemented by DE obscure the discontinuities.

Our analysis shows that the DE research design choices systematically reduce power, thereby obscuring or eliminating evidence of discontinuities due to earnings management because the DE research designs (1) do not account for the effect of firm size as a covariate, (2) place high weight on results for firms in the smallest size quartile and low weight on results for firms in the upper three quartiles, and (3) do not account for firm-size-related differences in the amount of earnings that can be managed at a cost lower than the benefits. To illustrate the importance of power in interpreting results, consider a research design choice with an obvious and well-known effect on power, namely sample size. Non-significant results from replications using a much smaller sample size than was used in the original test should not be interpreted as contradicting the significant results from the original test nor as confirmation of the null hypothesis. The same principle applies in interpreting results from the DE research designs. Because the DE designs reduce power, the lack of significant evidence of discontinuities based on the DE research designs should not be interpreted as contradicting previous significant results nor as confirming the null.

Another way to evaluate the DE explanations is to examine whether the presence or absence of discontinuities is explained by the presence or absence of scaling or selection effects. Gilliam et al. (2013, hereafter GHP) illustrate this approach. GHP conjecture that the implementation of the Sarbanes-Oxley Act (SOX) and other events around SOX have decreased the incidence of earnings management and document that the discontinuity at zero earnings disappears after SOX. As GHP point out, if discontinuities are explained by the data artifacts proposed by DE, then the data artifacts must have changed such that they once created the zero-earnings discontinuity before 2002 but no longer created the discontinuity after 2002. GHP do not find such a change in data artifacts after 2002, thereby casting doubt on the DE artifactual explanations.[2]

The approach in our paper and the approach in GHP are complementary, and each approach has its advantages. One of the main advantages to GHP's approach is that it identifies effects of a regulatory change that is explicitly linked to the incentives for earnings management. GHP show that evidence of discontinuities disappears when regulatory changes reduce the incentives for earnings management, even though the same scaling and selection artifacts continue. The main advantage of our approach is that we directly test the DE assertions and explicitly show that there is no evidence in DE that scaling and selection cause discontinuities. Our analyses and the analyses in GHP represent separate yet parallel approaches that both support the conclusion that the DE scaling and selection effects cannot explain discontinuities.

The paper is organized as follows. Section 2 provides background on theory and evidence of earnings management to meet benchmarks. Section 3 provides empirical evidence showing why the DE research designs predictably lead to non-significant discontinuity evidence. Section 4 concludes.

2. Background

Earnings management often is inferred based on evidence that the distribution of reported earnings differs from what would have been expected in the absence of earnings management. Specifically, BD suggest that earnings management to meet benchmarks creates discontinuities in the distribution of reported earnings. These discontinuities provide evidence that *some* firms have taken actions consistent with earnings management but do not specifically identify *which* firms have taken these actions. Further, because this approach focuses on reported total earnings, the set of actions included in the definition of earnings management is broad, including all accounting (accrual and disclosure) actions and all "real" operating, investing, and financing actions that affect reported earnings.

Economic theory suggests that managers take actions when they believe the benefits exceed the costs. When pre-managed earnings are below the benchmark by a small enough amount that the cost of managing earnings upward to meet the benchmark is less than the benefits, earnings will be managed upward to meet the benchmark. The cost of managing earnings includes the costs of accounting actions or the opportunity cost of suboptimal real operating, investing, or

[2] Similarly, Jacob and Jorgensen (2007), Donelson et al. (2013), and Jorgensen et al. (2013) show that discontinuities exist in earnings measures where there are incentives to manage earnings but do not exist in closely-related earnings measures that are subject to exactly the same scaling and selection artifacts but where there is little or no incentive to manage earnings, thereby also casting doubt on the DE explanations.

financing transactions used to achieve higher reported earnings. The benefits of managing earnings are primarily improved terms of transactions with stakeholders.[3] Earnings management occurs when perceived costs are less than the perceived benefits. When costs increase with the amount of earnings management while the benefits of meeting the benchmark are fixed, there is an upper bound on the amount of earnings that can be managed at a cost less than the benefits. For pre-managed earnings below the benchmark by no more than this bound, it is optimal to manage earnings to meet the benchmark because the costs are less than the benefits. For simplicity, we refer to this upper bound on the amount by which a firm can manage earnings to meet the benchmark at a cost less than the benefits as the "maximum earnings management."

While it is not possible to specify the precise relation between firm size and the maximum earnings management for the firm, it is reasonable to expect that the amount increases with firm size. The benefits from meeting a benchmark are likely larger for larger firms that have more and larger transactions with stakeholders. While it is more costly to manage earnings by a larger dollar amount, the materiality threshold used by auditors generally increases with firm size, so larger firms are able to make accounting choices that affect earnings by larger amounts without exceeding the materiality threshold.[4] Further, larger firms will have available a larger menu of operating, investing, and financing decisions that affect reported earnings, and this larger set of choices affords larger firms the ability to use real choices to manage earnings by larger amounts. Thus, it is plausible that the maximum earnings management for a $10 billion firm will be proportionally larger than the maximum earnings management for a $10 million firm.

Factors other than size can also affect the benefits and costs of earnings management. As benefits decrease or costs increase due to these other factors, the maximum amount by which earnings can be managed at a cost less than the benefits moves toward zero. Consequently, earnings management occurs for fewer observations, and discontinuities disappear. Thus, we expect less evidence of discontinuities when firms realize lower benefits or face higher costs of earnings management. For example, firms in financial difficulty likely perceive lower benefits of meeting an earnings target (because concerns about financial difficulty reduce any potential improvement in terms of transactions due to meeting the earnings benchmark) and face higher costs (because financial difficulty is associated with decreased flexibility to take actions that increase reported earnings). Therefore, the maximum earnings management is likely to be lower for firms in financial difficulty, resulting in lower rates of earnings management.[5]

3. DE "alternative explanations" for discontinuities

DE1 and DE2 assert that "deflation, sample selection, and a difference between the characteristics of profit and loss observations" are "much more likely" explanations for discontinuities than the widely-accepted theory that discontinuities are the result of earnings management to meet benchmarks. This section provides analysis and empirical evaluation for each of the three DE explanations.

For our analysis, we include all available observations in the annual Compustat file FUNDA for years 1990–2009. Although our sample period does not correspond exactly with either of the sample periods used in DE1 or DE2, there is substantial overlap. Moreover, because we use a sample period of comparable length to that in DE1 and somewhat shorter than the 30-year period in DE2, our samples sizes are generally comparable to, or smaller than, those in DE. Therefore, the difference between the significant results from our research designs versus the insignificant results from the DE research designs cannot be attributed to the effect of sample size on the power of tests. Following BD and DE, we delete banks, financial institutions, and firms in regulated industries.

3.1. Scaling as an alternative explanation

In this section, we evaluate the first alternative explanation proposed by DE, scaling, which they summarize as follows:

> "The arguments that net income must be deflated 'in an attempt to homogenize firms' [Degeorge et al., 1999, p. 16] or 'because firms are drawn from a broad range of firm sizes' [Burgstahler and Dichev, 1997, p. 102] seem reasonable at first glance. But these studies do not provide a priori reasons why heterogeneity or differences in firm size might affect the empirical analyses. More importantly, the implicit assumption is that deflation will not distort the underlying distribution of net income. Our results suggest that it does. Furthermore, since the hypotheses in the extant literature focus on management of earnings to avoid small losses, and since the arguments underlying these hypotheses do not suggest that earnings are managed relative to beginning-of-period price,

[3] As discussed in BD Section 4, examples of improvements in terms of transactions when benchmarks are met include lower wage and benefit demands from current and potential employees, better terms from suppliers and creditors, and higher valuation by shareholders.

[4] Auditors often define materiality thresholds as a percentage of expected earnings, sales, assets, or another measure of firm size (AICPA Audit and Accounting Manual, 1996; Nelson et al., 2005).

[5] BD provide evidence that there is less evidence of discontinuities in earnings levels and changes among firms with a weaker record of recent profitability, consistent with the hypothesis that these firms realize lower benefits or face higher costs of managing to meet benchmarks. Beatty et al. (2002) show that the rate of earnings management is lower among private banks than among public banks, consistent with the hypothesis that private banks have lower incentives to manage to meet benchmarks.

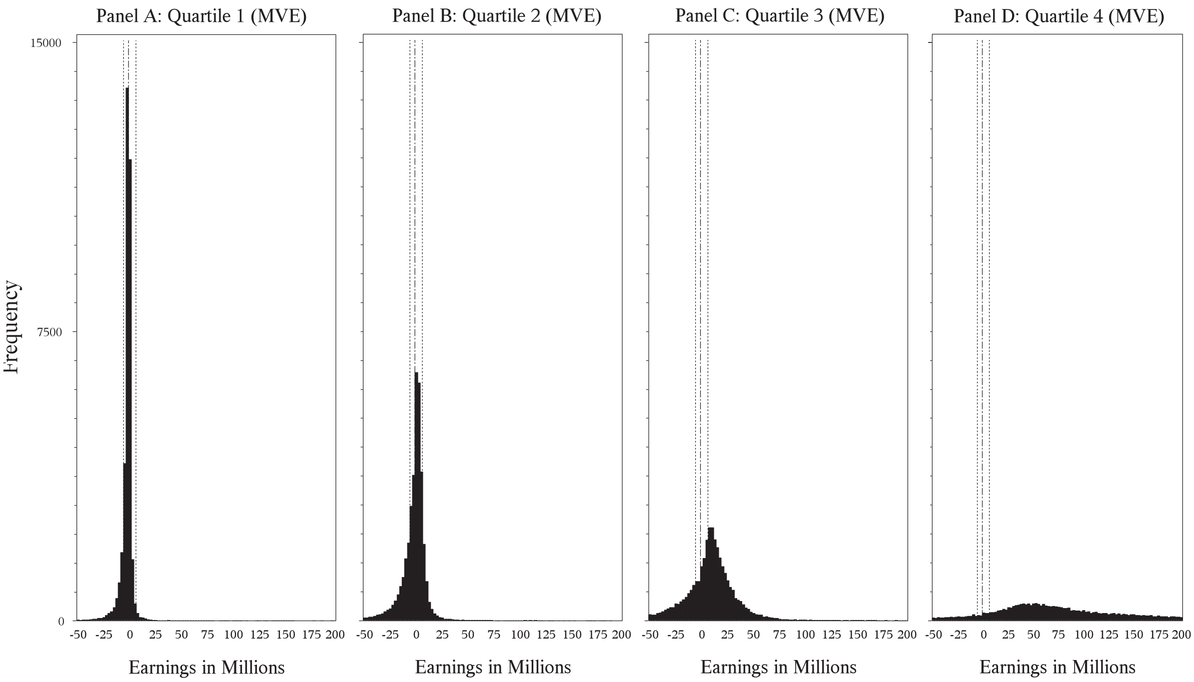
we question whether it is ever appropriate to examine deflated earnings in an earnings management context." (DE1 p. 559)

We begin by considering the research design rationale for scaling. Contrary to the arguments in DE, the explicit assumption of scaling is that it is necessary to "distort the underlying distribution" to achieve two goals: (1) to account for firm size as a covariate that explains variation in earnings (irrespective of earnings management) and (2) to account for the relation between firm size and the amount of earnings that can be managed at a cost lower than the benefits.

First, effective research design must account for the strong relation between firm size and earnings. Theory and empirical evidence clearly indicate that size explains a great deal of variation in earnings. For example, the mean earnings reported by $1 billion firms is far larger than the mean earnings reported by $10 million firms. Similarly, the standard deviation is strongly related to size – the probability of an earnings observation $10 million or more away from (either above or below) the mean earnings is far larger for a $1 billion firm than for a $10 million firm. We show that by failing to account for the effect of firm size, the DE design places high weight on observations for the smallest firms and, for reasons explained in the next few paragraphs, also results in low power tests for the smallest firms.

Second, effective research design must also account for size-related differences in the amount of earnings that can be managed at a cost less than the benefits of meeting the benchmark. When the maximum earnings management varies with firm size, effective research design must allow the interval width in earnings histograms to vary with firm size. The interval width determines the prominence of the trough below the benchmark and the peak above the benchmark. An interval width much wider than the maximum amount that typically can be managed to meet benchmarks will (1) include in the interval immediately below the benchmark many observations that are too far below the benchmark to be managed, decreasing the prominence of the trough, and (2) include in the interval above the benchmark many observations that have not been managed, decreasing the prominence of the peak. On the other hand, an interval width much narrower than the maximum amount that typically can be managed to meet benchmarks will mean that (1) a large proportion of observations two or three (or more) intervals below the benchmark will be managed, so that the trough will not be concentrated in just the interval below the benchmark, and (2) observations managed above the benchmark are more likely to be managed two or three (or more) intervals above the benchmark, decreasing the prominence of the peak in the interval immediately above the benchmark. Therefore, an effective research design cannot use the same interval width for $1 billion firms as for $10 million firms.

Because the maximum earnings management can vary across observations for reasons other than firm size, there is no interval width that will be appropriate for every observation in the distribution. Instead, it is necessary to choose an interval width that fits reasonably well with the typical amount of earnings that can be managed at a cost less than the benefits.



This figure shows histograms of unscaled earnings (Compustat item NI), partitioned into quartiles based on beginning market value of equity (the product of lagged Compustat items PRCC_F and CSHO). The 25th, 50th, and 75th percentiles of MVE that define the four quartiles are $22.2m, $106.3m, and $586.8m, respectively. All panels use the range -$50m to $200m of unscaled earnings with interval widths of $2.5m. The center dashed line indicates the zero benchmark, and the outer dotted lines indicate the range -$5m to $7m used in DE1 Figure 3 and DE2 Figures 1, 6, and 8

Fig. 1. Overview of distribution of unscaled earnings by quartiles formed on beginning market value of equity (MVE).

As shown in Sections 3.1.1 and 3.1.2, if the interval width does not correspond at least approximately to the typical maximum amount of earnings that can be managed by firms in the distribution, any discontinuity due to earnings management will be obscured or eliminated.

The following sections show that significant discontinuities exist in distributions of unscaled earnings and earnings per share, contrary to the DE assertion that discontinuities are due to scaling. For both unscaled earnings (Section 3.1.1) and earnings per share (Section 3.1.2), the analysis proceeds in three steps: First, we partition the population into quartiles based on firm size (measured as MVE in Section 3.1.1 or PPS in Section 3.1.2 to correspond to the per share measures in DE) to illustrate the strong relation between earnings and firm size. Second, we show the design using the DE interval width yields no evidence of discontinuities for either firms from the smallest size quartile or firms from the two largest size quartiles. We also explain how the DE design effectively places almost all the observational weight on the smallest size quartile, where the DE design yields no evidence of discontinuities. Third, we show how a research design using interval widths that are adjusted in proportion to firm size leads to statistically significant discontinuities in the three upper size quartiles and inconclusive evidence for the smallest size quartile.

3.1.1. *Unscaled earnings*

Fig. 1 provides a descriptive overview of the effect of firm size on earnings distributions. We expect a positive relation between size and both the mean and dispersion of reported earnings distributions because larger firms tend to have larger expected future earnings and there is a positive relation between future earnings and reported earnings. We use MVE as the primary measure of size in the analysis because (1) valuation models provide a direct theoretical link between MVE and expected future earnings, and (2) we expect MVE to be more closely related to differences in the maximum amount of earnings that typically can be managed at a cost less than the benefits than alternative measures of firm size, such as total revenues, total assets, and book value of equity.[6] Further, although benefits of earnings management to meet benchmarks are likely to be positively related to any of the alternative size measures, costs of earnings management are likely to be most closely related to market value of equity, under the additional assumption that the flexibility to manage current earnings is proportional to the present value of expected future earnings.

Fig. 1 shows distributions of unscaled earnings for a range that includes earnings of most of the firms on Compustat, –\$50 to \$200 million, and a histogram interval width of \$2.5 million.[7] The four panels correspond to the four quartiles of firm size, where size is measured as the beginning market value of equity (MVE).[8] Firms from the smallest size quartile have median MVE of \$7.5 million while firms from the upper three quartiles have median MVEs of \$50.1 million, \$235.2 million, and \$1,983.1 million, respectively. Thus, the range of size across the four quartiles is substantial—in round numbers, the median MVE increases by a multiple of about 5 across each quartile.

Fig. 1 and the corresponding descriptive statistics in Table 1 confirm the expected strong relation between MVE and both the location and dispersion of unscaled earnings. For the lowest MVE quartile, the distribution of unscaled earnings has a negative median and mean and is highly concentrated in the vicinity of the zero benchmark. For progressively larger MVE quartiles, the medians and means become increasingly positive and the dispersion of the earnings distributions increases. The dotted vertical lines in each panel indicate the range of unscaled earnings examined in the DE research design, unscaled earnings observations between –\$5 million and \$7 million. The preponderance of observations included in the DE design are from the first quartile while relatively few of the observations from the larger MVE quartiles are included, a point we explore in more detail in Fig. 2.

Fig. 2 Panels A–D show distributions of unscaled earnings using an interval width of \$100,000, the width used in the DE research design. A design that uses a constant interval width of \$100,000 for firms regardless of size is appropriate only when the maximum amount of earnings that can be managed at a cost lower than the benefits is \$100,000 irrespective of firm size. When this amount is substantially different from \$100,000 for some firms, a design using the \$100,000 interval width will obscure or eliminate discontinuities. Among firms where the maximum earnings management is much smaller than \$100,000, many of the observations in the \$100,000 interval immediately below the benchmark will be too far below the benchmark to be optimally managed and the rate of earnings management will be small, obscuring the discontinuity. Among those firms where the maximum earnings management is much larger than \$100,000, many of the observations in the intervals more than \$100,000 below the benchmark will be managed, again obscuring or eliminating the discontinuity.

The results in Panels A–D are consistent with the predicted effects of using a single interval width for all size quartiles. Results in Panel B show statistically significant evidence of a discontinuity, consistent with what is expected if \$100,000 corresponds approximately to the maximum amount of earnings that typically can be managed at a cost less than the benefits for observations in Q2. The left standardized difference is significantly negative and the right standardized

[6] Most of the results reported below also hold for alternative size measures, consistent with results previously reported in the literature showing that discontinuities are observed in distributions of earnings scaled by any of these alternative measures of size. See for example, BD p. 102.

[7] BD found that a large proportion of the small number of exact zero values of unscaled earnings (Compustat item NI) were errors. To avoid incorrectly treating data errors as values that meet the benchmark of zero, all observations with exact zero values of unscaled earnings are deleted throughout the analysis (though values exactly equal to zero are relatively rare so this choice has no qualitative effect on our conclusions). Note that this issue arises for unscaled earnings but not for EPS, where exact zero values are relatively common and are not subject to a high rate of errors.

[8] Because calculation of MVE requires market price, the results in Figs. 1 and 2 are subject to a selection effect that restricts the analysis to firms that have market prices. Selection effects are addressed in Section 3.3.

Table 1
Descriptive statistics.

	N	1st Pctl	25th Pctl	50th Pctl	75th Pctl	99th Pctl	Mean	Std dev	Skewness
Figs. 1 and 2: **Unscaled earnings (in millions of dollars), by MVE Quartile**									
Panel A or AA	37,687	−40.226	−2.393	−0.398	0.579	14.300	−1.813	6.402	−3.196
Panel B or BB	37,686	−94.648	−5.503	0.965	4.548	27.669	−2.867	16.017	−3.045
Panel C or CC	37,687	−262.915	−6.947	9.100	20.865	98.249	0.703	46.728	−2.880
Panel D or DD	37,686	−1,533.650	32.180	95.352	297.764	6,428.000	359.243	996.787	3.873
Panel E	150,746	−429.257	−2.849	2.027	24.234	2,460.190	74.085	330.489	5.366
Figs. 3 and 4: **EPS (in dollars), by PPS quartile**									
Panel A or AA	37,909	−6.180	−0.290	−0.060	0.010	1.130	−0.305	0.915	−4.140
Panel B or BB	37,344	−6.740	−0.600	0.010	0.400	2.200	−0.271	1.284	−2.308
Panel C or CC	37,588	−6.890	0.080	0.730	1.260	3.680	0.465	1.527	−2.022
Panel D or DD	37,673	−7.340	0.940	1.720	2.680	11.910	1.861	2.407	0.371
Panel E	150,514	−6.750	−0.200	0.230	1.200	6.570	0.426	1.756	−0.295
Fig. 5: **Earnings scaled by market value of equity (in dollars), by PPS Quartile**									
Panel A	37,667	−16.400	−0.497	−0.139	0.027	5.317	−0.564	2.222	−4.659
Panel B	37,661	−1.514	−0.132	0.002	0.079	0.527	−0.065	0.283	−2.374
Panel C	37,668	−0.566	0.006	0.054	0.089	0.302	0.031	0.121	−2.135
Panel D	37,661	−0.248	0.033	0.058	0.082	0.270	0.054	0.066	−1.093
Panel E	150,657	−3.799	−0.092	0.035	0.080	0.908	−0.102	0.548	−4.495
Fig. 7: **Unscaled earnings (in millions of dollars), by sales quartiles**									
Panel A	43,595	−72.650	−4.916	−1.039	0.080	13.018	−4.793	11.709	−3.480
Panel B	43,595	−99.273	−2.685	1.245	4.805	37.787	−1.356	17.017	−2.980
Panel C	43,596	−231.678	−1.440	8.661	24.294	164.952	8.769	48.140	−1.361
Panel D	43,595	−1,628.900	17.321	89.000	290.391	6,296.060	340.172	974.022	3.842

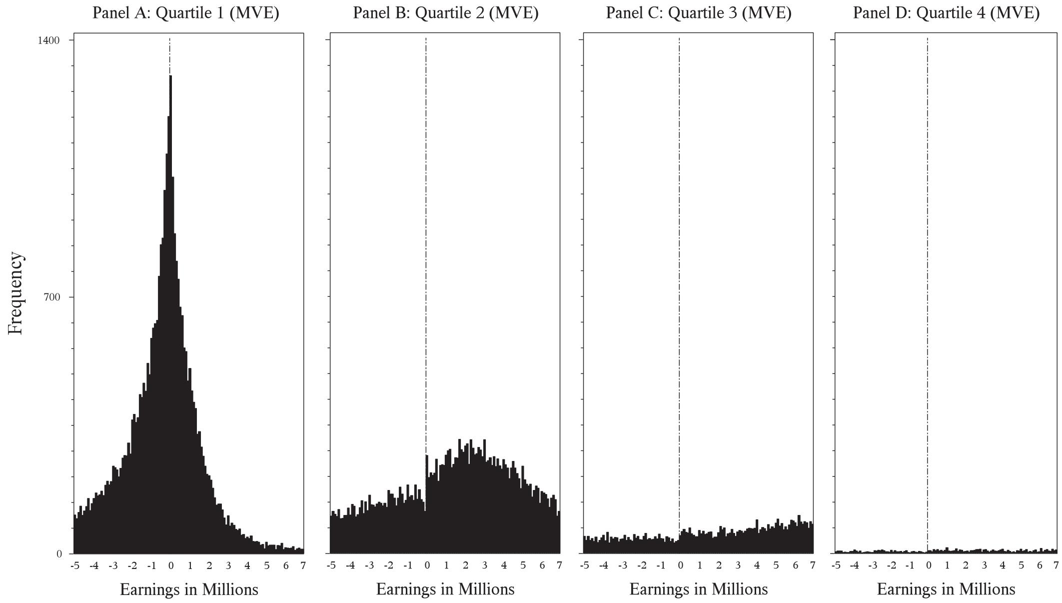
Notes: Table 1 tabulates the descriptive statistics for unscaled earnings, EPS, and market-value-scaled earnings for observations from Compustat for years 1990 to 2009, corresponding to Figs. 1–5, and 7. Figs. 1 and 2 plot unscaled earnings (Compustat item NI), partitioned into size quartiles based on beginning market capitalization (the product of lagged Compustat items PRCC_F and CSHO), so the descriptive statistics above for Figs. 1 and 2 pertain to unscaled earnings. For Figs. 1 and 2, Panels A and AA represent the smallest market capitalization quartile; Panels B and BB represent the second market capitalization quartile; Panels C and CC represent the third market capitalization quartile; and Panels D and DD represent the largest market capitalization quartile. Figs. 3 and 4 plot EPS (Compustat item EPSFX), partitioned into quartiles based on beginning price per share (lagged Compustat item PRCC_F), so the descriptive statistics for Figs. 3 and 4 pertain to EPS. For Figs. 3 and 4, Panels A and AA represent the smallest price per share quartile; Panels B and BB represent the second price per share quartile; Panels C and CC represent the third price per share quartile; Panels D and DD represent the largest price per share quartile. Fig. 5 plots earnings scaled by beginning market capitalization (Compustat item NI divided by the product of lagged PRCC_F and lagged CSHO), partitioned into quartiles based on beginning price per share (lagged Compustat item PRCC_F), so the descriptive statistics above for Fig. 5 pertain to earnings scaled by beginning market capitalization. For Fig. 5, Panel A represents the smallest price per share quartile; Panel B represents the second price per share quartile; Panel C represents the third price per share quartile; Panel D represents the largest price per share quartile. Fig. 7 plots unscaled earnings (Compustat item NI), partitioned into quartiles based on prior-year sales (lagged Compustat item SALE), so the descriptive statistics above for Fig. 7 pertain to unscaled earnings. For Fig. 7, Panel A represents the smallest sales quartile; Panel B represents the second sales quartile; Panel C represents the third sales quartile; Panel D represents the largest sales quartile.

difference is significantly positive. In contrast, results for Q3 and Q4 in Panels C and D do not show evidence of significant discontinuities, consistent with expectations if $100,000 is much smaller than the maximum amount of earnings that typically can be managed for firms in Q3 and Q4. For Q1, consistent with expectations if $100,000 is much larger than the maximum amount of earnings that typically can be managed for firms in Q1, results in Panel A show inconclusive evidence of a discontinuity—although the right standardized difference is nominally significant, the evidence is inconclusive because the peak of the distribution falls in the interval immediately above the benchmark.[9]

Fig. 2 Panel E shows explicitly how the combination of the four quartiles in Panels A–D results in the DE aggregate unscaled earnings distribution that does not show evidence of a discontinuity.[10] The bottom layer of each histogram bar corresponds to the very small number of observations from Q4 (from Panel D), the next layer corresponds to the slightly larger number of observations from Q3 (Panel C), the next layer to observations from Q2 (Panel B), and the top layer to observations from Q1 (Panel A). Because evidence of a discontinuity in Q1 is inconclusive and because the aggregate

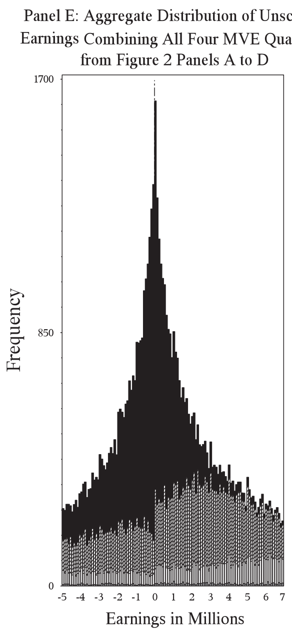
[9] When the peak of a distribution falls in the interval immediately above the benchmark, the standardized difference for the interval is typically significant. However, the significant standardized difference should be interpreted as inconclusive rather than as significant evidence of a discontinuity because (1) the assumption of smoothness is obviously violated for the interval at the peak and (2) it is plausible that the location of the distribution of pre-managed earnings just happened to fall in the interval to the right of the benchmark. This point is generally recognized in the literature. For example, in the Dechow et al. (2003) discussion on page 375 of their Figs. 8 and 9a, they point out that when the peak is at the zero benchmark "(t)his makes interpretation of the kink more difficult since it is unclear what number to expect at the peak of the distribution." Similarly, on page 377 they note: "When the mode is at zero it is much more difficult to evaluate whether the 'kink' is unusual." As another example, Jacob and Jorgensen (2007, p. 381) note that the standardized difference test assumes the expected frequency is the "mean of the actual frequency in the two neighboring partitions. This test is not appropriate to test for a discontinuity at the peak of the distribution."

[10] The aggregate distribution in Fig. 2 Panel E corresponds to distributions in DE1 Fig. 3 Panel B and DE2 Figs. 1, 6, and 8.



This figure shows histograms of unscaled earnings (Compustat item NI) partitioned into quartiles based on beginning market value of equity (the product of lagged Compustat items PRCC_F and CSHO). All panels use interval widths of $100,000 and the range of -$5m to $7m as in DE1 Figure 3 and DE2 Figures 1, 6, and 8. The vertical dashed line indicates the zero benchmark. Standardized differences, as defined in Burgstahler and Dichev (1997), are:

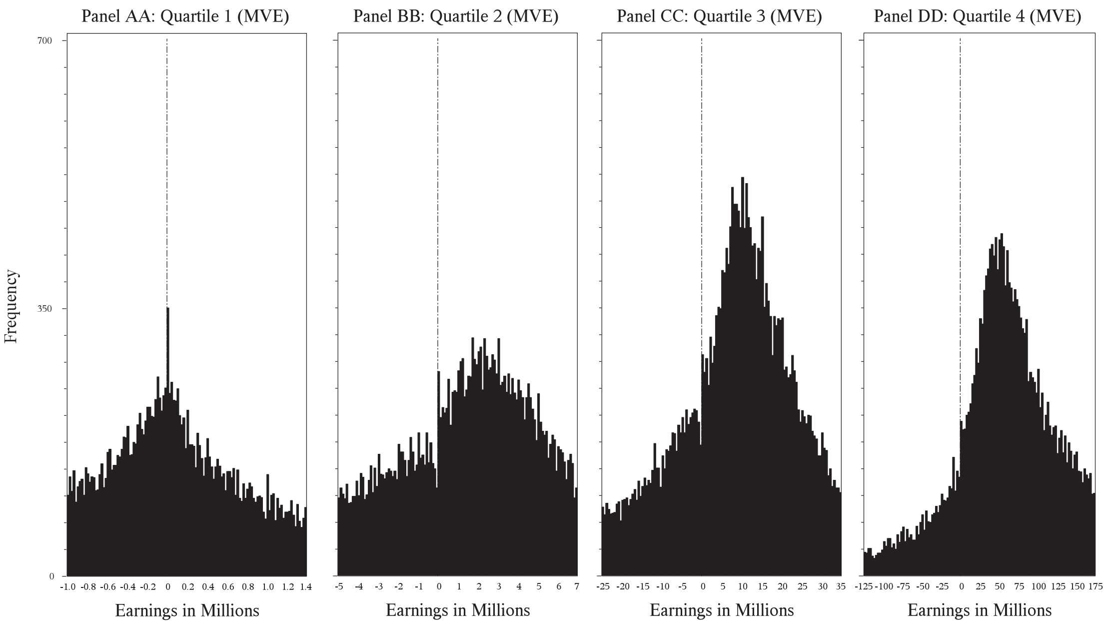
	Left Std Diff	Right Std Diff
Panel A	-0.107	4.507
Panel B	-6.015	5.690
Panel C	-0.665	0.117
Panel D	-0.392	0.603



This figure shows the histogram of unscaled earnings (Compustat item NI), combining all four MVE quartiles from Panels A to D of Figure 2. The lowest layer of the histogram is MVE Quartile 4 (Panel D), the second lowest layer is MVE Quartile 3 (Panel C), the third lowest layer is MVE Quartile 2 (Panel B), and the top layer is MVE Quartile 1 (Panel A). Standardized differences for the aggregate distribution, as defined in Burgstahler and Dichev (1997), are:

Left Std Diff	Right Std Diff
-2.186	6.353

Fig. 2. (Panels A – D) Distribution of unscaled earnings, by quartiles formed on beginning MVE using Durtschi and Easton's fixed interval width and range, which ignore differences in firm size (MVE). (Panel E) Aggregate distribution of unscaled earnings combining all four MVE quartiles from Panels A – D. (Panels AA – DD) Distribution of unscaled earnings, by quartiles formed on beginning MVE using interval width and range proportional to the median firm size (MVE) in each quartile.



This figure shows histograms of unscaled earnings (Compustat item NI), partitioned into quartiles based on beginning market value of equity (the product of lagged Compustat items PRCC_F and CSHO). The range and interval widths are adjusted across panels, such that the interval width is approximately 0.25% of the median market value of equity for the observations in the panel, resulting in interval widths of $20,000, $100,000, $500,000, and $2,500,000 in Panels AA, BB, CC, and DD, respectively. The vertical dashed line indicates the zero benchmark. Standardized differences, as defined in Burgstahler and Dichev (1997), are:

	Left Std Diff	Right Std Diff
Panel AA	-2.399	4.998
Panel BB	-6.015	5.690
Panel CC	-4.323	3.536
Panel DD	-2.772	2.503

Fig. 2. Continued.

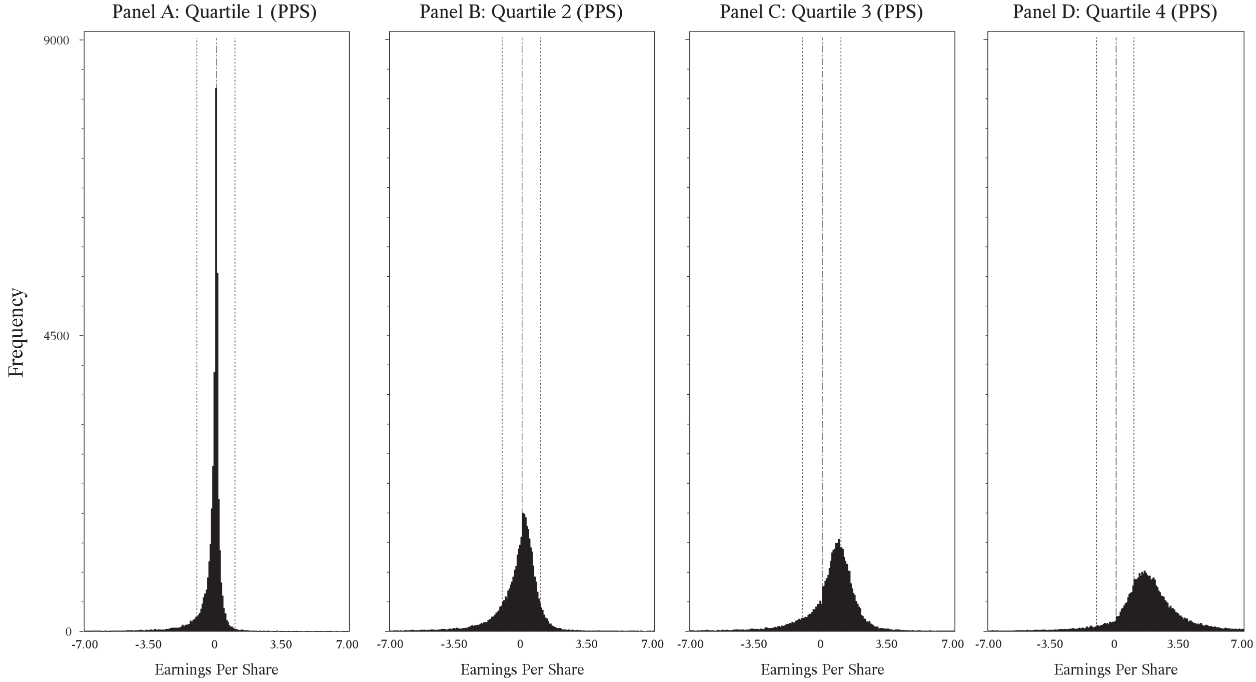
distribution is dominated by the large number of observations from Q1, evidence of a discontinuity in the aggregate distribution is also inconclusive.

We next consider distributions of unscaled earnings using a research design where interval widths are adjusted in proportion to MVE, under the assumption that the maximum earnings management varies with MVE. Three factors guide our attempt to adjust interval width in proportion to MVE—comparability with the interval widths used in DE, comparability with interval widths used in previous research, and materiality guidelines.

First, we choose interval widths that are simple multiples or fractions of the width used in the DE design so that the intervals in our design can be easily compared to the intervals in the DE design. Second, we choose interval widths that correspond roughly to the interval widths previously used in the discontinuity literature. For example, BD use interval widths of.5% of MVE for distributions of earnings levels and.25% of MVE for distributions of earnings changes and subsequent studies have commonly used similar interval widths. Third, we consider materiality, the amount of financial statement error permitted by auditing and reporting standards. Materiality guidelines are often specified as a percentage of firm size measures, such as assets, revenue, or earnings. To relate MVE to one of these common size measures, assume that MVE is a multiple of long-run expected earnings. Then materiality as a percentage of earnings, such as 5% (AICPA Audit and Accounting Manual, 1996; Newton, 1977; Leslie, 1985; Nelson, 1993; Bernard and Pincus, 1996; DeZoort et al., 2003; Nelson et al., 2005), translates into an approximate proportion of MVE. For example, if MVE is typically 10 to 20 times the long-run expected earnings, then 5% of expected earnings is about $5\% \times 1/10 = 0.5\%$ of MVE (for MVE that is 10 times the long-run expected earnings) or $5\% \times 1/20 = 0.25\%$ of MVE (for MVE that is 20 times the long-run expected earnings). From this perspective, 0.25–0.5% of MVE, the interval width commonly used in previous research, represents an amount of earnings management that typically could be accomplished at reasonably low cost given materiality thresholds.

Based on these three considerations, we use the DE interval width of $100,000 for Q2, the quartile where the DE interval width is closest to 0.25% of the median quartile MVE. For the other quartiles, we adjust interval width by simple multiples of the DE interval width and in approximate proportion to the ratio of MVE for the quartile to the MVE for Q2. Specifically, because the median MVEs in successive quartiles increase by a factor of approximately 5, we adjust the interval widths by factors of 5. The result is an interval width of $20,000 for Q1 in Panel AA (1/5 of the Q2 width), $500,000 for Q3 in Panel CC (5 times the Q2 width) and $2,500,000 for Q4 in Panel DD (25 times the Q2 width).[11]

[11] Discontinuity results depend on an interval width that corresponds approximately, but not precisely, to the maximum amount of earnings that typically can be managed. The interval widths used here are 0.27%, 0.20%, 0.21%, and 0.13% for the four quartiles, respectively so they all differ from 0.25% of MVE by no more than a factor of 2, i.e., they are no less than 1/2 of 0.25% and no more than 2 times 0.25%. In contrast, the DE interval width of $100,000



This figure shows histograms of EPS (Compustat item EPSFX), partitioned into quartiles based on beginning price per share (lagged Compustat item PRCC_F). The 25th, 50th, and 75th percentiles of PPS that define the four quartiles are $2.50, $9.31, and $22.00, respectively. All panels use the range -$7.00 to $7.00 of EPS with interval widths of $0.07 per share. The center dashed line indicates the zero benchmark. The outer dotted lines indicate the range -$1.00 to $1.00 used in DE1 Figure 1 Panel B, Figure 2, and Figures 6 through 10.

Fig. 3. Overview of distribution of EPS by quartiles formed on beginning price per share (PPS).

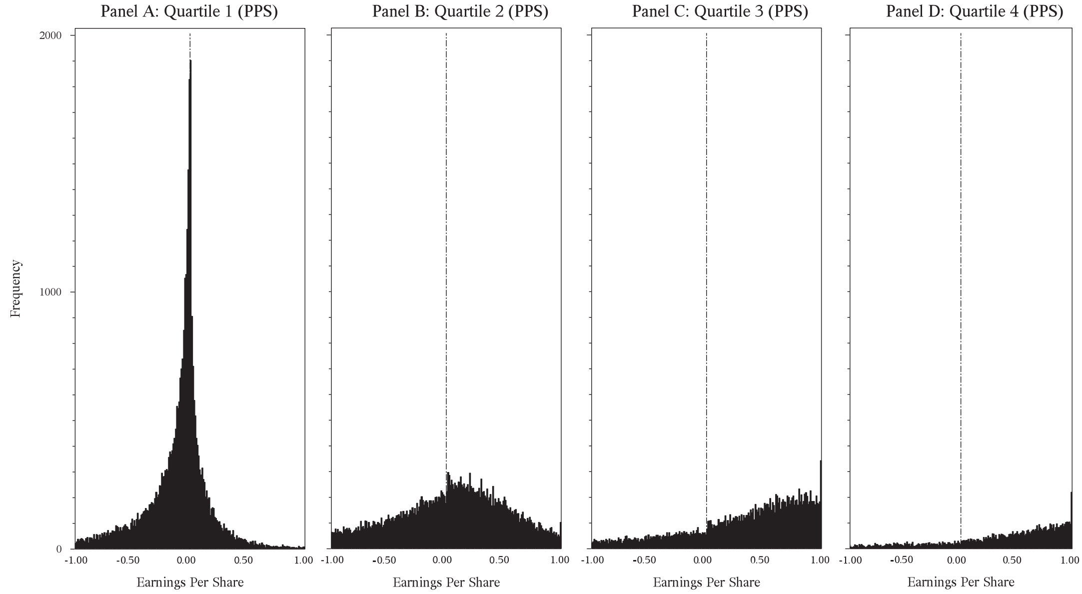
Using these interval widths that reflect size-related differences in the maximum earnings management, the results in Fig. 2 Panels AA–DD show highly significant discontinuities at the zero earnings benchmark in all four quartiles. For the upper three quartiles in Panels BB–DD, the benchmark is clearly to one side of the peak of the distribution and the interpretation of the highly significant discontinuities at the zero benchmark is clear-cut. For the first quartile in Panel AA where the benchmark falls near the peak of the distribution, results must be interpreted carefully. The number of observations in the first interval above the benchmark is so large relative to the numbers in either of the adjacent intervals that it seems unlikely that the large number in the first interval above the benchmark would be expected in any plausible null hypothesis distribution of unmanaged earnings. Thus, a slightly aggressive interpretation that invokes an additional assumption about plausible null hypothesis distributions is that the spike in the interval immediately above zero in Q1 is evidence of earnings management to meet the zero benchmark. However, the more cautious interpretation is that the results for Q1 are inconclusive because the peak of the Q1 distribution falls in the interval immediately above the benchmark. Therefore, under the more cautious interpretation, the results for Q1 are inconclusive.

3.1.2. Earnings scaled by number of shares (earnings per share)

Next, we perform a parallel analysis for earnings per share (EPS) because DE1 “argue that firms do not manage earnings deflated by financial variables; and, although it is possible that firms manage earnings per share, the distribution of earnings per share does not show evidence of earnings management.” (DE2, p. 1265) Because there is no discontinuity in the distribution of EPS in DE1 Fig. 1 Panel B but a distinct discontinuity in the distribution of price-scaled earnings in DE1 Fig. 1 Panel A, DE1 infer that scaling *causes* the discontinuity. However, we show below that in the EPS distributions shown in DE, discontinuities are obscured by DE's research design choices, for reasons similar to those discussed above for unscaled earnings.

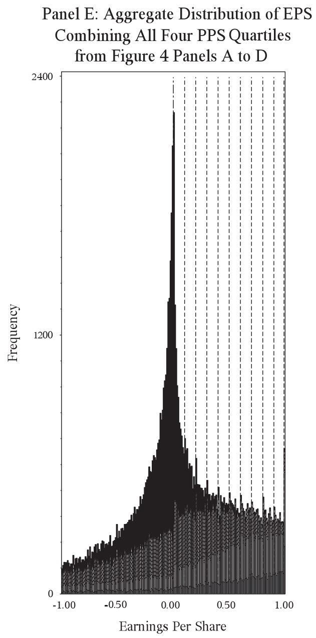
We again start with an overview of the distributions by quartile, where the quartiles are now defined in terms of the scaler advocated by DE, price per share (PPS). Fig. 3 shows distributions of EPS partitioned into quartiles based on beginning PPS, where the dotted lines indicate the range of EPS included in the DE distributions, –$1.00 to $1.00. As shown in Table 1, the mean and median EPS for the lowest PPS quartile are negative. Because of the expected positive relation between PPS and EPS, and similar to the results in Fig. 1, the distributions of EPS for firms in the larger PPS quartiles have larger means and greater dispersion.

(*footnote continued*)
differs from 0.25% of MVE by much larger factors for Q1, Q3, and Q4. The DE interval width exceeds 0.25% of Q1 median MVE by a factor of more than 5 at 1.33%, and is about 1/5 of 0.25% of median MVE in Q3 at 0.04% and about 1/50 of 0.25% of median MVE in Q4 at 0.005%.



This figure shows histograms of EPS (Compustat item EPSFX), partitioned into quartiles based on beginning price per share (lagged Compustat item PRCC_F). All panels use the interval width of $.01 as in DE1 Figure 1 Panel B, Figure 2, and Figures 6 through 10. The vertical dashed line indicates the zero benchmark. The first (second) interval right of the dashed line corresponds to EPS of $0.00 ($0.01). Standardized differences, as defined in Burgstahler and Dichev (1997), are

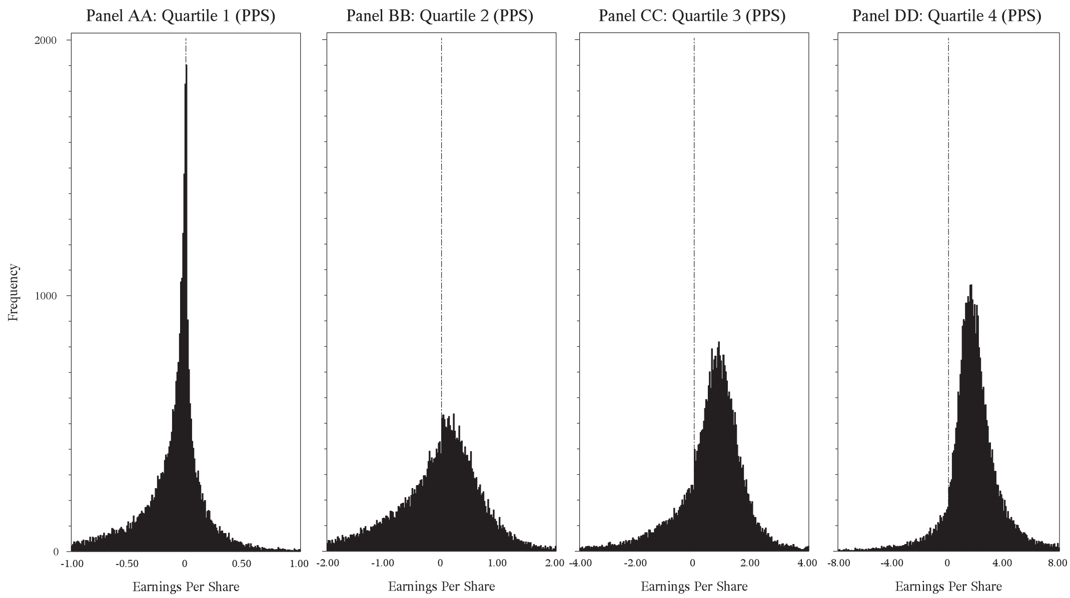
	Left Std Diff	Right Std Diff
Panel A	2.705	10.589
Panel B	-2.168	-0.758
Panel C	-1.646	-0.372
Panel D	-1.559	0.781



This figure shows the histogram of EPS (Compustat item EPSFX), combining all four PPS quartiles from Panels A to D of Figure 4. The lowest layer of the histogram is PPS Quartile 4 (Panel D), the second lowest layer is PPS Quartile 3 (Panel C), the third lowest layer is PPS Quartile 3 (Panel B), and the top layer is PPS Quartile 1 (Panel A). The vertical dashed lines indicate 10 cent multiples of EPS (e.g., $.10, $.20, $.30, ..., $1.00). Standardized differences, as defined in Burgstahler and Dichev (1997), are:

Benchmark	$0.00	$0.10	$0.20	$0.30	$0.40	$0.50	$0.60	$0.70	$0.80	$0.90	$1.00
Left Std Diff	1.426	-1.569	-2.873	-1.759	-3.348	-0.952	-1.743	-2.764	-3.461	-1.115	-4.184
Right Std Diff	9.472	1.919	4.132	2.384	4.500	1.708	2.367	2.183	3.782	2.059	2.913

Fig. 4. (Panels A – D) Distribution of EPS, by quartiles formed on beginning price per share (PPS) using Durtschi and Easton's fixed interval width and range, which ignore differences in PPS. (Panel E) Aggregate distribution of EPS combining all four PPS quartiles from Panels A–D. (Panels AA–DD) Distribution of EPS, by quartiles formed on beginning price per share (PPS) using interval width and range proportional to the median pps in each quartile.



This figure shows histograms of EPS (Compustat item EPSFX), partitioned into quartiles based on beginning price per share (lagged Compustat item PRCC_F). The range and interval widths are adjusted across panels, such that the interval width is approximately 0.25% of the median PPS for the observations in the panel, resulting in interval widths of $0.01, $0.02, $0.04, and $0.08 in Panels AA, BB, CC, and DD, respectively. The vertical dashed line indicates the zero benchmark. Standardized differences, as defined in Burgstahler and Dichev (1997), are:

	Left Std Diff	Right Std Diff
Panel AA	2.705	10.589
Panel BB	-3.747	2.325
Panel CC	-4.407	4.387
Panel DD	-2.110	1.579

Fig. 4. Continued

Fig. 4 Panels A–D show the distributions of EPS for each of the four PPS quartiles, using the DE interval width of $0.01 and the DE range of –$1.00 to $1.00 for all quartiles, which does not allow for differences in the amount of EPS that can be managed at a cost lower than the benefits. The distributions using the DE design provide only limited evidence of discontinuities, consistent with what is expected if $0.01 does not correspond to the typical maximum amount of earnings that can be managed for firms in these distributions. Although both the second and third quartiles provide visual evidence of discontinuities, the standardized differences are insignificant or only moderately significant.

Fig. 4 Panel E shows explicitly why the aggregate distribution of EPS reported in DE does not show evidence of a discontinuity. The Panel E aggregate distribution is the sum of the Panel D distribution (the bottom layer of each histogram bar), the Panel C distribution (the second layer), and the Panels B and A distributions (the third and fourth layers, respectively).[12] As with unscaled earnings, the preponderance of observations included in the DE aggregate EPS distribution are from the smallest size quartile. Specifically, in the intervals immediately adjacent to the zero benchmark, there are about 10 times as many observations from Q1 as from Q2, and roughly 40 times as many as from Q3, and 80 times as many as from Q4. Because it is dominated by observations from Q1, and because evidence of a discontinuity at the benchmark in Q1 is inconclusive because the peak of the distribution is in the interval immediately above the benchmark, the aggregate distribution is also inconclusive.

Before turning to results for the design that adjusts interval widths to reflect size-related differences in the maximum amount of earnings that can be optimally managed, we note another prominent feature of the EPS distribution that is consistent with the theory that earnings are managed to meet benchmarks but inconsistent with the DE alternative explanations for discontinuities. Carslaw (1988), Thomas (1989), Das and Zhang (2003), Dechow et al. (2003), and Grundfest and Malenko (2009) posit and report evidence of earnings management to achieve rounding in EPS, such as earnings management to increase earnings so that EPS rounds up from positive amounts ending in multiples of nine cents to the next highest multiple of 10 cents per share. Consistent with these previous results, the dashed vertical lines in Fig. 4 Panel E indicate the intervals corresponding to EPS benchmarks at multiples of 10 cents per share (i.e., $0.10, $0.20, …, $1.00 per share). Many of the discontinuities at the various 10 cent per share benchmarks are significant.[13] Of particular note are the highly significant left and right standardized differences at the $1.00 EPS benchmark, consistent with particularly strong

[12] The dashed line at the center of the distribution indicates the EPS=$0.00 benchmark. The other dashed lines, indicating EPS benchmarks at multiples of $0.10, are discussed in the next paragraph.

[13] Seven of ten left standardized differences are significantly negative at the 0.05 level while nine of ten right standardized differences are significantly positive. A more powerful test for the "rounding up" form of earnings management can be constructed based on the average of the standardized differences across the ten individual non-zero benchmarks and this test yields results that are still more significant. The standardized average computed from the ten left standardized differences in Fig. 5 is -7.53 and the standardized average for the ten right standardized differences is 8.85.

incentives to manage earnings such that earnings per share rounds up to a whole dollar.[14] The significant discontinuities at non-zero benchmarks cannot be explained by the scaling explanation offered by DE and, as discussed further below, are also inconsistent with the other DE explanations.[15]

Returning to the issue of discontinuities at the zero benchmark, we reexamine the EPS distributions using interval widths that are adjusted to reflect size-related differences in the maximum amount of earnings that typically can be managed. We set the interval width for Q1 equal to \$0.01, the interval width used in DE. (Although the \$0.01 interval width is 1.33% of median PPS in Q1, which is much larger than 0.25% of PPS, a narrower interval width would introduce additional issues due to calculation of EPS in fractional cents.) For the other three quartiles, we set the interval widths to \$0.02, \$0.04, and \$0.08, respectively, which are simple multiples of the DE interval width yet are roughly equal to 0.25% of median PPS for each quartile.[16] Fig. 4 Panels AA–DD show the resulting distributions of EPS. Quartiles 2, 3, and 4 show significant discontinuities and are discussed first. Then we turn to results for Quartile 1, which require additional exploration.

For Q2, Q3, and Q4 in Panels BB–DD, there are significant discontinuities and the interpretation of the evidence for these quartiles is straightforward. For Q2 and Q3, the left and right standardized differences are all significant at the 0.01 level. For Q4 in Panel DD, the left standardized difference is significant at the 0.05 level and the right standardized difference is significant at the 0.10 level.

For Q1 in Panel AA using an interval width of \$0.01, the evidence is inconclusive because the peak of the distribution falls in the interval immediately above the zero benchmark. Thus, the distribution for firms in Q1 neither confirms nor contradicts the theory that earnings are managed to meet the benchmark. However, further investigation reveals that the lack of evidence of a discontinuity among firms in Q1 is driven by a lack of discontinuity among firms with very low PPS < \$1.00. Specifically, unreported results show there is no discontinuity in the distribution of EPS for firms with PPS < \$1.00 but there is a significant discontinuity for firms with PPS > \$1.00.[17] The lack of discontinuity for firms with PPS < \$1.00 is not surprising, as these firms face substantial costs imposed by institutional factors. For example, shares with PPS < \$1.00 are typically not allowed to trade on major exchanges, many mutual funds operate under rules limiting investments in shares with low PPS, and low PPS shares usually are not eligible for margin purchases. Thus, firms that have not been able to maintain PPS above \$1.00 are more likely to be in financial difficulty, a conjecture supported by additional descriptive statistics reported in Table 2. About 15% of firms in PPS Q1 have no revenue and about 26% have negative book value of equity. These proportions are even larger among the subset of Q1 firms with PPS < \$1.00, as about 20% of these firms have no revenue and about 35% have negative book value. Firms in financial difficulty are likely to realize lower benefits and face higher costs of earnings management, resulting in a lower maximum amount of earnings that typically can be managed and therefore a lower rate of earnings management. To maintain comparability with DE, our analysis retains these observations without positive revenue, assets, or book value. However, because these firms likely have a lower rate of earnings management, it is important to remember that retention of these observations dilutes evidence of discontinuities due to earnings management, and this dilution is likely to affect primarily results for PPS Q1.

3.1.3. Summary of the effects of scaling

In this section, we have shown why there are not significant discontinuities in the distributions of unscaled earnings and EPS reported by DE. The DE research designs, which do not account for the substantial effect of firm size as a covariate and do not account for the relation between size and the amount of earnings that can be managed at a cost lower than the benefits, obscure evidence of discontinuities. In contrast, a research design that accounts for these two fundamental effects *but without scaling* reveals significant discontinuities in distributions of unscaled earnings and EPS.

Although the size-partitioned research designs are useful to definitively establish that discontinuities are not the result of scaling, the conceptual and practical advantages of a design based on scaled earnings should also be recognized. Scaling adjusts each observation by its own individual size, whereas partitioning into size quartiles only partially accounts for the effect of size, holding size only roughly constant within each quartile. Further, scaling allows all observations to be aggregated into a single distribution, resulting in a more powerful, unified test rather than four separate distributions and four separate tests, each based on one-fourth the total sample size.[18] The results above (and additional results discussed in the next section) show that there is no empirical support for conjectures that scaling might explain the discontinuities in earnings distributions. Thus, in order to achieve a larger effective sample size and a closer matching of interval width and

[14] The discontinuity in the aggregate distribution at EPS = \$1.00 is also visually apparent in each of the individual quartile distributions shown in Panels B–D. (There is no evidence of a discontinuity at EPS = \$1.00 in Panel A because there are very few observations from the first PPS quartile with EPS close to \$1.00.)

[15] DE1 footnote 20 acknowledges that discontinuities exist at multiples of 10 cents per share but does not explain how these discontinuities could be explained by any of the "alternative, much more likely, explanations" advocated in DE.

[16] Specifically, these interval widths represent 0.374% of median PPS for Q2 (\$0.02/\$5.34), 0.271% for Q3 (\$0.04/\$14.75) and 0.237% for Q4 (\$0.08/\$33.75), which all differ from.25% of median PPS by no more than a factor of 2. For comparison, the DE interval width of \$0.01 represents 0.18% for Q2, 0.066% for Q3, and 0.03% for Q4. Thus, the DE interval width of \$0.01 differs from 0.25% of median PPS by a factor of about 1/4 in Q3 and about 1/8 in Q4. In Q2, our interval width and the DE interval width both differ from 0.25% of PPS by less than a factor of 2, and both yield significant discontinuities.

[17] Untabulated results show that the same conclusion applies to distributions of scaled earnings. Specifically, there is little evidence of a discontinuity in distributions of scaled earnings for firms with PPS < 1.00 while there is a significant discontinuity for firms in Q1 with PPS > \$1.00.

[18] Increased power that results from combining results into a single distribution is illustrated by the still more significant results discussed below for the aggregate distribution of MVE-scaled earnings in Fig. 5.

Table 2
Percentage of sample with missing, zero, or negative values for assets, sales, and book value.

	PPS Q4	PPS Q3	PPS Q2	PPS Q1	PPS Q1 PPS ≥ $1	PPS Q1 PPS < $1	Missing PPS
Number of observations	37,673	37,588	37,344	37,909	15,705	22,204	35,962
Panel A: Total assets							
% Firms With Missing Total Assets	0.04%	0.09%	0.04%	0.05%	0.04%	0.05%	0.16%
% Firms With Zero Total Assets	0.00%	0.00%	0.05%	0.99%	0.12%	1.61%	0.51%
% Firms With Negative Total Assets	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Total	**0.05%**	**0.09%**	**0.09%**	**1.04%**	**0.17%**	**1.66%**	**0.66%**
Panel B: Sales							
% Firms with missing sales	0.03%	0.07%	0.02%	0.03%	0.04%	0.02%	0.07%
% Firms with zero sales	0.20%	0.79%	3.04%	14.56%	7.18%	19.78%	7.35%
% Firms with negative sales	0.03%	0.07%	0.11%	0.15%	0.12%	0.18%	0.09%
Total	**0.27%**	**0.93%**	**3.17%**	**14.75%**	**7.35%**	**19.98%**	**7.50%**
Panel C: BVE							
% Firms With Missing BVE	0.23%	0.24%	0.15%	0.08%	0.06%	0.10%	0.39%
% Firms With Zero BVE	0.00%	0.01%	0.01%	0.18%	0.07%	0.25%	0.17%
% Firms With Negative BVE	1.34%	2.10%	5.90%	25.70%	13.03%	34.66%	19.36%
Total	**1.56%**	**2.35%**	**6.06%**	**25.96%**	**13.16%**	**35.02%**	**19.92%**

Notes: Table 2 tabulates Compustat data for years 1990 to 2009. Panel A presents the percentage of the sample with missing, zero, or negative values for total assets. Panel B presents the percentage of the sample with missing, zero, or negative values for sales. Panel C presents the percentage of the sample with missing, zero, or negative values for book value. The first column summarizes observations from price per share quartile 4 (largest price per share). The second column summarizes observations from price per share quartile 3. The third column summarizes observations from price per share quartile 2. The fourth column summarizes observations from price per share quartile 1 (smallest price per share). The fifth column summarizes observations from price per share quartile 1 with price per share values of $1 or higher. The sixth column summarizes observations from price per share quartile 1 with price per share values of less than $1. The seventh column summarizes observations with missing price per share data.

the amount of earnings that can be managed at a cost less than the benefits for each individual firm, we recommend that future researchers use earnings scaled by MVE. While other size scalers often yield similar results, we believe the direct link between MVE and long-run expected earnings in valuation models makes MVE the best choice for a size scaler.

3.2. Relation between earnings and market prices as an alternative explanation

DE1 Fig. 6 shows that the PPS associated with observations of small positive EPS tend to be higher than the PPS associated with observations of small negative EPS. Based on this finding, DE1 propose a mechanism that could create a discontinuity in the distribution of price-scaled earnings:

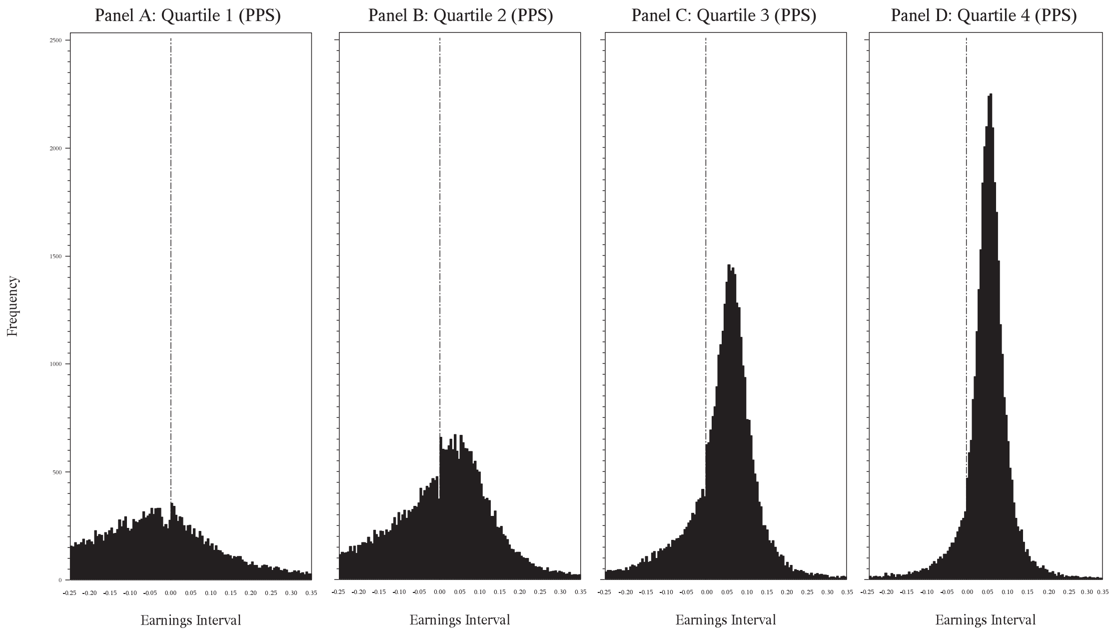
> “These results suggest that the firms reporting a small profit tend to have a higher beginning-of-year price, and, when net income is deflated by this higher price, relatively more of the earnings observations are drawn toward zero, whereas the lower beginning-of-year price for firms reporting a loss tends to push relatively more of the earnings observations away from zero.” (p. 573)

DE assert that the discontinuity at the zero benchmark in scaled earnings is explained by the higher price scalers associated with positive unscaled earnings than with negative unscaled earnings, which we refer to as selective scaling. However, as explained further below, the evidence in DE1 Fig. 6 does not show that selective scaling occurs across all observations that make up the distribution of scaled earnings. Rather, selective scaling occurs in just one subset of observations and this subset plays only a small role in the discontinuity in the distribution of scaled earnings. The discontinuity in scaled earnings is due primarily to observations where there is no evidence of selective scaling.[19]

3.2.1. Discontinuities in distributions of scaled earnings with larger PPS

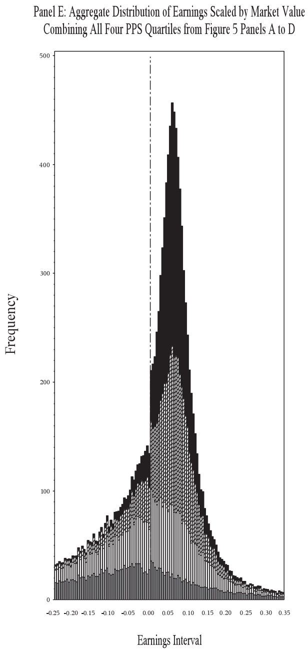
We begin by documenting which segments of the population account for the discontinuity in scaled earnings by extending the analysis of unscaled earnings in Figs. 3 and 4. We segment the distribution of scaled earnings into four component distributions corresponding to the four PPS quartiles. Fig. 5 Panels A–D show that there are discontinuities in the scaled earnings distributions from all four quartiles. Most small scaled earnings observations (i.e., scaled earnings in the intervals surrounding the zero benchmark where the discontinuity occurs) are from PPS Q2 and Q3, followed in order by Q4 and Q1. Focusing specifically

[19] DE2 Fig. 2 provides closely related evidence showing that small negative earnings observations tend to be scaled by lower market capitalization than small positive earnings observations. For the sake of brevity, the analysis in this section addresses only the relation between EPS and PPS documented in DE1 Fig. 6. However, a parallel analysis of the relation between unscaled earnings and market capitalization yields similar conclusions.



This figure shows histograms of earnings scaled by beginning market value of equity (Compustat item NI scaled by the product of lagged PRCC_F and lagged CSHO), partitioned into quartiles based on beginning PPS (lagged Compustat item PRCC_F). Standardized differences, as defined in Burgstahler and Dichev (1997), are:

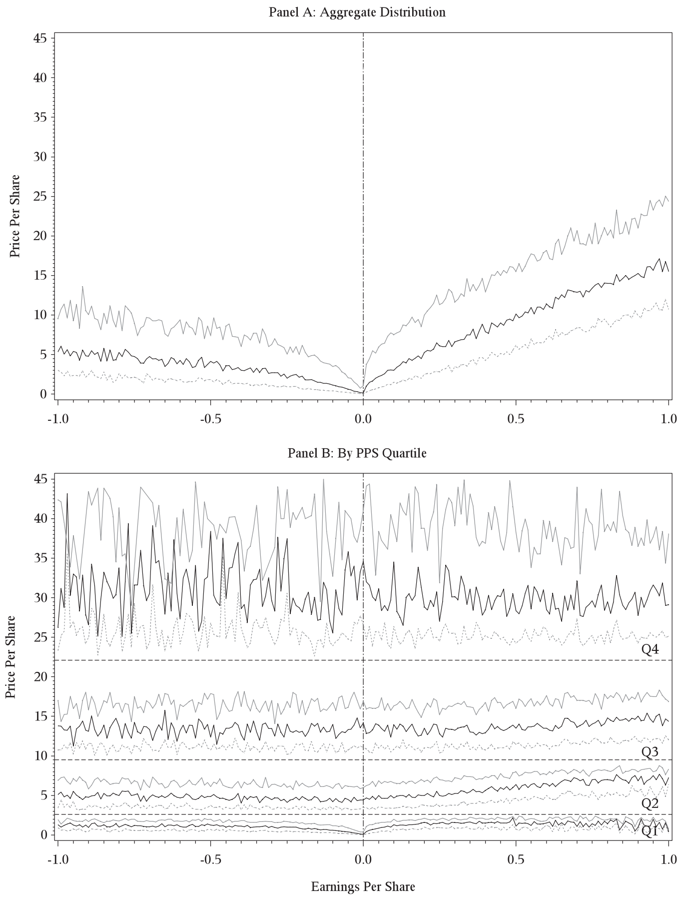
	Left Std Diff	Right Std Diff
Panel A	-1.021	2.083
Panel B	-7.607	5.705
Panel C	-5.433	3.907
Panel D	-2.793	0.704



This figure shows the histogram of earnings scaled by beginning market value of equity (Compustat item NI scaled by the product of lagged PRCC_F and lagged CSHO), combining all four PPS quartiles from Panels A to D of Figure 5. The lowest layer of the histogram is PPS Quartile 1 (Panel A), the second lowest layer is PPS Quartile 2 (Panel B), the third lowest layer is PPS Quartile 3 (Panel C), and the top layer is PPS Quartile 4 (Panel D). Standardized differences for the aggregate distribution, as defined in Burgstahler and Dichev (1997) are:

Left Std Diff	Right Std Diff
-8.870	6.496

Fig. 5. (Panels A–D) Distribution of earnings scaled by beginning market value, by quartiles formed on beginning price per share (PPS). (Panel E) Aggregate distribution of earnings scaled by market value, combining all four pps quartiles from Panels A to D.



Each figure plots the 25th (gray dotted line), 50th (black solid line), and 75th (gray solid line) percentiles of the distribution of beginning PPS associated with each one cent value of EPS. Panel A presents percentiles for the distribution that combines all observations, corresponding to DE1 Figure 6. Panel B presents the percentiles within each of the four PPS quartiles, and the dashed horizontal lines correspond to the values of PPS that separate the four quartiles, PPS=\$2.50, \$9.31, and \$22.00

Fig. 6. Distribution of beginning price per share for each one-cent earnings per share portfolio.

on the discontinuity, the difference between the number of observations in the interval immediately above versus the interval immediately below the zero benchmark is due most to observations from Q2 and Q3 followed in order by Q4 and Q1. The distribution of scaled earnings that combines results across all values of PPS (shown in Panel E and discussed in more detail later) has 762 more observations in the interval immediately above the benchmark than in the interval immediately below the benchmark, where the corresponding differences in the four component quartiles (shown in Panels A–D) are 80, 286, 241, and 155, respectively. Thus, the proportions of the discontinuity in scaled earnings due to each of the four quartiles are 10%, 38%, 32%, and 20%, respectively. The relative contributions from each quartile to the overall discontinuity are important in evaluating the selective scaling explanation, as discussed in the next section.

3.2.2. Evidence of selective scaling by PPS quartile

We next evaluate whether there is evidence of selective scaling for observations in all size quartiles or whether selective scaling is concentrated in only one quartile of the population. We begin by replicating the selective scaling evidence from DE1. Consistent with DE1, Fig. 6 Panel A shows that the PPS associated with small negative EPS appear to be consistently lower than the PPS associated with small positive EPS. Confirming the visual evidence of selective scaling, the difference between the mean PPS associated with EPS $= -0.01$ versus EPS $= 0.01$ in Panel A is significant at the 0.0001 level.

Although the evidence of selective scaling reported in DE1 and replicated here uses all earnings observations, it is important to recognize that the results for small magnitude EPS are dominated by observations with small PPS. To see this, compare Fig. 4 Panels A–D, where the preponderance of observations with small negative and small positive EPS are from the smallest quartile of PPS while only a small fraction of these small EPS observations are from the upper PPS quartiles. Thus, the selective scaling evidence in DE1 and in our Fig. 6 Panel A may not be due to selective scaling throughout the population, but could instead be due to selective scaling only among the smallest PPS observations. Because the results in Fig. 5 show that observations with the smallest values of PPS account for only a small portion (i.e., 10%) of the discontinuity in scaled earnings, it is essential to determine whether there is evidence of selective scaling for observations from the three upper quartiles of PPS, which account for 90% of the discontinuity.

Fig. 6 Panel B evaluates evidence of selective scaling in each of the individual quartiles based on beginning PPS. In Panel B, horizontal dashed lines indicate the three PPS values that separate the four quartiles, \$2.50, \$9.31, and \$22.00. Within each of the four quartiles, three lines plot the 25th, 50th, and 75th percentiles of PPS corresponding to each one-cent value of EPS between $-$\$1.00 and $+$\$1.00.[20] The plots in Panel B suggest that evidence of selective scaling is confined to Q1 while in Q2, Q3, and Q4 there is no systematic difference between the PPS associated with small negative EPS versus small positive EPS. The visual evidence in Panel B is confirmed by statistical tests. The difference between the mean PPS associated with EPS $= -0.01$ versus EPS $= 0.01$ is significant at the 0.0001 level for Q1 but none of the differences are significant at even the 0.10 level for Q2, Q3, or Q4. Thus, when observations are segmented into PPS quartiles, only Q1 shows evidence of the selective scaling effect reported in DE1, while the remaining three quartiles show no evidence of the selective scaling required for the DE mechanism to explain discontinuities in scaled earnings. Therefore, the DE selective scaling mechanism cannot account for the 90% of the discontinuity in scaled earnings due to observations from Q2, Q3, and Q4.[21]

To evaluate how much of the remaining 10% of the discontinuity is accounted for by selective scaling, we go one step further and separate the observations in PPS Q1 into those with PPS $<$ \$1.00 versus those with PPS $\geq$ \$1.00. The untabulated results show that of the 80 observation difference between the number of observations above versus below the discontinuity in Q1, about 84% (67/80) is due to observations with PPS $\geq$ \$1.00 while only 16% (13/80) is due to observations with PPS $<$ \$1.00, and (2) the difference between the mean PPS associated with EPS $= -0.01$ versus EPS $= 0.01$ is significant at the 0.0001 level for observations with PPS $<$ \$1.00 but is not significant at even the 0.10 level for observations with PPS $\geq$ \$1.00. Thus, evidence of selective scaling is concentrated in the portion of Q1 that contributes only 16% of the 10% of the discontinuity that is attributable to Q1 or only about 1.6% (16% $\times$ 10%) of the total discontinuity in scaled earnings.

Fig. 5 Panel E graphically illustrates the small role of the segment of the observations where there is evidence of selective scaling. The bottom layer of each histogram bar represents observations from PPS Q1. There is little evidence of discontinuity in the distribution of scaled earnings observations from PPS Q1 represented in this bottom layer, yet Q1 is the segment where any effect of selective scaling should be the strongest. More importantly, the inclusion of these observations from Q1 has almost no effect on the discontinuity in the combined distribution. Instead, most of the discontinuity in the combined distribution of scaled earnings is due to observations where there is no evidence of selective scaling, the observations in Q2, Q3, and Q4.

3.3. Sample selection as an alternative explanation

DE speculate that discontinuities in scaled earnings distributions might be attributable to selection effects because scaling by market value requires beginning market price. They report two closely-related empirical results: (1) A greater proportion of loss observations than profit observations are eliminated when missing price observations are excluded from the sample, and (2) discontinuities become weaker when missing price observations are included in the sample. Based on these results, DE infer that sample selection effects explain discontinuities. For example, they conclude, based on the distribution of earnings per share in DE1 Fig. 2 Panel B, that the discontinuities observed in distributions of scaled earnings are explained by selection criteria:

> "It follows from Fig. 2, panel B that, since a larger proportion of loss observations are deleted from the sample because of an inability to calculate beginning-of-year market capitalization, the discontinuity of the (deflated) earnings distribution at zero may be partially due to this sample selection criterion" (DE1, p. 566).

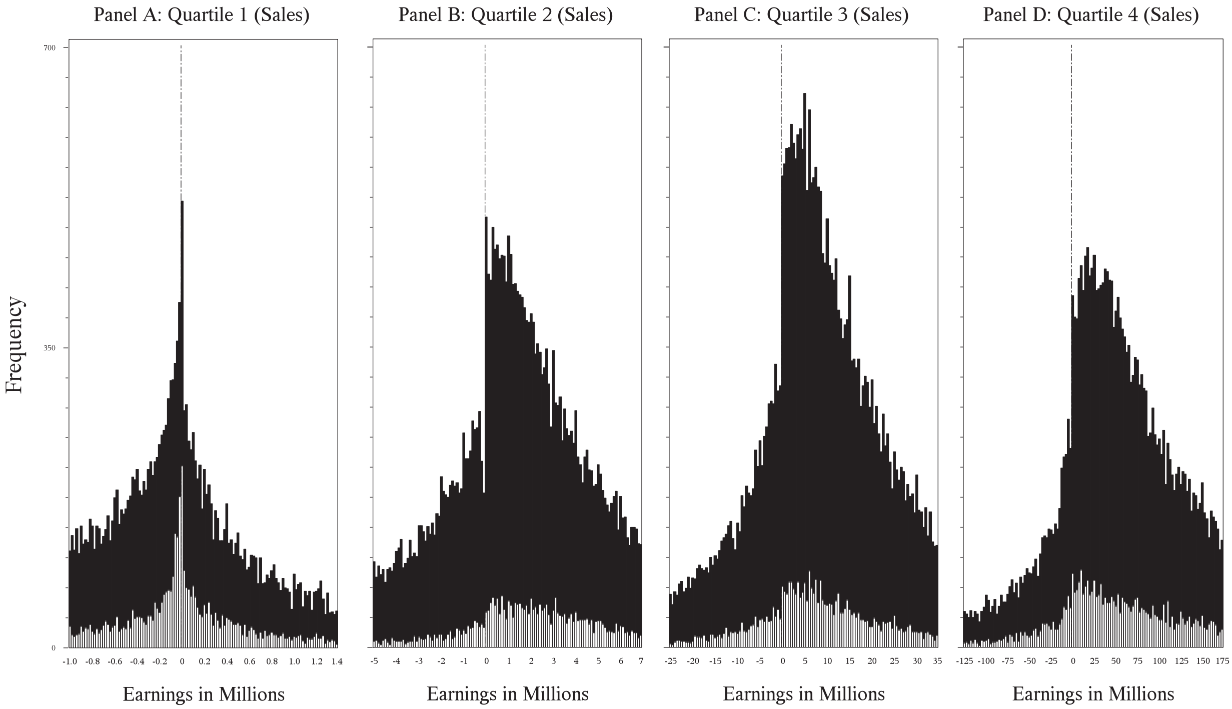
In this section, we evaluate sample selection as an explanation for discontinuities. We explain how the two empirical results reported by DE follow directly from the fact that earnings distributions for firms *without prices* have a large proportion of small negative earnings observations. This fact does not explain why discontinuities exist in earnings distributions for firms *with prices*.

The effects of selection criteria can be described in terms of (1) a core population that is always included, (2) an additional segment to be included or excluded depending on selection criteria, and (3) the extended population, comprising both the core and the segment. Using this terminology, the extended population is the result of inclusion of the segment with the core or, equivalently, the core is the result of exclusion of the segment from the extended population.[22] Applying

[20] Because the number of observations in the plotted range of EPS decreases substantially in higher PPS quartiles (as can be seen in Fig. 4 Panels A–D), there is substantially more random variation in the plotted percentiles for the higher PPS quartiles. For example, Fig. 4 Panel D shows that there are only a few small unscaled EPS observations between \$$-$1.00 and \$1.00 for Q4, so the 25th, 50th, and 75th percentiles of EPS in Fig. 6 Panel B show substantial random variation for Q4.

[21] As an alternative way to evaluate the selective scaling explanation, we conduct two additional tests and find that the mean PPS and mean MVE associated with small *scaled* losses in the interval immediately below the benchmark are not significantly different than the mean PPS and mean MVE associated with small *scaled* profits in the interval immediately above the benchmark, respectively. While these alternative tests require an additional assumption that selective scaling leads to differences in prices associated with positive versus negative *scaled* earnings, the results are consistent with the conclusion that selective scaling does not explain the discontinuity in scaled earnings. We thank the reviewer for suggesting these additional tests.

[22] Beaver et al. (2007) point out that whether selection based on the availability of price data represents a sample selection bias depends on whether the research focus is on listed firms or some other set of firms. Beaver et al. (2007) argue that researchers should exclude companies without price data if



This figure shows histograms of unscaled earnings (Compustat item NI), partitioned into quartiles based on prior-year sales (lagged Compustat item SALE). The range and interval widths are the same as in Figure 2 Panels AA to DD (i.e., interval widths of $20,000, $100,000, $500,000, and $2,500,000 in Panels A, B, C, and D, respectively). The vertical dashed line indicates the zero benchmark. The bottom layer represents observations without beginning price data, and the top layer represents observations with beginning price data. Standardized differences, as defined in Burgstahler and Dichev (1997), are:

	Left Std Diff	Right Std Diff
Panel A	-1.465	6.895
Panel B	-9.475	7.606
Panel C	-5.261	4.177
Panel D	-5.299	4.278

Fig. 7. Distribution of unscaled earnings, by quartiles formed on prior-year sales stacked histograms with nonmissing price observations (top layer) and missing price observations (bottom layer).

this terminology to the discontinuity setting, it is clear that the core distribution of earnings for firms with prices is different from the segment distribution for firms without prices, and that these differences explain why there is no discontinuity in the extended distribution consisting of firms with prices and firms without prices. In some cases, researchers may only be interested in conclusions that apply to the extended population. However, in many cases researchers will be interested in conclusions about the core population of firms with prices.

When the segment has a weaker discontinuity than the core (defined as a lower ratio of the number of observations in the interval immediately above the benchmark to the number in the interval immediately below the benchmark in the segment than in the core), the effects of including or excluding the segment are direct and predictable. Inclusion of the segment results in an extended distribution with a lower ratio and weaker discontinuity than the core. Similarly, exclusion of the segment leaves the core distribution and, because the segment has a weaker discontinuity than the core distribution, exclusion of the segment also by definition eliminates a larger proportion of observations below the benchmark than the proportion in the core distribution.[23]

(*footnote continued*)
the population of interest is companies that are publicly listed. Following similar logic, we define the core population as the observations with price data and we define the segment as the observations without price data. If the purpose is to draw conclusions about the population of publicly listed firms, then it is appropriate to treat firms with price as the core population and firms without price as a separate segment. More importantly, there are reasons to expect the segment of firms without price to differ from core firms with price, both in terms of their earnings management behavior and in terms of basic characteristics of the distribution of earnings (such as the mean and variances depicted in Fig. 7). From this perspective, it is appropriate to examine the core and segment distributions separately.

[23] To illustrate these effects with a numerical example, consider a core with a strong discontinuity (with 75 observations in the interval immediately below the benchmark and 175 observations in the interval immediately above the benchmark, and a ratio of $175/75 \cong 2.33$) and a segment with the weakest possible discontinuity, namely no discontinuity (with 125 observations below and 125 observations above the benchmark, and a ratio of observations above versus below the discontinuity of $125/125=1$). Inclusion of the segment results in an extended distribution with a weaker discontinuity and lower ratio than in the core. In the extended distribution, the ratio is $(175+125)/(75+125)=1.5$; in the core, the ratio is $175/75=2.33$. Thus, inclusion of the segment weakens the discontinuity in the extended distribution. Similarly, exclusion of the segment results in elimination of a larger proportion of observations below the benchmark than the proportion in the core distribution. Specifically, exclusion of the segment eliminates $125/200=62.5\%$ of observations below the benchmark, a proportion that is larger than the $125/300=41.7\%$ of observations from the extended distribution eliminated in the interval above the benchmark.

If discontinuities are attributable to the exclusion of observations that are missing price, then distributions of unscaled earnings should not contain discontinuities when missing price observations are included. We have seen that it is necessary to account for the effects of size in order to discern discontinuities in distributions of unscaled earnings, as in Fig. 2 Panels AA–DD. While we cannot directly evaluate the effect of missing price on the distribution of scaled earnings (because price-scaled earnings can only be calculated for non-missing price observations), we can evaluate the effect of missing price observations on the discontinuities in distributions of unscaled earnings using an alternative size scaler that does not require price, namely sales.

Fig. 7 shows distributions of unscaled earnings similar to Fig. 2 Panels AA–DD, except the quartiles are defined by prior-year sales (as an alternative measure of firm size), so that Fig. 7 can provide evidence on the effect of excluding observations with missing price. In each quartile, the bottom layer represents observations with missing beginning price data, while the top layer represents observations with nonmissing beginning price data. As in Fig. 2 Panels AA–DD, Q2, Q3, and Q4 in Fig. 7 all have significant discontinuities. Further, it is clear that the missing price observations play a very limited role in these discontinuities. The plots in Fig. 7 Panels B–D show that most of the discontinuity is due to observations with price while the observations that are missing price contribute little to the discontinuity. The standardized differences in these three quartiles are highly significant whether the missing price observations are included or excluded.

The effects of including or excluding the segment of firms with missing price follow directly from the fact that there is less evidence of discontinuities for firms with missing price. Inclusion of the segment results in a weaker discontinuity in the extended distribution and exclusion of the segment results in elimination of a larger proportion of observations below the benchmark.[24] However, these direct effects of a weaker discontinuity for the segment of firms that are missing price only show that distributions for the segment of missing price observations have weaker discontinuities. The lack of evidence of a discontinuity in the distribution of unscaled earnings for firms with missing prices in the smallest size quartile does not explain the existence of discontinuities in the upper quartiles, where there are significant discontinuities whether missing price observations are included or excluded.

4. Conclusion

A vast body of literature documents discontinuities in earnings distributions at prominent earnings benchmarks. Evidence in previous papers shows that the discontinuities differ across segments of the population (for example, BD show that discontinuities differ across segments with differing recent earnings history) and additional evidence in this paper shows that there is less evidence of discontinuities for firms in financial difficulty as proxied by very low price per share. Discontinuity evidence has been widely interpreted as consistent with the theory that managers take actions to avoid small losses. This interpretation is supported by survey evidence indicating that managers are willing to incur real costs to meet benchmarks.

Two papers by Durtschi and Easton show that the discontinuity in distributions of earnings can be eliminated by (1) using unscaled earnings or earnings scaled by number of shares, *and* (2) adopting modified research designs that emphasize results for firms with low prices and that do not account for size-related differences in the amount of earnings that can be managed at a cost lower than the benefits. Based on these results, DE infer that discontinuities in earnings distributions are "driven by sample selection bias and scaling" and that "deflation, sample selection, and a difference between the characteristics of profit and loss observations" are "much more likely" explanations for discontinuities than the widely-accepted theory that discontinuities are the result of earnings management to meet benchmarks.

The assertions in DE caused many researchers to question the appropriateness of using discontinuities as a proxy for earnings management, as has been common in previous research (e.g., Marquardt and Wiedman, 2004; Roychowdhury, 2006; Frank and Rego, 2006; Bergstresser and Philippon, 2006; Gleason and Mills, 2008; Dhaliwal et al., 2004). However, the discussion and analysis in our paper show that discontinuities are not attributable to any of the DE explanations. Results in Gilliam et al. (2013), which show that the zero-earnings discontinuity disappears after the Sarbanes-Oxley Act of 2002 (SOX) but the data artifacts have not changed, cast further doubt on the DE explanations. In particular, Gilliam et al. (2013) show that evidence of discontinuities is eliminated following 2002 but scaling and selection artifacts are held constant. For these reasons, the evidence in DE does not provide a valid reason to abandon the use of discontinuity evidence to identify firms that are more likely to have managed earnings.

Based on theory and principles of research design, we show how the DE research design choices obscure evidence of discontinuities because they (1) do not account for the extremely important effect of firm size as a covariate, (2) place inordinate weight on results for small firms, (3) and use a research design that does not recognize that the amount of earnings that can be managed at a cost lower than the benefits is much smaller for small firms. As a result, the DE designs

[24] These effects are further amplified by the DE research designs which place high weight on observations with small-magnitude unscaled earnings because most of the observations in the narrow range emphasized by the DE designs are from the smallest size quartile, as can be seen by close inspection of Fig. 7. Fig. 7 immediately reveals that there are more observations in the vicinity of the zero benchmark in the smallest size quartile than in the upper three quartiles. However, the difference in the number of observations is actually much larger than it first appears because the three upper quartiles of Fig. 7 use intervals that are wider than those in Q1 by factors of 5, 25, and 125, respectively. Therefore, the relative numbers of observations in the intervals surrounding zero in these three quartiles are roughly 1/5, 1/25, and 1/125 as large as they might first appear. In sum, there are far more missing price observations in the vicinity of the zero benchmark in Q1 than in the three upper quartiles.

move almost all the evidential weight to segments of the population where evidence of discontinuities is predictably weak or inconclusive. DE interpret inconclusive evidence as evidence that contradicts the alternative hypothesis that earnings are managed, but inconclusive evidence can neither confirm nor contradict any hypothesis. The more appropriate interpretation is that inconclusive and insignificant evidence of discontinuities at the zero benchmark in DE is the direct and predictable result of their research designs.

References

American Institute of Certified Public Accountants (AICPA), 1996. AICPA Audit and Accounting Manual. AICPA, New York, NY.

Beatty, A., Ke, B., Petroni, K., 2002. Earnings management to avoid earnings declines across publicly and privately held banks. The Accounting Review 77, 547–570.

Beaver, W., McNichols, M., Nelson, K., 2007. An alternative interpretation of the discontinuity in earnings distributions. Review of Accounting Studies 12, 525–556.

Bergstresser, D., Philippon, T., 2006. CEO incentives and earnings management. Journal of Financial Economics 80, 511–529.

Bernard, R.A., Pincus, K.V., 1996. The relationship between materiality thresholds and judgments of fraud risk. Managerial Finance 22, 1–15.

Burgstahler, D., and M. Eames. "Management of Earnings and Analysts' Forecasts to Achieve Zero and Small Positive Earnings Surprise." Journal of Business Finance and Accounting. Vol. 33, No.5-6, (June/July 2006): 633–652.

Burgstahler, D. And E. Chuk, 2014. Detecting earnings management using discontinuity evidence. Unpublished working paper. University of Washington.

Burgstahler, D., Dichev, I., 1997. Earnings management to avoid earnings decreases and losses. Journal of Accounting and Economics 24, 99–126.

Carslaw, C., 1988. Anomalies in income numbers: evidence of goal oriented behavior. The Accounting Review 63, 321–327.

Das, S., Zhang, H., 2003. Rounding-up in reported EPS, behavioral thresholds, and earnings management. Journal of Accounting and Economics 35, 31–50.

Daske, H., Gebhardt, G., McLeay, S., 2006. The distribution of earnings relative to targets in the European Union. Accounting and Business Research 36 (3), 137–167.

Dechow, P., Richardson, S., Tuna, I., 2003. Why are earnings kinky? An examination of the earnings management explanation. Review of Accounting Studies 8, 355–384.

Degeorge, F., Patel, J., Zeckhauser., R., 1999. Earnings management to exceed thresholds. The Journal of Business 72, 1–33.

DeZoort, F.T., Hermanson, D.R., Houston, R.W., 2003. Audit committee support for auditors: the effects of materiality justification and accounting precision. Journal of Accounting and Public Policy 22, 175–199.

Dhaliwal, D., Gleason, C., Mills, L., 2004. Last-chance earnings management: using the tax expense to meet analysts' forecasts. Contemporary Accounting Research 21, 431–459.

Dichev, I., Graham, J., Harvey, C., Rajgopal, S., 2013. Earnings quality: evidence from the field. Journal of Accounting and Economics 56, 1–33.

Donelson, D., McInnis, J., Mergenthaler, R., 2013. Discontinuities and earnings management: evidence from restatements related to securities litigation. Contemporary Accounting Research 30, 242–268.

Durtschi, C., Easton, P., 2005. Earnings management? The shapes of the frequency distributions of earnings metrics are not evidence ipso facto. Journal of Accounting Research 43, 557–592.

Durtschi, C., Easton, P., 2009. Earnings management? Erroneous inferences based on earnings frequency distributions. Journal of Accounting Research 47, 1249–1281.

Frank, M., Rego, S., 2006. Do managers use the valuation allowance account to manage earnings around certain earnings targets? Journal of the American Taxation Association 28, 43–66.

Gilliam, T., F. Heflin, and J. Patterson, 2013. Evidence that the zero-earnings discontinuity has disappeared. Journal of Accounting and Economics, forthcoming.

Gleason, C., Mills, L., 2008. Evidence of differing market responses to beating analysts' targets through tax expense decreases. Review of Accounting Studies 13, 295–318.

Graham, J., Harvey, C., Rajgopal, S., 2005. The economic implications of corporate financial reporting. Journal of Accounting and Economics 40, 3–73.

Grundfest, J., Malenko, N., 2009. Quadrophobia: strategic rounding of EPS data. Working paper.

Hayn, C., 1995. The information content of losses. Journal of Accounting and Economics 20, 125–153.

Jacob, J., Jorgensen, B., 2007. Earnings management and accounting income aggregation. Journal of Accounting and Economics 43, 369–390.

Jorgensen, B., Lee, Y., Rock, S., 2013. The shapes of scaled earnings histograms are not due to scaling and sample selection: evidence from distributions of reported earnings per share. Contemporary Accounting Research..

Leslie, D.A., 1985. Materiality: The Concept and Its Application to Auditing. The Canadian Institute of Chartered Accountants, Toronto, Canada.

Leuz, C., Nanda, D., Wysocki, P., 2003. Earnings management and investor protection: an international comparison. Journal of Financial Economics 69, 505–527.

Marquardt, C., Wiedman, C.I., 2004. How are earnings managed? An examination of specific accruals. Contemporary Accounting Research 21, 461–491.

Nelson, M.W., 1993. The effects of error frequency and accounting knowledge on error diagnosis in analytical review. The Accounting Review 68, 804–824.

Nelson, M.W., Smith, S.D., Palmrose, Z., 2005. The effect of quantitative materiality approach on auditors' adjustment decisions. The Accounting Review 80, 897–920.

Newton, L.K., 1977. The risk factor in materiality decisions. The Accounting Review 52, 97–108.

Roychowdhury, S., 2006. Earnings management through real activities manipulation. Journal of Accounting and Economics 42, 335–370.

Thomas, J., 1989. Unusual patterns in reported earnings. The Accounting Review 64, 773–787.