MANIPULATING SALES REVENUE TO USER REFERENCE POINTS IN PRE- AND POST-SARBANES OXLEY ERAS

Charles E. Jordan*
Stanley J. Clark

INTRODUCTION

Earnings management represents the most frequently researched topic within financial accounting in the last two decades. The use of earnings management and its almost cavalier acceptance by financial statement preparers and users during the 1990’s led to the significant corporate reporting debacles of the early 2000’s (e.g., Enron, WorldCom, HealthSouth, etc.) and ultimately to the Sarbanes-Oxley Act (SarbOx) in 2002. Although no universal definition exists of earnings management, in general it embodies the deliberate manipulation of income for personal gain. For example, managers manipulate income to meet analysts’ forecasts (Payne and Robb, 2000), to increase their own wealth through bonus schemes tied to earnings (Guidry et al., 1999), or to improve share-price performance that enhances their wealth through stock-based compensation plans (Brown and Higgins, 2001).

Earnings management, which is also commonly referred to as earnings manipulation, generally occurs in one of two forms. The first is known as real earnings management and results from operating manipulations. For example, to boost earnings in the current period management may shift scheduled maintenance procedures into the next period. The second is referred to as discretionary earnings management and results from accounting manipulations accomplished through the use of judgmental accruals and estimates inherent in the financial reporting process. As an example, one of the many tactics employed by HealthSouth management in the early

The authors are, respectively, Professor of Accounting at University of Southern Mississippi and Associate Professor of Accounting at Middle Tennessee State University.
2000’s to meet analysts’ forecasts of earnings was using low estimates of bad debts on credit sales. Whether earnings management results from operating manipulations or accounting manipulations, Fischer and Rosenzweig (1995) suggest that financial statement users are misled because current earnings are not indicative of long-run profitability and, thus, the “stakeholders’ trust is violated (p. 434).”

This article discusses an intriguing subset of earnings management known as cosmetic earnings management (CEM), whereby income is manipulated upward to key cognitive user reference points. The article then extends the extant research on CEM to examine the manipulation of sales revenue in a like manner. Most research on revenue manipulation analyzes it as a means of accomplishing earnings management (e.g., Plummer and Mest, 2001; Marquardt and Wiedman, 2004; Caylor, 2010; Stubben, 2010). Little research examines sales revenue itself as the end target of manipulation. In this article, the manipulation of sales revenue is analyzed for both pre- and post-SarbOx periods with results indicating that revenue was managed upward to achieve user reference points before SarbOx but not after it. The following sections of the article provide a literature review on CEM and a discussion of the research methodology, including the development of hypotheses and a description of the sample data. The research findings are then presented, and the final section contains conclusions and limitations of the research.

**LITERATURE REVIEW**

Although research results are somewhat mixed, ample evidence suggests that at least some forms of earnings management significantly declined subsequent to the implementation of SarbOx (e.g., Bedard et al., 2004; Chang and Sun, 2009; Davis et al., 2009; Kalelkar and Nwaeze, 2011). While the incidence of real earnings management increased post-SarbOx, it
appears that pre-SarbOx levels of discretionary earnings management decreased markedly (Cohen et al., 2008; Bartov and Cohen, 2009). In a post-SarbOx survey of professionals and managers, Grasso et al. (2009) demonstrate that this group judges accounting manipulations (i.e., discretionary earnings management) much more harshly than operating manipulations (i.e., real earnings management), which helps explain the results just noted by Cohen et al. and Bartov and Cohen that, post-SarbOx, levels of discretionary earnings management decreased while the frequency of real earnings management increased.

**Cosmetic Earnings Management**

CEM occurs when unmanipulated income falls slightly below a user reference point (e.g., $395 million) and management takes actions to boost earnings just enough to cross the threshold and increase the first digit by one (i.e., to barely above $400 million). Brenner and Brenner (1982) note that humans possess a limited amount of memory and tend to remember only the most relevant information about a number; the first (i.e., left-most) digit in a number is viewed as the most important with increasingly less significance afforded the second, third, and so forth digits. Furthermore, when remembering numbers, people tend to round down rather than up (Carslaw, 1988). Accordingly, investors would typically remember earnings of $395 million as three hundred something million rather than almost $400 million (Skousen et al., 2004).

When unmanipulated income falls just below a user reference point, rather than allowing investors to round this number down in their recollection of earnings, managers instead find ways to increase income to just above the breakpoint. CEM might seem like a relatively unimportant form of earnings management since it involves only marginal increases in income; however, as Thomas (1989) notes, even small changes in earnings near user reference points can cause significant changes in how a company is valued. A comparable analogy outside of
accounting would be a prospective graduate student who scores 590 on the GMAT but tells his professors he scored 600. While the manipulation of the reported score is relatively slight, it likely results in a significant and disproportionate impact on the professors’ perceptions of the student’s abilities.

Numerous studies document that CEM occurred in the U.S. and abroad from the 1980’s through the early 2000’s (e.g., Carslaw, 1988; Thomas, 1989; Niskanen and Keloharju, 2000; Van Caneghem, 2002; Kinnunen and Koskela, 2003; Skousen et al., 2004; Guan et al., 2006). Generally, these studies examine large samples of companies and find that firms with positive income possess a significantly higher rate of zeros and lower frequency of nines than expected in the second digital position of earnings. In the number 7,956, seven appears in the first digital position with nine in the second position and so on. In the above studies, typically, numbers other than nine and zero (i.e., one through eight) appear in the second earnings position at rates approximating their expected distributions. The higher rate of zeros and lower frequency of nines than expected in the second earnings position led the researchers to conclude that when unmanipulated earnings fall just below a user reference point that management bumps income up just enough to cross the breakpoint and increase the first digit by one. Van Caneghem (2002) demonstrates that short-term discretionary accruals represent the tool normally used to achieve the earnings boosts needed to accomplish CEM.

In a recent study, Jordan and Clark (2011) test for the presence of CEM in the U.S. during pre- and post-SarbOx periods. Like the previous CEM studies, they find clear evidence of this manipulative behavior in the earlier time period; however, no signs of CEM appear in the post-SarbOx era. In the Cohen et al. (2008) and Bartov and Cohen (2009) studies discussed previously, earnings management effected through discretionary accruals decreased significantly
after SarbOx. This lends credibility to the findings of Jordan and Clark. That is, since CEM is accomplished primarily through current discretionary accruals, it comes as no surprise that this form of earnings management declined dramatically after SarbOx.

RESEARCH DESIGN

Development of Research Questions

As noted in the introduction, little research examines sales revenue as the end object of manipulation. However, several studies highlight the importance of revenue as an individual component of accrual earnings. For example, growth rates in revenue provide an important measure of firm quality and value (e.g., Dechow and Dichiv, 2002; Zheng and Strangeland, 2007). Changes in revenue actually outperform both changes in earnings and changes in operating cash flows in terms of share price prediction (Jordan et al., 2007). Ertimur et al. (2003) found that investors react more strongly to unexpected revenue results than they do to unexpected cost-savings or expense amounts. For companies that traditionally report low or negative earnings (e.g., internet firms), evidence indicates that market values are more closely associated with revenue levels than with earnings (Bowen et al., 2002; Davis, 2002). Cullen et al. (2008) show that the greater a company’s string of past and expected future net losses or past and expected future negative operating cash flows, the greater the likelihood the firm will manipulate (i.e., overstate) revenues to maintain or enhance market valuation.

Given the importance of sales revenue as a stand-alone performance measure and the many ways it can be manipulated (e.g., see Stallworth and DiGregorio, 2004), a question arises concerning whether sales revenue is manipulated upward to achieve user reference points in a manner similar to the cosmetic earnings management described earlier. Only one study (Jordan et al., 2009) examines this issue and does so for a very limited sample (i.e., the Fortune 1,000
companies for the year 2006). Although their results do not follow the true patterns demonstrated in CEM studies (i.e., significantly more zeros and fewer nines than expected in the second sales revenue position), the authors find some evidence of manipulation to achieve user reference points in revenue. Their limited sample and rather inconclusive findings, though, leave the issue unresolved and lead to the two research questions or hypotheses for the current study. The first question concerns whether cosmetic sales management (CSM) is practiced within a broad cross-section of U.S. companies, and thus the first hypothesis is stated as follows (in the null form):

\[ H_1: \text{The occurrence of the numbers zero through nine will appear in the second revenue position at rates approximating their expected distributions and, therefore, no evidence will exist of revenue manipulation to achieve user reference points.} \]

The alternative hypothesis supports the existence of CSM as numbers in the second revenue position will not conform to their anticipated frequencies. Instead, there will be an unusually high distribution of zeros and an abnormally low frequency of nines in the second revenue position.

The second question deals with whether SarbOx and the increased scrutiny of financial reporting following the high profile scandals around the turn of the millennium produced any apparent effect on the incidence of CSM. Thus, the second hypothesis is stated as follows (in the null form):

\[ H_2: \text{There will be no difference between pre- and post-SarbOx periods in the incidence of manipulating revenue to achieve user reference points.} \]

The alternative hypothesis supports the notion that the incidence of CSM will be less in the post-SarbOx period than in the pre-SarbOx period. This would be expected not only because of the harsh penalties for fraudulent financial reporting under SarbOx but also because of the
heightened attention associated with dubious forms of financial reporting following the major financial scandals in the early 2000’s.

Notice that CSM and CEM are two separate phenomena. Ample previous research examines and documents the incidence of CEM in practice, both before and after SarbOx, while little work addresses the occurrence of CSM in either period. The current study attempts to fill this void in the literature and, accordingly, focuses solely on CSM.

**Methodology and Sample Data**

The methodology used in this study to detect CSM echoes the techniques employed in the CEM studies discussed earlier. The second digital position within a firm’s net sales revenue figure represents the key indicator of CSM. If companies whose unmanipulated revenue falls only slightly below a user reference point are managed upward just enough to cross the threshold and increase the first digit by one, then when examining a large sample of firms there should exist a higher than expected frequency of zeros and lower than anticipated rate of nines in the second revenue position. Digits other than zero and nine should occur in the second revenue position at rates approximating their expected distributions. The absence of CSM would be suggested if all digits, zero through nine, appear in the second revenue position at rates approximating their anticipated frequencies.

Obviously, a crucial aspect of testing for CSM is ensuring that the expected frequencies of digits (zero through nine) in the second revenue position are appropriate. Benford (1938) demonstrates that numbers do not appear in the first two digital positions of naturally occurring data in equal distributions. More specifically, the numbers one through nine do not have a one-ninth chance of appearing in the first digital position. Likewise, the numbers zero through nine do not occur proportionately in the second digital position. Instead, small numbers appear in the
first two digital positions of naturally occurring data at disproportionately higher rates than do large numbers. Now known simply as Benford’s Law, Table 1 provides expected frequencies for numbers appearing in the first two positions of naturally occurring data.

**Table 1**
Benford’s expected digital frequencies.

<table>
<thead>
<tr>
<th>Digit</th>
<th>Position of digit in number</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First</td>
</tr>
<tr>
<td>0</td>
<td>11.97%</td>
</tr>
<tr>
<td>1</td>
<td>30.10%</td>
</tr>
<tr>
<td>2</td>
<td>17.61%</td>
</tr>
<tr>
<td>3</td>
<td>12.49%</td>
</tr>
<tr>
<td>4</td>
<td>9.69%</td>
</tr>
<tr>
<td>5</td>
<td>7.92%</td>
</tr>
<tr>
<td>6</td>
<td>6.70%</td>
</tr>
<tr>
<td>7</td>
<td>5.80%</td>
</tr>
<tr>
<td>8</td>
<td>5.12%</td>
</tr>
<tr>
<td>9</td>
<td>4.58%</td>
</tr>
</tbody>
</table>


For example, Benford’s Law suggests that ones occur in the first digital position about 30.10 percent of the time while nines possess only a 4.58 percent probability of appearing in this position. Similarly, zeros appear in the second digital position around 11.97 percent of the time while nines occur in this position only 8.50 percent of the time. In positions right of the second digit, the numbers zero through nine appear at approximately proportional rates. A simple example demonstrates the logic behind Benford’s Law. Assume a company has $10 million in assets and the assets grow at the rate of 5 percent per year. At this compounded growth rate, it will take about 14 years for total assets to reach $20 million. However, it will take only eight more years before the first digit increases again (i.e., to $30 million) and will take only six years after that for total assets to reach $40 million. Thus, with each unit increase in the first digit, it takes a shorter period of time before the first digit increases again, until the first digit reaches one
again (e.g., $100 million), at which time the process begins anew. Thus, at any given time, for a randomly selected sample of companies taken from a population of firms with varying sizes, more of the entities would have total assets with low numbers as the first digit than would have high numbers in this digital position. Prior research demonstrates that unmanipulated financial statement data conform very well with Benford’s expected frequencies while manipulated financial reporting items do not (e.g., Nigrini, 1996; Nigrini and Mittermaier, 1997). Benford’s expected digital distributions have been used extensively in detecting the presence of CEM and, accordingly, represent the appropriate frequencies in testing for CSM as well.

To ascertain whether CSM occurs and if the passage of SarbOx and the increased scrutiny of financial reporting following the early 2000’s affected the incidence of CSM, data are collected for two time periods (i.e., 1992-1999 and 2003-2010). Data are not analyzed for the period 2000-2002 because it was during this time that the high profile scandals became publicly known and SarbOx was developed. Results for these years would likely have been convoluted. The two eight-year time periods chosen represent clear pre- and post-SarbOx eras. The samples consist of all firm years in the North America – Annual Compustat database excluding Canadian firms, inactive firms, and firms with negative or zero net sales revenue. The 1992-1999 time period had a total of 96,514 firm years. Firms were deleted from this total for the three reasons stated: Canadian firms - 8,145; inactive firms - 56,373; negative or zero net sales revenue - 2,062. The 2003-2010 time period had a total of 86,239 firm years. Firms were deleted from this total for the three reasons stated: Canadian firms - 12,801; inactive firms - 12,181; negative or zero net sales revenue - 17,503. The final pre- and post-SarbOx samples comprise 29,934 and 43,754 firm years, respectively. The pre-SarbOx sample is smaller than the post-SarbOx sample primarily because, relative to the later period, a larger number of firms that existed during the
earlier period are no longer active. Still, though, both samples are extremely large and representative of the companies existing during the pre- and post-SarbOx periods.

RESULTS

Table 2 presents results for the 1992-1999 time period. The table shows the number of times and rate at which each of the digits zero through nine appear in the second revenue position. For example, zeros occur in the second revenue position 3,816 times, which represents an observed frequency of 12.75 percent. Table 2 also shows Benford’s expected frequencies at which each of the digits zero through nine should appear in the second position of unmanipulated data. For example, zeros should occur in the second position 11.97 percent of the time. Finally, for each of the digits zero through nine, the table provides the z value and p-level for a one-sample proportions test comparing the observed frequency of each digit with its expected rate of occurrence. As an example, the proportions test comparing the observed and expected frequencies of zeros appearing in the second revenue position yields a z value and p-level of 4.15 and .000, respectively.

Table 2

<table>
<thead>
<tr>
<th>Second revenue digit</th>
<th>Observed count (n)</th>
<th>Observed frequency (%)</th>
<th>Expected frequency (%)</th>
<th>z value</th>
<th>p-level</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>3816</td>
<td>12.75</td>
<td>11.97</td>
<td>4.15</td>
<td>.000**</td>
</tr>
<tr>
<td>1</td>
<td>3298</td>
<td>11.02</td>
<td>11.39</td>
<td>-2.03</td>
<td>.043*</td>
</tr>
<tr>
<td>2</td>
<td>3260</td>
<td>10.89</td>
<td>10.80</td>
<td>.059</td>
<td>.953</td>
</tr>
<tr>
<td>3</td>
<td>3119</td>
<td>10.42</td>
<td>10.83</td>
<td>.339</td>
<td>.953</td>
</tr>
<tr>
<td>4</td>
<td>3020</td>
<td>10.09</td>
<td>10.03</td>
<td>.535</td>
<td>.735</td>
</tr>
<tr>
<td>5</td>
<td>2922</td>
<td>9.76</td>
<td>9.67</td>
<td>-.195</td>
<td>.845</td>
</tr>
<tr>
<td>6</td>
<td>2786</td>
<td>9.31</td>
<td>9.34</td>
<td>.241</td>
<td>.809</td>
</tr>
<tr>
<td>7</td>
<td>2718</td>
<td>9.08</td>
<td>9.04</td>
<td>-2.68</td>
<td>.388</td>
</tr>
<tr>
<td>8</td>
<td>2580</td>
<td>8.62</td>
<td>8.76</td>
<td>-2.68</td>
<td>.735</td>
</tr>
<tr>
<td>9</td>
<td>2415</td>
<td>8.07</td>
<td>8.50</td>
<td>-2.68</td>
<td>.007**</td>
</tr>
</tbody>
</table>

**Difference between the observed and expected frequencies is significant at the .01 level.
* Difference between the observed and expected frequencies is significant at the .05 level.

An alpha level of .01 is used for statistical testing rather than a less strenuous level (e.g., .05 or .10) to increase the likelihood that any statistically significant differences noted between the observed and expected frequencies result from manipulative behavior rather than random
chance. For example, if testing at a .10 alpha level it would be expected that at least one of the
ten digits would possess an observed frequency that differs significantly from its expected
distribution simply due to random occurrence. Similarly, testing at a .05 alpha level, while more
rigorous, still results in a high likelihood (i.e., a 50 percent probability) that one of the ten digits
would differ significantly from its expectation merely due to random chance. Testing at the .01
alpha level gives very strong evidence that any significant difference between a digit's observed
and expected frequencies was caused by manipulative behavior. The results in Table 2 reveal a
clear pattern of CSM for the pre-SarbOx period. In particular, nines and zeros occur in the
second revenue position at rates significantly different from their expected frequencies. Nines
appear much less often than expected (i.e., p-level = .007), while zeros occur much more
frequently than anticipated (i.e., p-level = .000). All other digits (i.e., one through eight) appear
in the second revenue position at rates approximating their expected distributions. For the pre-
SarbOx period, this provides compelling evidence that when unmanipulated revenue fell just
below a user reference point, management manipulated revenue upward just enough to cross the
threshold and increase the first digit by one.

Table 3 provides results for the 2003-2010 period in the same manner as presented for the
earlier period in Table 2. However, the findings in Table 3 for the post-SarbOx period stand in
stark contrast with those from the pre-SarbOx era. More specifically, again testing at a .01 alpha
level, every digit (i.e., zero through nine) for the post-SarbOx sample occurs in the second
revenue position at a rate approximating its expected frequency. The absence of any statistically
significant digital anomalies in the second revenue position suggests that CSM does not exist in
the post-SarbOx era.
Table 3
Digital frequencies for second revenue position for 2003-2010 (post-SarbOx) sample.

<table>
<thead>
<tr>
<th>Second revenue digit</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed count (n)</td>
<td>5359</td>
<td>5088</td>
<td>4756</td>
<td>4463</td>
<td>4339</td>
<td>4233</td>
<td>4066</td>
<td>3838</td>
<td>3835</td>
<td>3777</td>
</tr>
<tr>
<td>Observed frequency (%)</td>
<td>12.25</td>
<td>11.63</td>
<td>10.87</td>
<td>10.20</td>
<td>9.92</td>
<td>9.67</td>
<td>9.29</td>
<td>8.77</td>
<td>8.76</td>
<td>8.63</td>
</tr>
<tr>
<td>Expected frequency (%)</td>
<td>11.97</td>
<td>11.39</td>
<td>10.88</td>
<td>10.43</td>
<td>10.03</td>
<td>9.67</td>
<td>9.34</td>
<td>9.04</td>
<td>8.76</td>
<td>8.50</td>
</tr>
<tr>
<td>z value</td>
<td>1.79</td>
<td>1.57</td>
<td>-0.68</td>
<td>-1.57</td>
<td>-0.788</td>
<td>0.022</td>
<td>-0.339</td>
<td>-1.96</td>
<td>0.036</td>
<td>0.993</td>
</tr>
<tr>
<td>p-level</td>
<td>0.073</td>
<td>0.116</td>
<td>0.956</td>
<td>0.116</td>
<td>0.431</td>
<td>0.974</td>
<td>0.735</td>
<td>0.0504</td>
<td>0.971</td>
<td>0.321</td>
</tr>
</tbody>
</table>

Cullen et al. (2008) note that companies reporting negative earnings are more likely to manipulate revenue to maintain market value than are companies with positive earnings. Thus, the results in Tables 2 and 3 could be explained if a larger portion of the pre-SarbOx companies had negative earnings relative to the post-SarbOx sample. However, this is not the case. A much lower portion (i.e., 26.3 percent) of the pre-SarbOx firms reported net losses relative to the post-SarbOx sample (with 34.6 percent of the companies presenting negative earnings). In general, companies in the pre-SarbOx period were substantially more profitable than those in the post-SarbOx era in that the median return on investment (ROI) for the earlier time period of 3.03 percent is approximately double the median ROI (i.e., 1.52 percent) for the later period. With companies struggling to generate profits in the post-SarbOx period, it would appear management had more incentive to window dress the financial statements during this era than they did in the earlier period. Yet, the findings in this study paint a markedly different picture showing that CSM seems to have vanished during a period when, otherwise, one might have expected it to flourish. A logical explanation for the apparent disappearance of CSM in recent years is, of course, the implementation of SarbOx and the heightened awareness of financial statement manipulation occurring after the high profile reporting debacles of the early 2000’s.
In summary, the first hypothesis that the numbers in the second revenue position will appear at their expected frequencies and no revenue manipulation occurs must be rejected for the pre-SarbOx period. Instead, clear evidence supports the alternative hypothesis as this sample has an unusually high frequency of zeros and an abnormally low distribution of nines in the second revenue position. This suggests that when unmanipulated revenue fell just below a user reference point, management took action to increase revenue just enough to cross the threshold and increase the first revenue digit by one. The findings also support rejection of the second hypothesis (i.e., that there will be no difference between pre- and post-SarbOx samples in the incidence of CSM). The results indicate the alternative hypothesis be accepted as evidence exists of manipulation to achieve user reference points in the pre-SarbOx period but not in the post-SarbOx era.

**CONCLUSION AND LIMITATIONS**

This study provides strong evidence that CSM occurred prior to SarbOx but vanished in the post-SarbOx period. Along with the studies showing an overall decline in earnings management after SarbOx, the current findings suggest that managers and accountants have adopted a much more altruistic view of financial reporting than existed in the pre-SarbOx era. What is not clear, however, is exactly what caused this trend toward unbiased reporting.

Davis et al. (2009) suggest that the increased scrutiny and threat of legal sanctions resulting from SarbOx likely led to a reduction of financial statement manipulation by managers and to auditors’ decreased tolerance of it. Grasso et al. (2009), though, speculate that the high profile reporting scandals may have done more to change attitudes about the ethicality of biased financial reporting than did SarbOx. Another plausible explanation exists in the form of Staff Accounting Bulletin (SAB) 101, which was issued by the SEC in 1999 and provided additional
guidance on applying GAAP in relation to revenue recognition. Its purpose was to tighten, at least somewhat, the rather vague revenue recognition criteria used in practice, thus presumably making it more difficult to manipulate both revenue and earnings. In all likelihood, a combination of the above factors created the changed attitudes toward manipulating revenue and earnings and resulted in the unbiased financial reporting currently being documented in this and other studies.

The apparent disappearance of CSM in the post-SarbOx period for the general sample examined in this study does not necessarily imply that it no longer occurs within specific subgroups (e.g., particular industries, companies with a history of negative earnings or operating cash flows, clients of specific audit firms, etc.). Further research would be needed to determine this. In addition, it is possible the elimination of CSM noted here is unrelated to the events occurring at or shortly after the turn of the millennium (i.e., high profile reporting scandals, SarbOx, or SAB 101) but instead is due to some unknown phenomenon. The inability to determine a specific cause for the more transparent financial reporting represents a limitation of this study.
References


