Does Earnings Management Matter for Strategy Research?*

R. Anthony Gibbs\textsuperscript{1}, Timothy S. Simcoe\textsuperscript{2}, and David M. Waguespack\textsuperscript{1}

\textsuperscript{1}Robert H. Smith School of Business, University of Maryland
\textsuperscript{2}Boston University Questrom School of Business and NBER

May 4, 2021

Abstract

Strategic management research frequently seeks to explain variation in organizational performance using metrics such as accounting profits scaled by firm assets (ROA). A concern with accrual-based accounting methods, perhaps best illustrated by a large discontinuity in the distribution of ROA around zero for U.S. public firms, is that operational and accounting practices will artificially inflate/deflate accounting profit. In this manuscript we establish that such \textit{earnings management} is common, introduces non-classical noise, and distorts our understanding of broad drivers of firm performance. We conclude with analysis showing that an alternative performance measure, Cash Flows from Operations on Assets (OCFOA), offers a robust vehicle for checking results using accounting profits.

Keywords: earnings management, return on assets, performance, measurement error, bunching

\textsuperscript{*}Contact: rgibbs12@umd.edu; tsimcoe@bu.edu; dwaguesp@rhsmith.umd.edu. For helpful comments, we thank J.C. Suárez Serrato and seminar participants at the University of Maryland, University of Michigan Ross School of Business, and BU Questrom School of Business.
1 Introduction

Investors, managers, and scholars all rely on accounting-based measures of public firm performance. A cursory search on Google Scholar, for example, yields over 22,000 papers, distributed over multiple fields, containing the terms “Compustat” and “Return on Assets.” This prevalence partly reflects a long tradition of using accounting data to study both the drivers of profitability (McGahan and Porter, 1997, 2002), and the persistence of performance (Rumelt et al., 1991; D’Aveni et al., 2010). Moreover, reporting of accounting measures is mandated for publicly traded firms, providing scholars with metrics that are comparable, convenient, and broadly accepted as important.

At the same time, there is a substantial accounting literature on earnings management, defined as reporting that aims to “mislead some stakeholders about the underlying economic performance of the company or influence contractual outcomes that depend on reported accounting numbers” (Healy and Wahlen, 1999). Although our impression is that scholars studying firm performance are generally aware that accounting adjustments can obscure the link between real and reported performance, we find very few citations to the relevant accounting research in fields such as strategy, finance, and economics. We speculate that this omission reflects the fact that manipulation, an activity that is by definition hidden, is hard to systematically assess. Moreover, scholars may implicitly assume that any “noise” in accrual accounting will balance out within the firm over time, and that the market will detect non-trivial misreporting.

In this manuscript we first establish that earnings management is common, introduces non-classical noise, and distorts our understanding of broad drivers of firm performance. We begin with a simple model of incentives to manipulate earnings that predicts bunching in reported earnings just above the zero returns threshold. We then turn to data from Compustat and Execucomp, and document a discontinuity in the distribution of Return on Assets (ROA) at zero profits, and employ a bunching estimator (Chetty, 2012; Kleven and Waseem, 2013) to estimate that approximately 15 percent of firm-year observations are shifted from negative to positive profitability. While striking, this shift in the distribution of ROA could reflect endogenous effort (i.e., striving harder when within striking distance of a goal) as well as accounting tricks. Therefore, we next demonstrate that for Cash Flows...

---

1We use the term “earnings manipulation” to describe discretionary reporting decisions permitted under Generally Accepted Accounting Principles (GAAP) that strategically inflate or deflate accounting profits. The term does not imply fraud.

2For example, at the time of this writing, the seminal study of earnings management by Burgstahler and Dichev (1997) had been cited over 5,000 times. It received one citation in Strategic Management Journal, three in the Academy of Management Journal, five in the Journal of Finance, and none in the Quarterly Journal of Economics.
from Operations on Assets (OCFOA), an alternative accounting measure that is arguably less subject to manipulation, only approximately four percent of observations shift from the negative to positive region. Finally, we conduct a decomposition of variance, in the spirit of canonical analyses found in Schmalensee (1985), Rumelt et al. (1991), and McGahan and Porter (1997, 2002), comparing results based on ROA versus OCFOA. We find that earnings management may obfuscate 10 percent or more of the variance in earnings that scholars can predict using these factors; and moreover, the manipulation changes the relative importance of industry-, firm-, and CEO-level factors.

The issues of performance measurement and the match to theoretical constructs are of long-standing concern in management scholarship (Winter, 1995; Lieberman, 2021). An emerging stream of recent work has addressed deficiencies in accounting measures, such as the distinction between average and marginal profit maximization (Levinthal and Wu, 2010; Shapira and Shaver, 2014), short- versus long-term value creation (Wibbens and Siggelkow, 2020), cleavages between value creation and capture (Lieberman et al., 2017), and the ability of firms to leverage non-owned assets (Barney, 2019). We contribute to this line of work by documenting how accrual accounting may systematically obscure understanding of the relationship between firm policies and outcomes, and by offering the relatively simple solution of checking results with an alternative accounting measure. While we embed our variance decomposition analysis within a broad line of empirical inquiry, we believe there is much potential for strategy scholars to examine whether and how more specific firm actions are influenced by earnings manipulation.

2 Theory: Earnings Manipulation and Bunching

This section presents a simplified model of earning management, based on the more general treatment in Kleven (2016). Our model includes a single firm whose true performance is a random variable denoted by \( \pi \). The CEO observes her firm’s performance and makes a report \( R = \pi + a \), where \( a \) represents accounting adjustments. In our empirical context, \( R \) corresponds to publicly reported accounting-based performance measures.

We assume that adjustments incur a quadratic cost \( c(a) = \gamma a^2/2 \), so unbiased reporting is free, and reporting costs increase (at an increasing rate) with the size of any adjustments. In practice, the costs of earnings management may show up in a wide variety of ways, such as a loss in credibility, managerial distraction, the direct costs of an audit, increased financial constraint, or the cost of “unwinding” an adjustment by under-stating future profits. By adopting a reduced-form quadratic cost function, we are emphasizing expositional clarity and convenience over realism.
The CEO chooses adjustments, $a$, to maximize her payoff, which takes the following form:

$$\max_a U(a; \pi) = (\pi + a) + B \cdot 1_{\{R \geq 0\}} - c(a)$$  \hspace{1cm} (1)$$

where $1_{\{R \geq 0\}}$ is an indicator function that equals one if and only if the report, $R$, is non-negative. The CEO’s payoff increases linearly with $R$, to capture the idea that she would generally like to report better performance. Because she also pays a quadratic adjustment cost, $c(a)$, however, there is a limit to the size of any distortions. The parameter $B$ is a “bonus” paid to the CEO for a non-negative report. This bonus could represent an actual payout, a reduced probability of termination, or simply a psychological benefit associated with “not losing money.” Regardless of the underlying cause, the bonus produces a discontinuous jump in the marginal benefits of earnings management when $R = 0$. This jump is called a “notch” in the public finance literature.

As a baseline model of earnings manipulation, consider the CEO’s report in the absence of a notch (i.e., when $B = 0$). Given the linear quadratic structure of equation (1), the CEO’s first-order condition reveals that $a^* = \frac{1}{\gamma}$. The CEO will always make optimistic reports, and the size of her adjustments will naturally decline as the cost of mis-reporting, $\gamma$, grows larger.

Before considering how a notch affects the CEO’s report, it is useful to pause and consider the implications of this baseline model for empirical strategy research. Because the CEO always makes adjustments, a researcher never actually observes “true” performance. On the other hand, this may not matter very much. In particular, variation in underlying performance, $\pi$, maps directly into variation in the optimal report, $R = \pi + \frac{1}{\gamma}$. For example, in a statistical analysis that seeks to explain how some factor or decision $X$ impacts observed performance $R$, all reporting distortions can be swept away simply by including a constant term in the regression. Unfortunately, this argument only goes so far. In our model, $a^*$ is constant only because the marginal costs and benefits of adjustments are uncorrelated with $\pi$. In general, as we now illustrate for the case of a notch, mis-reporting might be correlated with $\pi$, $X$ or both, leading to well-known problems of omitted variables or simultaneity.

To see how this can happen, consider our baseline model with a notch induced by $B > 0$. The CEO now has an incentive to “reach” for the bonus by reporting $R = 0$ (or equivalently, $a = -\pi$), as long as the firm’s true performance is close enough to the reporting threshold. In the Appendix, we show that this happens when $\pi > \pi_L \equiv -\left(\frac{1}{\gamma} + \sqrt{\frac{2B}{\gamma}}\right)$. This implies that
the CEO’s optimal reporting strategy is:

\[ R^*(\pi) = \begin{cases} 
\pi + \frac{1}{\gamma} & \text{if } \pi \notin [\pi_L, -\frac{1}{\gamma}] \\
0 & \text{if } \pi \in [\pi_L, -\frac{1}{\gamma}] 
\end{cases} \]

Figure 1 graphs this optimal reporting strategy, and illustrates the distribution of \( R \) when true performance is normally distributed. As illustrated in right panel, there is a “hole” in the distribution of reports just below \( R = 0 \), and a spike or mass-point at zero, because all of the firms with true performance in the interval \([\pi_L, -\frac{1}{\gamma}]\) shift their reports upwards to zero. This is the key feature of the model that we will examine in our data.

Figure 1: Optimal Reporting (left) and the Distribution of Reports (right)

Although the predictions of this simple model are very stark, they can be relaxed. For example, if we allow the marginal cost of adjustment, \( \gamma \), to vary across firms or introduce an idiosyncratic fixed cost of earnings manipulation, then some CEO’s may choose to make slightly negative reports. We do not pursue those extensions here because the purpose of the simple model is not to capture every feature of the data set described below. Rather, our aim is to illustrate a set of incentives that can generate bunching in reported profits. We then use the actual bunching observed in our data to illustrate how earnings management can distort empirical strategy research.

A final point about the model that merits some discussion is the interpretation of the CEO’s choice. Up to this point, we have labeled the variable \( a \) “adjustments” and assumed that it represents earnings manipulation. Although we find that interpretation plausible, one could easily re-label \( a \) “managerial effort” and argue that a better interpretation of any observed bunching is a try-harder effect induced by the same notch in the CEO’s payoff function. To address that concern, we introduce a second performance measure that is harder
to manipulate, and show that there is a systematic difference in the amount of bunching across these two outcomes. Because that approach is fundamentally empirical, we now turn to a description of the data.

3 Context: Public Firm Performance Data

Return on Assets (ROA) features prominently in strategy research on the drivers of organizational performance. For example, out of approximately 860 empirical articles published in the *Strategic Management Journal* between 2011 and 2020, we found that 238 articles (27%) reference ROA. The popularity of ROA as an outcome variable in the empirical strategy literature is due to at least three factors. First, ROA is comparable across firms of different sizes, and in theory represents the capability of managers to generate value from a stock resources (Barney, 1991, for example). Second, ROA is a key outcome variable used by investors, making it reasonable to assume that managers also focus on that outcome. Third, and perhaps most importantly, the underlying components of ROA — Net or Operating Income and Total Assets — are part of the mandated reporting requirements for publicly traded U.S. firms, and are therefore readily available to scholars through the Compustat database.

Because our aim in this paper is to illustrate the potential importance of earnings management for Strategy research that takes ROA as an outcome, we also use Compustat data. Table 1 below reports descriptive statistics, and Table 2 correlations, for selected variables from the Compustat database using data from 1992-2018. Each table considers two samples. The first sample comprises all firms publicly traded in the United States (N=210,797). The second sample (N=171,328) excludes firms in the financial sector (standard industrial classification [SIC] codes in the 6000s) or public administration (SIC codes in the 9000s) as is common in many academic studies that utilize ROA. Both samples are unbalanced panels, with firms entering in 1992 or the year they became public, and exiting in 2018 or the year they ceased being public.

Most of the variables used in our analysis are quite standard. Net Income, Total Assets, and OCF (Net cash flows from operating activities) are incorporated into Compustat from the firm’s annual 10-K filings with the SEC. OCF “...is the cash profit the company would have reported had it constructed its income statement on a cash basis rather than an accrual basis” (Easton et al. 2013, p. 2-17).

OCF plays an important role in our analysis, and it can be calculated in two ways: the direct method (i.e., noting the cash received or cash paid for all operating transactions), and
Table 1: Summary Statistics

<table>
<thead>
<tr>
<th>Variables</th>
<th>Observations</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fiscal Year</td>
<td>210,797</td>
<td>2005</td>
<td>7.670</td>
<td>1992</td>
<td>2019</td>
</tr>
<tr>
<td>Total Assets</td>
<td>210,797</td>
<td>9,127.191</td>
<td>82,896.37</td>
<td>0</td>
<td>3,771,200</td>
</tr>
<tr>
<td>Net Income</td>
<td>210,797</td>
<td>161.497</td>
<td>1,306.279</td>
<td>-99,289</td>
<td>99,806.04</td>
</tr>
<tr>
<td>ROA</td>
<td>210,797</td>
<td>-0.034</td>
<td>0.224</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>OCF</td>
<td>210,797</td>
<td>369.056</td>
<td>2,570.763</td>
<td>-110,560</td>
<td>166,671.5</td>
</tr>
<tr>
<td>OCFOA</td>
<td>210,797</td>
<td>0.031</td>
<td>0.193</td>
<td>-1</td>
<td>0.999</td>
</tr>
<tr>
<td>Earnings Smoothing</td>
<td>145,479</td>
<td>2.692</td>
<td>6.311</td>
<td>0</td>
<td>520.836</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variables</th>
<th>Observations</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fiscal Year</td>
<td>171,328</td>
<td>2005</td>
<td>7.670</td>
<td>1992</td>
<td>2019</td>
</tr>
<tr>
<td>Total Assets</td>
<td>171,328</td>
<td>3,688.508</td>
<td>16,842.77</td>
<td>0</td>
<td>551,669</td>
</tr>
<tr>
<td>Net Income</td>
<td>171,328</td>
<td>145.215</td>
<td>1,211.2</td>
<td>-98,696</td>
<td>98,806.04</td>
</tr>
<tr>
<td>ROA</td>
<td>171,328</td>
<td>-0.042</td>
<td>0.236</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>OCF</td>
<td>171,328</td>
<td>352.571</td>
<td>1,853.95</td>
<td>-16,856</td>
<td>81,266</td>
</tr>
<tr>
<td>OCFOA</td>
<td>171,328</td>
<td>0.032</td>
<td>0.202</td>
<td>-1</td>
<td>0.999</td>
</tr>
<tr>
<td>Earnings Smoothing</td>
<td>117,585</td>
<td>2.109</td>
<td>2.243</td>
<td>0.090</td>
<td>20.544</td>
</tr>
</tbody>
</table>

SICs in 6000s and 9000s Omitted

<table>
<thead>
<tr>
<th>Variables</th>
<th>Observations</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fiscal Year</td>
<td>171,328</td>
<td>2005</td>
<td>7.670</td>
<td>1992</td>
<td>2019</td>
</tr>
<tr>
<td>Total Assets</td>
<td>171,328</td>
<td>3,688.508</td>
<td>16,842.77</td>
<td>0</td>
<td>551,669</td>
</tr>
<tr>
<td>Net Income</td>
<td>171,328</td>
<td>145.215</td>
<td>1,211.2</td>
<td>-98,696</td>
<td>98,806.04</td>
</tr>
<tr>
<td>ROA</td>
<td>171,328</td>
<td>-0.042</td>
<td>0.236</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>OCF</td>
<td>171,328</td>
<td>352.571</td>
<td>1,853.95</td>
<td>-16,856</td>
<td>81,266</td>
</tr>
<tr>
<td>OCFOA</td>
<td>171,328</td>
<td>0.032</td>
<td>0.202</td>
<td>-1</td>
<td>0.999</td>
</tr>
<tr>
<td>Earnings Smoothing</td>
<td>117,585</td>
<td>2.109</td>
<td>2.243</td>
<td>0.090</td>
<td>20.544</td>
</tr>
</tbody>
</table>

the indirect method of starting from Net Income and removing all non-cash gains or losses.\[^3^]

At a conceptual level, Net Income – the numerator of ROA – represents the profit or loss of a business using accrual-based accounting, while OCF represents the profit or loss from operations using a cash basis.\[^4^\] Specifically, using OCF as a measure of firm performance rather than income-based measures removes the effect of investing and financing effects, the effects of interest, taxes, and special items, and the effects of non-cash book transactions such as depreciation, amortization, or book-value changes in asset or liability valuation. In addition to these specific items that would appear as journal entries in the corporate accounts, OCF is also not sensitive to broad accounting policy decisions such as the choice of inventory valuation method (e.g., LIFO vs. FIFO), when revenue is recognized, or allowances for potential outcomes (such as anticipated customer returns). Because it is less sensitive to various discretionary choices that managers can use to influence reported profit, OCF should

\[^3^\]Specifically, the items that are removed are typically depreciation/amortization, changes to current non-cash assets (such as accounts receivable, inventory), and changes to current non-cash liabilities (such as accounts payable).

\[^4^\]Some scholars use Operating Income or adjusted income such as Earnings Before Interest, Taxes, Depreciation, or Amortization (EBITDA) to calculate ROA. These other income-based measures relieve some of the potential error from earnings management, as they strip out certain sources of accounting-based discretion, but OCF excludes more potential sources for accounting-based manipulation by restricting fully to a cash basis.
be less vulnerable to accounting-based earnings management than ROA.

Table 2: Cross-Correlation Table

<table>
<thead>
<tr>
<th>Variables</th>
<th>Fiscal Year</th>
<th>Total Assets</th>
<th>Net Income</th>
<th>ROA</th>
<th>OCF</th>
<th>OCFOA</th>
<th>Earnings Smoothing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fiscal Year</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Assets</td>
<td>0.068</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Net Income</td>
<td>0.071</td>
<td>0.357</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ROA</td>
<td>-0.046</td>
<td>0.027</td>
<td>0.106</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OCF</td>
<td>0.075</td>
<td>0.379</td>
<td>0.546</td>
<td>0.062</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OCFOA</td>
<td>-0.036</td>
<td>0.007</td>
<td>0.072</td>
<td>0.697</td>
<td>0.079</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>Earnings Smoothing</td>
<td>-0.021</td>
<td>0.109</td>
<td>0.046</td>
<td>0.101</td>
<td>0.010</td>
<td>0.043</td>
<td>1.000</td>
</tr>
</tbody>
</table>

SICs in 6000s and 9000s Omitted

<table>
<thead>
<tr>
<th>Variables</th>
<th>Fiscal Year</th>
<th>Total Assets</th>
<th>Net Income</th>
<th>ROA</th>
<th>OCF</th>
<th>OCFOA</th>
<th>Earnings Smoothing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fiscal Year</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Assets</td>
<td>0.115</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Net Income</td>
<td>0.071</td>
<td>0.605</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ROA</td>
<td>-0.068</td>
<td>0.082</td>
<td>0.123</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OCF</td>
<td>0.100</td>
<td>0.875</td>
<td>0.748</td>
<td>0.095</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OCFOA</td>
<td>-0.040</td>
<td>0.080</td>
<td>0.090</td>
<td>0.718</td>
<td>0.112</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>Earnings Smoothing</td>
<td>-0.097</td>
<td>-0.002</td>
<td>0.021</td>
<td>0.246</td>
<td>-0.002</td>
<td>0.174</td>
<td>1.000</td>
</tr>
</tbody>
</table>

ROA in a given year for a given firm is calculated by the authors, following convention, by dividing Net Income by the Total Assets from the prior year. Similarly, OCFOA is calculated by dividing OCF by the Total Assets of the prior year. By construction, OCF and Net Income are strongly correlated, as are the two performance measures ROA and OCFOA. Although OCFOA is not widely used as a performance measure in the strategy literature, we found only six instances in our corpus of SMJ articles, it is clearly linked to operational performance, and for the reasons described above, less subject to accounting manipulation than ROA.

Earnings Smoothing, the final variable listed in Tables 1 and 2, is well known to accounting scholars (Leuz et al., 2003; Dechow et al., 2010) but less common in strategy research.

5There is evidence that firms also use methods in addition to accruals to engage in earnings management (Zang, 2012; Roychowdhury, 2006; Graham et al., 2005). Mismeasurement caused by these other types of activities may not be detected by our analysis. Thus, our estimates are likely to represent a conservative lower bound on the potential impact earnings-management-induced measurement error have on ROA.

6Total Assets from the prior year is used in order to avoid time reversal, for instance such that declines in Net Income or OCF early in the year prompt asset depreciation later in the year.

7See Figure C.1
It is defined as standard deviation of OCFOA divided by the standard deviation of ROA, calculated over the trailing 12 quarters (and therefore computed from quarterly rather than annual data). Earnings Smoothing is constructed such that a higher ratio indicates smoother earnings relative to the underlying cash flows. Many managers prefer smooth earnings paths (Graham et al., 2005), and the intuition behind this variable is that a large discrepancy between variation in operating cash flows and variation in accounting earnings may signal that a firm is intentionally smoothing earnings by boosting profit during poorer quarters and stashing away profits during good ones. It is important to note that while Earnings Smoothing provides some evidence that earnings are being intentionally managed from period-to-period, it does not provide information on whether any specific period’s earnings have been shifted, nor what the “true” counterfactual earnings should have been.

Before turning to the analysis, we briefly review the rationale for accrual accounting which, when used properly, can add useful information to reported earnings. For example, suppose a firm incurs a monthly rental expense of $X that is paid in cash 30 days after the 1st of each month. Under cash-based accounting, the firm would show monthly expenses of $X, $0, and $2X for January, February, and March, respectively. In contrast, because of the matching principle, accrual accounting would show an expense of $X in all three months. Because the company incurred the liability when it used the facility, the accrual accounting method shows a truer picture of the financial impact of this use than the cash-based method. In econometric models that use monthly panel data, we might therefore expect ROA to produce a better fit than OCFOA. Similar arguments can be applied to a wide variety of investment and financing activities.

On the other hand, accrual accounting implies a degree of managerial discretion that can be used to obfuscate underlying performance. Suppose, for example, that a firm generates a cash-based loss of $Y in one month by selling product A, and a cash profit of $Y the next month selling product B. If the firm makes an accrual to inflate profits in the first month (e.g., by making a more aggressive prediction about its receivables), and then unwinds that accrual in the next month, the pattern of returns would be $0, $0 under accrual accounting and -$Y, $Y based on cash. Consequently, a regression of “product sold” on profitability would produce no clear result if ROA is used as the outcome variable, but would show that...

---

8 Although there are other measures for earnings management/earnings manipulation that hold value (for example, the Modified Jones method (Dechow et al., 1995)), comparing variation of earnings to variation in cash flows has helpful features for our purpose. Unlike methods that rely on identifying and isolating discretionary accruals, this method covers both “real” income smoothing and “artificial” income smoothing (Ronen and Yaari, 2008). Additionally, this method does not require the existence of a “non-manipulated” period for each firm from which to derive their non-discretionary accrual patterns. Finally, the ratio of standard deviations is more intuitive for a non-accounting audience than detecting anomalies in specific accrual accounts would be.
product B is associated with greater profit when using OCFOA. This latter example also illustrates why the intuition that earnings manipulation simply “averages out” is not correct. Even if all adjustments are eventually reversed, earnings management can generate bias in statistical analyses when it is correlated with other variables, such as a particular manager or strategy.\footnote{When firms exit a data set (e.g., through bankruptcy, acquisition, or going private) we also may not observe the “unwinding” of all accounting adjustments.}

The preceding discussion suggests that accrual-based accounting can provide a better picture of performance over time by matching operational decisions to their financial consequences, and smoothing out idiosyncratic and “lumpy” cash flows. At the same time, accruals may obscure true performance, at least for a while. Ultimately, the information content of ROA relative to OCFOA is therefore an empirical question whose answer will depend, among other factors, on the amount of earnings manipulation and its causes.

4 The Amount of Earnings Management

Accounting scholars are well aware that there is a discontinuity in reported earnings at the zero-profit threshold, and that this “kink” also appears when earnings are scaled by share-price (Hayn 1995) or shareholders’ equity (Burgstahler and Dichev 1997). Figure 2 illustrates this discontinuity using an ROA histogram.

Figure 2: ROA Histogram: U.S. Public Firms 1992-2018

The left panel of Figure 2 is based on the full sample of all U.S. Public Firms from 1992 through 2018. There is a clear spike in the reported ROA distribution at zero (the vertical solid line). The right panel omits firms with a primary SIC code in the financial, insurance, or public administration industries (SICs in the 6000s and the 9000s). Although the large
spike at zero becomes less pronounced in the right panel, there is still a sharp increase in
the probability distribution just above zero. Many empirical studies choose to omit firms in
the financial sector, and this graph suggests there is a logic to that decision, although (as
we show below) it does not eliminate the measurement problem.

In the accounting literature, earnings management is generally accepted as the explana-
tion for the discontinuous jump in the distribution of reported earnings just above zero (e.g.,
Burgstahler and Chuk [2017]). We are aware of no prior study, however, that estimates
how much earnings management occurs around that threshold. To address this gap, and
to provide some sense of the overall the size of the potential measurement problem for em-
pirical strategy research, we use a set of methods developed to analyze economic behavior
around discontinuities in incentives (Chetty, 2012; Kleven and Waseem, 2013; Kleven, 2016).
In particular, Diamond and Persson (2016) suggest a methodology for assessing how much
probability mass is shifted across a threshold where there is a “notch” in incentives (as in the
simple model presented above). We apply their method to the ROA distribution in Figure 2.

At the core of this methodology is a model of the probability distribution of ROA (denoted
by $x$) that takes the following form:

$$P = \sum_{m=1}^{K} \beta_m x^m + \sum_{x=L}^{U} \alpha_x + \sum_{x=0}^{U} \gamma_x + \epsilon$$  (2)

where $P$ is a count of observations at $ROA = x$; the $\beta_m$ are coefficients of a $K^{th}$ order
polynomial in $x$; the parameters $\alpha_x$ ($\gamma_x$) measure the missing (excess) mass due to earnings
manipulation below (above) the zero-profit threshold; and $\epsilon$ is an econometric error term.
Intuitively, this regression uses a flexible polynomial to estimate the un-manipulated coun-
terfactual ROA distribution on the interval $[L,U]$, and the dummies $\alpha_x$ and $\gamma_x$ provide a
flexible fit to the actual data in that manipulated region. This model assumes that (1) there
is a “manipulation zone” around zero – specifically inside the interval $[L,U]$ — where the
ROA measure is distorted, (2) outside of that interval we observe an accurate measure of
ROA, and (3) the counter-factual (unmanipulated) distribution of ROA is continuous on the
interval $[L,U]$, so we can extrapolate from a polynomial estimated on data outside of the
manipulation zone to impute the counterfactual values within.

To complete this empirical model of earnings manipulation requires that we select values
for the parameters $K$, $U$, and $L$. To do so, we use the cross-validation algorithm proposed
in Diamond and Persson (2016), which consists of the following steps:

10Based on citations, this fact does not appear to be widely known to strategic management scholars. For
example, Hayn (1995) has not been cited and Burgstahler and Dichev (1997) is cited by only one article in
Strategic Management Journal.
1. Discretize the underlying ROA data. In practice, we use 200 bins of equal width between ROA values of -1 and 1 (i.e., each bin covers .005 units of ROA).

2. Construct five random samples, by selecting \( N \) observations (with replacement) from the actual ROA data. In each random sample, we treat 80% of the observations as a training data set, and 20% as a holdout sample.

3. Perform a grid search, looping over feasible values of \((K, L, U)\), and for each triple
   
   (a) Estimate equation (2) for given values \((K, L, U)\) on the full dataset. Test the hypothesis that \( \sum_{x=L}^{U-1} \alpha_x = \sum_{x=0}^{U} \gamma_x \) (i.e., the “missing” mass below zero equals the “excess” mass above). If that test rejects at the 10% level or better, move to the next triple.

   (b) If we cannot reject the hypothesis that missing mass equals excess mass, then estimate equation (2) using the values \((K, L, U)\) on each of the five training samples, and compute the mean squared prediction error (MSE) for the associated holdout sample. Store the sum of the MSE across all five test samples.

4. Choose the values \((K, L, U)\) that produced the lowest aggregate MSE at Step 3, and re-estimate that model on the full data set.

The results of this five-fold cross-validation procedure are displayed in Figure 3. The upper \((U)\) and lower \((L)\) bounds of the region of ROA manipulation are indicated by dashed lines. Gray circles indicate the number of firm-year observations in each ROA bin. Black diamonds represent the counterfactual estimate for that bin imputed from our model.

Figure 3: Imputed Vs. Actual ROA

The left panel in Figure 3 plots the actual versus predicted distribution of ROA for the full sample, where the cross-validation procedure selected a 12th degree polynomial with
$L = -0.15$ and $U = 0.08$. For that sample, our model implies that 15.5 percent of all firm-year observations were shifted from negative to positive ROA.

The right panel in Figure 3 shows results if we exclude financial and public-sector firms from our sample. For this sample, the best-fit model was a 15-degree polynomial, with $L = -0.15$ and $U = 0.10$. The model implies that 10.5 percent of all non-financial firm-year observations were shifted from negative to positive ROA. This is almost 30 percent less earnings manipulation than we estimate for the full sample, which suggests that manipulation among financial firms, which only comprise about 20 percent of the full sample, could be quite substantial. Nevertheless, our baseline estimates suggest that around 1 in 10 observations in a paper that employs Compustat ROA is prone to systematic measurement error, even when excluding the financial sector. In Appendix B we show that an alternative methodology that replaces the polynomial in equation (2) with a function of the density of OCFOA (under the assumption that OCFOA is not manipulated), yields similar results, at least for the non-financials.

4.1 Earnings Management vs. Endogenous Effort

At the end of Section 2 we noted that there are at least two explanations for the discontinuity in ROA at the zero-profit threshold: earnings manipulation and a “try-harder” effect. Up to this point, we have focused on measuring the scale of the discontinuity (i.e., what share of all reporting is moved from negative to positive) and discussed those results in terms of earnings manipulation. We now consider two complementary approaches that help to rule out explanations other than earnings manipulation. The first method uses the Earnings Smoothing measure described above, and the second exploits the idea that OCFOA is harder to manipulate than ROA.

Figure 4 shows a binned scatterplot of the mean of Earnings Smoothing conditional on ROA. We have overlaid on this graph a fitted regression line with confidence intervals, and indicated the manipulation region identified as described above using dashed vertical lines. For both the full sample and the sample excluding financial-sector firms, we observe a sharp (discontinuous) increase in earnings smoothing when ROA is just above zero. This indicates that when firms report small positive values of ROA, they also tend to exhibit a sudden increase in the ratio of the variance in accounting earnings to the variance in OCF. Moreover, because these variances are computed within-firm (over the trailing 12 quarters), the evidence of earnings manipulation in Figure 4 is not simply an implication of the baseline

\[11\] In the literature on bunching, round numbers and psychologically important thresholds are called focal points. The effort-based explanation for bunching near focal points has been advanced in other contexts, such as the distribution of marathon finishing times (Allen et al., 2017).
discontinuity illustrated in Figure 3. Put simply, the firms bunching just above zero in the ROA distribution are also characterized by an unusually low level of earnings volatility relative to their cash flows.

If we expand our gaze, moving away from the discontinuity at zero ROA to consider the entire manipulated region of the ROA distribution, it becomes clear that Earnings Smoothing is lower at negative levels of ROA, and higher when ROA is positive. This is a natural consequence of accounting conventions. Firms with higher underlying profitability are less constrained in their ability to smooth earnings, because some financial slack is required in order to reallocate resources. After peaking at an ROA of 5 to 10 percent, the relationship between ROA and Smoothing turns negative, perhaps because managers feel less pressure to manipulate earnings when the business is performing well.

To the extent that our measure of Earnings Smoothing captures what it purports to measure, Figure 4 provides direct evidence against the hypothesis that bunching in the ROA distribution at zero is caused by endogenous effort rather than earnings manipulation. As another test, however, we can apply our cross-validation approach directly to OCFOA to estimate the amount of “cash flow manipulation” at the same threshold. Under the maintained assumption that it is more difficult for CEOs to manipulate cash flow than accounting earnings, we would expect to find less evidence of OCFOA manipulation. Figure 5 shows the results of that exercise.

The top two panels in the Figure compare ROA to OCFOA manipulation for the full sample, and the bottom two panels compare ROA to OCFOA manipulation for the non-financials. It is clear even from visual inspection that the size of the discontinuity around zero and the subsequent bunching above zero is dramatically reduced by using the cash-basis approach.

---

12In the Appendix, we provide the histograms corresponding to each panel in Figure C.2
Figure 5: Comparison of Imputed Vs. Actual ROA and OCF
performance measure of OCFOA rather than the accrual-basis performance measure of ROA. For the full sample, our estimates imply that four percent of the observations are “shifted” from negative to positive OCF. In the non-financial sample, we estimate that the amount of earnings manipulation is negative. Instead of “missing” mass below zero, there are slightly more negative observations that were predicted. This evidence, we interpret as essentially no sign of left-to-right OCFOA manipulation.

4.2 Robustness

The analyses in this section yield three basic facts. First, there is a substantial amount of earnings manipulation (on the order of 15% of all firm-year observations) around the zero-profit threshold. Second, manipulation is especially prevalent among firms in the financial sector. Third, there is much less manipulation of OCFOA, and essentially none for non-financial firms. We have considered a number of supplemental analyses and robustness checks that further support these findings.

First, we checked whether Earnings Smoothing was continuous at the zero-OCFOA threshold, and whether OCFOA was continuous at the zero-ROA threshold. In the Appendix, we show that for non-financial firms, there is no evidence of smoothing to achieve positive cash flow, and that real-earnings management (i.e., manipulation of OCF to achieve positive ROA) is confined to the financial sector. Both results are consistent with our findings that accounting earnings are more prone to manipulation than cash flows.

Second, as an alternative to the specification in equation (2) that relies on functional form to estimate the counterfactual density of ROA in the interval \([L, U]\), we developed a model that uses OCFOA to predict ROA. This approach rests on the maintained assumption that OCFOA is not manipulated, and as a result, works better for the sample that excludes financial-sector firms. The results, provided in Appendix B, indicate that around 6 percent of firm-year observations in our non-financial sample are manipulated.

Finally, there is a concern that the missing mass in our figures might be caused by a liquidation option for struggling firms. In particular, if those firms most likely to post accounting losses leave the dataset due to bankruptcy, acquisition by another firm, or being taken private, that could produce a “hole” in the earnings distribution just below zero. This hypothesis does not explain the bunching of reported earnings just above zero. Nevertheless, we have replicated our main results on a dataset that excludes firms firms that exit the Compustat before the end of the sample period (regardless of whether the exit was due to bankruptcy, liquidation, leveraged buyout, etc.) with substantially similar results.

---

\(^{13}\)See Figure C.3.

\(^{14}\)See Figure B.1.
5 Earnings Management and Empirical Strategy Research

Having established that there is a large amount of earnings manipulation near the zero-profit threshold, the question remains whether this “matters” for empirical strategy research. To address this question, we return to an old but influential line of studies that seeks to attribute variation in performance to firm, industry, and macro-economic factors (e.g. Schmalensee, 1985; Rumelt et al., 1991; McGahan and Porter, 2002). Our goal is not to replicate prior studies, or to address any of the methodological shortcomings of variance decomposition that are well-documented in previous studies. Rather, we aim to show how the results of this type of analysis change when we move from ROA to OCFOA as a measure of firm performance.

Our analysis will consider two ways in which earning manipulation might matter. First, it may add “classical” measurement error that reduces the overall explanatory power of a model. Second, and more importantly in our view, earnings management might be correlated with other variables (e.g., if it is more prevalent in specific industries, and linked to certain CEOs). To the extent that earnings manipulation is correlated with other factors, it has the potential to introduce bias into analyses that use ROA as an outcome.

The foundational studies in this literature estimated models that might include year, industry, firm, and/or business-unit fixed effects. By comparing the model R-squared for different combinations of variables, it is possible to compute how much total variance is explained by each of the observed factors. One limitation of using OCFOA in this context is that operating cash flows are not required to be reported at the business segment level, and therefore a direct replication of the classic studies is not possible. In particular, our “industry” effects are based on the primary SIC code assigned to the firm as a whole, rather than to an individual business unit. On the other hand, we can extend upon the early papers by using the Execucomp data set to include CEO fixed effects, following later scholars in this literature stream (e.g. Mackey, 2008).

Our analysis is based on the following model for the generation of reported accounting profit:

\[ r_{t,j,i,k} = \mu + \gamma_t + \alpha_i + \beta_j + \delta_k + \varepsilon_{t,j,i,k} \]  \hspace{1cm} (3)

In this equation, \( r_{t,j,i,k} \) is either the ROA or OCFOA reported in a given year \( t \) by a specific firm \( j \) operating within industry \( i \) and led by CEO \( k \). \( \mu \) is the average accounting profit over the entire sample (the constant in the regression models), and the other variables represent...
fixed effects for the year ($\gamma_t$), the industry ($\alpha_i$), the firm ($\beta_j$), and the CEO ($\delta_k$), as well as the error term ($\epsilon_{t,i,j,k}$).

For the sequential ANOVA model, we incrementally added fixed effects for year, industry, firm, and CEO to gauge the marginal contribution of R-squared gained with the addition of each set of fixed effects. This approach was used by scholars earlier in this literature stream, but has a significant flaw of being sensitive to the order in which the fixed effects are added, as noted prominently by McGahan and Porter (2002) and Mackey (2008). In sequential ANOVA, variance that could be explained by either of two nested levels of fixed effects will be attributed to the first one added to the model.

This drawback is alleviated by the second approach, a simultaneous ANOVA model. In the simultaneous model, variance that could be explained by more than one factor is not attributed to either of them. This has the benefit of avoiding misattribution of explained variance, while it also has the drawback of leading to lower estimates of variance explained for each category, as the ambiguous cases are not attributed at all. However, the total R-squared for the full model with all fixed effects is not understated even if the category breakdown may be (i.e., the total R-squared for the model exceeds that of the sum of the categories).

Figure 6 presents the results of the explanatory value of the full models for both ROA and OCFOA, sequential and simultaneous, for both all industry and non-financial industry samples. The key finding here is that our ability to predict/explain variance in OCFOA exceeds that of ROA by approximately 10 percentage points across all specifications. As the entire point of accrual accounting is to add salient information and remove noise from cash-basis performance, the 10 percentage points of explained variance should be considered a fairly conservative lower bound for how much obfuscation appears to be introduced by
strategic accounting decisions. Not only are accruals not giving us a clearer picture of underlying financial performance (as they are supposed to do), they are actively worsening the signal-to-noise ratio in the most common measure of performance used in the strategy literature.

If earnings management introduces measurement error in ROA, under what conditions should we be concerned with bias rather than merely a loss of efficiency? If earnings management caused primarily classical measurement error in ROA, it would not cause us great concern when using ROA in our econometric models. When we used ROA as an outcome variable, this would simply reduce the efficiency and increase the standard errors around our coefficients. When ROA was used as an explanatory variable, it would attenuate the coefficient towards zero, which is often toward a more conservative interpretation, i.e., pulling our inference towards the null [Bound et al., 2001]. But unfortunately, there is reason to believe that the measurement error caused by earnings management on ROA is non-classical.

The key assumption of classical measurement error is that the error itself is uncorrelated with values of the measure, but also that the error is uncorrelated with other variables in the econometric model [Hyslop and Imbens, 2001]. We saw from the analysis in the previous sections that the prevalence/degree of manipulation was not evenly spread across all values of ROA. Indeed, it is concentrated enough in a region of ROA to cause visual discontinuities in the distribution.

In addition to the correlation with ROA itself, there is also reason to suspect that the measurement error from earnings management is correlated with other variables that may be in our econometric equations. An easy way to see this is in looking at the differential impact using OCFOA vs. ROA has on the amount of variance explained in each category of fixed effects in the ANOVA models. Figure 7 shows the breakdown of explained variance for each of the categories of the ANOVA—year, industry, firm, and CEO. Across the models, the relative explanatory power of firm and industry lowered when using ROA rather than OCFOA, while the relative explanatory power of CEO and year increased. In the all industries nested ANOVA model, this effect is large enough to change the rank order of CEO and industry effects by reversing their relative importance.

One interpretation of the results in Figure 7 is that certain CEOs are more likely to manipulate earnings, so that moving from ROA to OCFOA as the focal measure of firm performance causes the share of variance attributed to CEO effects to decline, and the share of variance explained by firm and industry-level factors to increase.

\[16\text{For tabular format, please see Table C.1}\]
6 Conclusions

We make three contributions in this study. First, we provide new evidence on the prevalence of earnings management, a well-known problem that has resisted precise measurement or quantification. We find evidence that 10 to 15 percent of firm-year observations in Compustat exhibit earnings manipulation. Our estimates also indicate that earnings manipulation is more prevalent in the financial sector, thereby providing a firmer empirical foundation for the “folk wisdom” that one might want to exclude financials when analyzing firm performance with accounting data. These finding augment the literature using regression discontinuity designs (Burgstahler and Chuk 2017), studies leveraging discretionary accruals (Dechow et al. 1995) or accrual reversals (Dechow et al. 2012), and survey designs targeting chief financial officers (Graham et al. 2005) by employing novel methods from the econometrics literature on bunching.

Second, these bunching methods are employed to evaluate alternative performance measures. We find that OCFOA exhibits less manipulation, and thus provides a method by which scholars can test the sensitivity of models including accounting profits for bias introduced by earnings manipulation. This study thus dovetails with current efforts to rethink and improve how we measure performance (Lieberman 2021, Lieberman et al. 2017, Wibbens and Siggelkow 2020).

Finally, we deploy our insight about OCFOA to re-evaluate a classic set of strategy papers that uses variance decomposition to understand the drivers of firm-performance (Schmalensee 1985, Rumelt et al. 1991, McGahan and Porter 1997, 2002). Our results suggest, counter-intuitively, that we can explain more of the total variance in cash-based rather than accrual-based accounting performance. Moving from ROA to OCFOA also reduces the amount of variance in firm-performance associated with CEO effects, which suggest

Figure 7: Decomposition of Variance of ROA Vs. OCF
that some CEOs are more likely to engage in manipulation than others.

Our findings have implications for empirical work where firm performance is measured using accounting profit. For many studies, restricting the sample to non-financial firms and utilizing OCFOA as a performance measure for accounting profitability offers a simple way to avoid potential econometric problems created by earnings management. More generally, researchers should carefully consider whether firms’ unobserved propensity to inflate profits could be correlated with key outcomes or explanatory variables. In some cases, such as when ROA serves as an ancillary control variable, this will not be especially problematic. But when ROA is the outcome, and key explanatory variables might be correlated with the propensity to manipulate, researchers should explore sensitivity to using OCFOA. Our findings also suggest that these issues may be particularly salient when exploring the relationship between CEO attributes and firm performance.

Although the strategic management literature provides an extensive toolkit for analyzing firm performance and value creation, we often take accounting and the measurement of these constructs for granted. In our view, empirical scholars could be more attuned to “how the sausage gets made” when it comes to performance measurement, because it is fertile ground for future research on how earnings measurement affects strategic decision-making within the firm and value creation across a broad spectrum of stakeholders.
References


Appendix A: Derivation of $\pi_L$

To find the lower threshold of the “hole” in reported earnings (i.e., $\pi_L$), we can look for solutions of $U(\frac{1}{\gamma}; \pi) = U(-\pi; \pi)$. At that point, the CEO is indifferent between making a larger adjustment that achieves the bonus $B$ or staying with the locally optimal report $a^* = \frac{1}{\gamma}$. Substituting into equation (1) and simplifying leads to the quadratic equation

$$\frac{\gamma}{2} \cdot \pi^2 + \pi + \left(\frac{1}{2\gamma} - B\right) = 0$$

The roots of this quadratic are $\pi = \frac{-1 \pm \sqrt{2\gamma B}}{\gamma}$. The larger root cannot be the solution, because for $\pi > \frac{1}{\gamma}$ the CEO would obtain the bonus under her “normal” reporting strategy $a^* = \frac{1}{\gamma}$. This implies that the solution for the lower threshold must be $\pi_L = \frac{-1 - \sqrt{2\gamma B}}{\gamma}$, or equivalently, $\pi_L \equiv -\left(\frac{1}{\gamma} + \sqrt{\frac{2B}{\gamma}}\right)$, as reported in the text.
Appendix B: Alternative Bunching Estimates

Blomquist and Newey (2018) critique the use of bunching estimators in public finance to estimate the tax elasticity of income using kinks or notches in the tax schedule. The core of their argument is that identification rests on functional form assumptions. In particular, within the region where outcomes are assumed to be manipulated, bunching methods impute counterfactual outcomes entirely from extrapolation, rather than any comparison of observed quantities. For example, if the true counterfactual distribution of ROA is highly non-linear around zero, then the estimates of earnings management that we report in Section 4 could be biased.

Setting aside any debate over the practical implications of this critique, there is a natural solution available in our empirical setting. If we assume that OCFOA is not manipulated, then the relationship between OCFOA and ROA helps identify the counterfactual distribution of ROA within the manipulated region. To implement this idea, we use the following model:

\[ P = \beta CF_{x+T} + \sum_{x=L}^{x=U-1} \alpha_x + \sum_{x=0}^{U} \gamma_x + \epsilon \]  

where \( P \) is the number of observations with ROA equal to \( x \), \( CF_{x+T} \) is the number observations with OCFOA equal to \( x + T \), and all other parameters are defined as in equation (2). Comparing (4) to (2), it should be clear that we have simply replaced the polynomial previously used to extrapolate ROA in the region \([L, U]\) with a linear function of the OCFOA that is “shifted” by \( T \) bins.

The reason we allow for the OCFOA distribution to be shifted relative to ROA is that operating activities are normally a profit center, so OCFOA generally exceeds ROA (e.g., due to taxes, depreciation, et cetera). We select a value of \( T \) using the cross-validation procedure described in Section 4 of the paper, searching over \( T \) rather than \( K \) (the degree of the polynomial). Figure B.1 shows the resulting histogram for all firms as well as the non-financial sample. As for earlier figures, the upper (\( U \)) and lower (\( L \)) bounds of the region of ROA manipulation are indicated by dashed lines, gray circles indicate the number of firm-year observations in each ROA bin, and black diamonds represent the counterfactual estimate for that bin imputed from the model.

For both panels, the cross-validation procedure selected a leftward shift of 5 bins for OCFOA (\( T = 5 \)). The left panel shows the results from the full sample, which has a lower bound (\( L \)) of -0.13, an upper bound (\( U \)) of 0.09, and a total amount of displaced probability mass of 16.8 percent. The right panel shows the results from the non-financial sample, which has a lower bound (\( L \)) of -0.14, an upper bound (\( U \)) of 0.09, and a total amount of
displaced probability mass of 6.4 percent. In both cases, our estimates of total earnings manipulation decline slightly, because the distribution of OCFOA has a more pronounced peak than the counterfactuals based on a polynomial approximation. This can be seen be comparing Figures B.1 and 3.

Finally, we note the spike in the predicted values of ROA just below zero in the left panel. This corresponds to a discontinuity in the distribution of OCFOA for financial firms that can also be observed in the top right panel of Figure 5. We interpret this spike as evidence of real earnings management (i.e., manipulation that also influences OCF) by financial firms. It suggests that the approach used in this Appendix will work better for the non-financial sector, whereas the standard approach of relying on a polynomial extrapolation may be more reliable for the full sample.
Appendix C: Supplemental Tables and Figures

Figure C.1: Binned Scatterplot of ROA and OCFOA
Figure C.2: Comparison of ROA and OCFOA Histograms
Figure C.3: Comparison of Smoothing Vs. ROA and OCFOA
Table C.1: Nested and Simultaneous ANOVA Results

All Industries - Nested ANOVA

<table>
<thead>
<tr>
<th>Category</th>
<th>ROA Variance Explained</th>
<th>ROA Percentage of Explained Var.</th>
<th>OCFOA Variance Explained</th>
<th>OCFOA Percentage of Explained Var.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year</td>
<td>.0138</td>
<td>2.61</td>
<td>.0032</td>
<td>0.51</td>
</tr>
<tr>
<td>Industry</td>
<td>.0700</td>
<td>13.23</td>
<td>.1520</td>
<td>24.34</td>
</tr>
<tr>
<td>Firm</td>
<td>.3071</td>
<td>58.02</td>
<td>.3494</td>
<td>55.95</td>
</tr>
<tr>
<td>CEO</td>
<td>.1384</td>
<td>26.15</td>
<td>.1199</td>
<td>19.20</td>
</tr>
<tr>
<td>Full Model</td>
<td>.5293</td>
<td>100</td>
<td>.6245</td>
<td>100</td>
</tr>
</tbody>
</table>

All Industries - Simultaneous ANOVA

<table>
<thead>
<tr>
<th>Category</th>
<th>ROA Variance Explained</th>
<th>ROA Percentage of Explained Var.</th>
<th>OCFOA Variance Explained</th>
<th>OCFOA Percentage of Explained Var.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year</td>
<td>.0149</td>
<td>3.65</td>
<td>.0033</td>
<td>0.89</td>
</tr>
<tr>
<td>Industry</td>
<td>.0212</td>
<td>5.19</td>
<td>.0254</td>
<td>6.86</td>
</tr>
<tr>
<td>Firm</td>
<td>.2128</td>
<td>52.07</td>
<td>.2120</td>
<td>57.22</td>
</tr>
<tr>
<td>CEO</td>
<td>.1598</td>
<td>39.10</td>
<td>.1298</td>
<td>35.03</td>
</tr>
<tr>
<td>Full Model</td>
<td>.5555</td>
<td>100</td>
<td>.6386</td>
<td>100</td>
</tr>
</tbody>
</table>

SICs in 6000s and 9000s Omitted - Nested ANOVA

<table>
<thead>
<tr>
<th>Category</th>
<th>ROA Variance Explained</th>
<th>ROA Percentage of Explained Var.</th>
<th>OCFOA Variance Explained</th>
<th>OCFOA Percentage of Explained Var.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year</td>
<td>.0147</td>
<td>2.79</td>
<td>.0019</td>
<td>0.31</td>
</tr>
<tr>
<td>Industry</td>
<td>.0624</td>
<td>11.85</td>
<td>.1155</td>
<td>19.07</td>
</tr>
<tr>
<td>Firm</td>
<td>.3112</td>
<td>59.07</td>
<td>.3626</td>
<td>59.86</td>
</tr>
<tr>
<td>CEO</td>
<td>.1385</td>
<td>26.29</td>
<td>.1257</td>
<td>20.75</td>
</tr>
<tr>
<td>Full Model</td>
<td>.5268</td>
<td>100</td>
<td>.6057</td>
<td>100</td>
</tr>
</tbody>
</table>

SICs in 6000s and 9000s Omitted - Simultaneous ANOVA

<table>
<thead>
<tr>
<th>Category</th>
<th>ROA Variance Explained</th>
<th>ROA Percentage of Explained Var.</th>
<th>OCFOA Variance Explained</th>
<th>OCFOA Percentage of Explained Var.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year</td>
<td>.0165</td>
<td>3.99</td>
<td>.0038</td>
<td>1.00</td>
</tr>
<tr>
<td>Industry</td>
<td>.0196</td>
<td>4.74</td>
<td>.0232</td>
<td>6.05</td>
</tr>
<tr>
<td>Firm</td>
<td>.2162</td>
<td>52.30</td>
<td>.2200</td>
<td>57.35</td>
</tr>
<tr>
<td>CEO</td>
<td>.1611</td>
<td>38.97</td>
<td>.1366</td>
<td>35.61</td>
</tr>
<tr>
<td>Full Model</td>
<td>.5544</td>
<td>100</td>
<td>.6215</td>
<td>100</td>
</tr>
</tbody>
</table>