The Effects of Firm Growth and Model Specification Choices on Tests of Earnings Management

Daniel W. Collins, Raunaq S. Pungaliya, and Anand M. Vijh *

Abstract

We show that commonly used Jones-type discretionary accrual models applied in quarterly settings do not adequately control for nondiscretionary working capital accruals that naturally occur due to firm growth. This biases tests of earnings management in many settings where the partitioning variable is correlated with firm growth (such as stock splits, SEOs, stock acquisitions, and stock-based compensation). Using data for a comprehensive sample of Compustat firms, we estimate the biases associated with popular alternative discretionary accrual model specifications with and without controls for performance (ROA), accruals’ noise reduction role, and non-linearity due to timely loss recognition, but not firm growth. We show that there is a severe problem of falsely rejecting the null hypothesis of no earnings management in samples over-represented by high growth or low growth firms when using performance-adjusted discretionary accruals. In contrast, discretionary accrual models that control for both performance and firm growth are well specified and do not sacrifice power. Including adjustments for accruals’ noise reduction and timely loss recognition roles further improves the model power.

March 2012

* Collins and Vijh are from Tippie College of Business, The University of Iowa, Iowa City, Iowa 52242-1994, USA. Pungaliya is from Sungkyunkwan University, Seoul, Korea. The authors appreciate comments by Ray Ball, Phil Burger, Frank Ecker, Merle Erickson, Dave Folsom, Cristi Gleason, Paul Hribar, Bruce Johnson, S.P. Kothari, Ed Maydew, and workshop participants at the University of Chicago, Duke University, The University of Iowa, and Lehigh University. Email: daniel-collins@uiowa.edu, raunaq@skku.edu, and anand-vijh@uiowa.edu.
The Effects of Firm Growth and Model Specification Choices on Tests of Earnings Management

1. Introduction

An extensive body of literature in accounting and finance uses Jones-type model discretionary accrual estimates to test for earnings management. This literature includes studies that test for evidence of earnings management around specific corporate events (e.g., initial public offerings (IPOs), seasoned equity offerings (SEO), mergers and acquisitions, proxy contests, share repurchases, and stock-splits) as well as studies that test for cross-sectional differences in earnings management as a function of firms’ contracting characteristics (e.g., stock-based management compensation arrangements and debt contracting environment).\(^1\) Much of the research to date fails to control for the effects of firm growth on estimates of discretionary accruals. Dechow, Kothari, and Watts (1998) develop an analytical model that highlights the fact that high sales growth firms require legitimate higher investments in working capital to deal with higher customer demand. Their model implies that growth-related changes in accruals should be treated as non discretionary because this component of accruals is predictable and common across growth firms. Thus, in the absence of controls for firm growth, standard Jones-type discretionary accrual estimates will be confounded with innate growth accrual effects.

McNichols (2000) is among the first to recognize the confounding effects of growth on discretionary accrual estimates. She posits and finds that firms with greater expected earnings growth are likely to have greater accruals than firms with less expected earnings growth. She concludes that “… earnings management measures based on the Jones and modified-Jones model approach are not sufficiently powerful or reliable to assess earnings management behavior in many contexts … I provide evidence of possible misspecification of these models, which have been used extensively in the literature to identify discretionality behavior” (p. 337). [Emphasis added]

Kothari, Leone, and Wasley (2005) examine the specification and power of Jones-type discretionary accrual models using annual data and show these accruals are correlated with firm performance. They find that both Jones model and modified-Jones model residuals adjusted by the

---

\(^1\) See Appendix 1 for several references to event studies as well as cross-sectional studies of earnings management.
residuals of same-industry firms matched on ROA yield reasonably well-specified tests of earnings management in most stratified random samples. Further, they conclude that “Performance-matched discretionary accruals exhibit only a modest degree of misspecification (emphasis added) when firms are randomly selected from an extreme quartile of stocks ranked on firm characteristics such as the book-to-market ratio, firm size, sales growth, and earnings yield.” (page 167). Despite the warnings about possible misspecification due to failure to control for firm growth issued by McNichols (2000), most earnings management studies over the past decade (see Appendix 1 for summary) follow the guidance provided in Kothari, Leone, and Wasley and use performance (ROA)-matched Jones-type model discretionary accrual estimates in tests of earnings management, implicitly assuming that any distortion due to firm growth is minimal.

In this paper we extend the analysis in McNichols (2000) in several ways. First, we identify multiple partitioning variables used in prior research that tests for earnings management and demonstrate how these partitioning variables are correlated with firm growth measures. We show how average discretionary accrual estimates change in these settings after controlling for firm growth. Next, using random samples of firms with no known earnings management stratified by growth, we provide evidence of severe bias (high Type I error rates) in tests for earnings management in quarterly settings that fail to control for growth. Further, we show that the tendency for over-rejection persists in models that correct for accruals’ noise reduction and timely loss recognition roles (Ball and Shivakumar, 2006) in samples over-represented by either high growth or low growth firms. Using simulation analysis, we demonstrate that Jones-type discretionary accrual estimates industry-matched on both performance (ROA) and sales growth result in well specified tests with high power of detecting modest levels (0.25% of total assets) of earnings management, particularly for models that explicitly take into account accruals’ noise reduction and timely loss recognition roles.

Concerns arise that matching on sales growth may “throw the baby out with the bathwater” when revenues are manipulated. Contrary to what one might expect, we demonstrate that matching on sales growth introduces very little downward bias (typically, less than 5 basis points) in discretionary accrual estimates when earnings are managed through revenue manipulation. We also demonstrate that reversal
methodology recently advanced by Dechow et al. (2012) as having greater power than matching procedures when applied in annual settings actually yields tests of lower power in quarterly settings where the number of quarters over which reversals occur is less certain and the analysis is confounded by seasonality.

The remainder of the paper is organized as follows. Section 2 outlines a generic discretionary accruals framework for evaluating potential bias in tests of earnings management. We document the degree of correlation between firm growth and a wide range of events and firm characteristics that prior research has hypothesized to be associated with earnings management and estimate the degree of bias in alternative Jones-type discretionary accrual measures that have been widely used in prior research. Section 3 demonstrates graphically and numerically the bias in Jones-type discretionary accrual estimates across a comprehensive sample of Compustat firm-quarters partitioned by ROA deciles and sales growth deciles. Section 4 compares Type I error rates for alternative Jones-model tests of income-increasing and income-decreasing earnings management across all Compustat firms and in samples with varying degrees of over-represented firms from extreme quintiles of sales growth, employee growth, and ROA. Section 5 uses simulation analysis to compare Type II error rates and power of alternative discretionary accrual models with varying levels of upward and downward earnings management seeded in the data, and provides evidence on the bias in alternative Jones-type discretionary accrual estimates with and without adjustment for firm growth. Section 6 presents simulation results that address the concern of whether matching on sales growth throws the baby out with the bathwater when earnings management is accomplished through revenue manipulation. Section 7 compares the ROA + SG matching procedure to the reversal methodology recently proposed by Dechow et al. (2012) and offers some simulation results on the relative power of these two approaches in quarterly settings. Section 8 concludes and summarizes the implications of our findings for future earnings management research.
2. Discretionary accruals and earnings management partitioning variables

2.1 Generic framework for assessing earnings management

An unbiased test of earnings management requires that measurement error in the discretionary accruals proxy be uncorrelated with the partitioning variable in the research design. To help understand the interpretation problems that are created when this condition is not met, McNichols and Wilson (1988) outline a general discretionary accruals framework that is relevant to assessing the potential bias in earnings management studies that use discretionary accruals estimates. In their framework, accruals are partitioned into a discretionary \( \text{DACC}^* \) and non-discretionary \( \text{NDACC}^* \) component, such that:

\[
\text{ACC} = \text{NDACC}^* + \text{DACC}^*
\]

where the * denotes the true, but unobservable, accrual components. Because \( \text{DACC}^* \) is unobservable, it is typically estimated using one of several alternative Jones-type discretionary accrual models. The estimate of discretionary accruals \( \text{DACC} \) that emerges from these models inevitably measures ‘true’ and unknowable discretionary accruals \( \text{DACC}^* \) with error \( \eta \):

\[
\text{DACC} = \text{DACC}^* + \eta
\]

Researchers have used this framework to test for earnings management surrounding such diverse events as SEOs, proxy contests, stock acquisitions, management buyouts, asset write-downs, and stock splits. Typically, the test is conducted by regressing discretionary accruals on a partitioning variable \( \text{PART} \), which is a dummy variable equaling one in the period(s) in which earnings management is hypothesized to occur. A theoretical model to test for earnings management can be represented as follows:

\[
\text{DACC}^* = \alpha + \beta \text{PART} + \varepsilon
\]

---

2 Examples of discretionary accruals models include the Healy (1985) model, the DeAngelo (1986) model, the Jones (1991) model, the modified-Jones model (Dechow, Sloan, and Sweeny, 1995), and the Kang and Sivaramakrishnan (1995) model. Because of their prevalence in accounting research, the analysis in this paper focuses on variations of the Jones model and modified-Jones model.
The significance of $\beta$ is used to draw inferences about the presence of earnings management (or lack thereof). However, because a discretionary accrual proxy, $DACC$, is used instead of $DACC^*$, McNichols and Wilson (1988) demonstrate that equation (3) can be rewritten as follows:

$$DACC = \alpha + \gamma PART + \nu,$$

where

$$\gamma = \beta + \rho_{PART, \eta} \times \sigma_\eta / \sigma_{PART} = \beta + \text{bias in } \gamma$$

Hence, given measurement error in $DACC$, the coefficient used to test for the presence of earnings management is biased. Moreover, this bias will be (1) increasing in the correlation between $\eta$ and $PART$, (2) increasing in the variance of $\eta$, and (3) decreasing in the variance of $PART$. More importantly, with a significant bias in the discretionary accrual estimate, an erroneous conclusion could be made that earnings are managed (i.e., observe non-zero values of $\gamma$) when, in fact, earnings are not managed at all (i.e., $\beta = 0$).

2.2 The correlation between alternative partitioning variables and firm growth measures

We examine the pervasiveness and magnitude of the bias that exists in extant earnings management studies that fail to control for firm growth, in two steps. First, we show the association between five key partitioning variables and firms ranked on two measures of firm growth—sales growth (SG) and employee growth (EG). Next, we quantify the error in Jones-type discretionary accrual estimates that are not adjusted for growth in three event-driven settings and estimate the potential bias as shown in equation (4).

The five partitioning variables we consider are stock splits, SEOs, stock-for-stock acquisitions, percentage of stock-based (executive) compensation, and abnormal insider selling. Prior research has hypothesized and shown (see list of studies in Appendix 2) each of these partitioning variables to be significantly associated with upward earnings management (i.e., significantly positive discretionary accruals as reflected in positive $\gamma$ in equation (4)).

For the first three partitioning variables, we start with a comprehensive sample of firm-quarters from 1991 to 2007 from the Compustat and CRSP databases and merge it with samples of firms that
announced stock splits, SEOs, and stock acquisitions.\(^3\) We require that the included firm-quarters have a CRSP share code of 10 or 11 and an asset value greater than $10 million. We also require that the quarterly earnings announcement date is available in Compustat. We exclude financial firms. The sales growth is calculated as the sales during the quarter with the earnings announcement date preceding the event date of interest divided by the sales during the same quarter of the previous year, minus one. The corresponding decile ranks are calculated each quarter using the data for all firm-quarters. Because the employee numbers are only available annually, we follow a parallel procedure with a comprehensive sample of firm-years to calculate employee growth deciles. Stock splits are identified from the CRSP database using distribution code of 5523 and a positive split factor, and SEOs and stock acquisitions are identified from the SDC database.

Figure 1 shows the frequency distribution of 2,646 stock splits, 2,951 SEOs, and 1,193 stock acquisitions across sales growth (SG) deciles (Panel A) and employee growth (EG) deciles (Panel B). As shown, there is a strong positive relation between all three partitioning events and both firm growth measures, with a large proportion of these events falling into the upper two decile ranks of firm growth.

[Insert Figure 1 here]

For the stock-based compensation and abnormal insider selling partitions, we start with a comprehensive sample of 41,383 firm-years (instead of firm-quarters) during 1991 to 2007 from the Compustat and CRSP databases and select a subset of firm-years for which stock-based compensation data are available from ExecuComp (1992 to 2007) or insider buying and selling data are available from Thomson Financial (1991 to 2007).\(^4\) The insider trading data pass through several filters commonly employed in previous literature.\(^5\) Stock based compensation is calculated as the Black-Scholes value of stock option grants plus the market value of restricted stock divided by total compensation and this quotient is multiplied by 100. Total compensation is defined as the value of stock options and restricted

\(^{3}\) The construction of the comprehensive sample of firm-quarters and the calculation of accrual measures is provided below in Section 3.1.

\(^{4}\) We use firm-years because executive compensation data is only available on an annual basis from ExecuComp.

\(^{5}\) We collect data as reported on form 4 filed with the SEC. We restrict to cleanse codes R and H, which indicate the highest level of confidence in data, and transaction codes P and S, which indicate open market or private purchase and sale of non-derivative or derivative security. We also restrict to transactions involving at least 100 shares.
stock plus salary and bonus. Following Beneish and Vargus (2002), firm-years characterized by abnormal insider selling are identified as follows. First, we sum the total sales and the total purchases of shares by the top five executives, calculate the difference, and divide by the total shares outstanding. Second, we check whether this scaled difference is greater than the corresponding median value for all firm-years with the same market value decile rank.

The left bars in Panels A and B of Figure 2 show the median stock-based compensation as a percentage of total compensation for firm-years ranked by sales growth decile and employee growth decile, respectively. The right bars show the percent of all firm-years for which there was abnormal insider selling for both decile rank growth measures. Although the positive relationship between these two partitioning variables and firm growth is not as strong as in Figure 1, both stock-based compensation and abnormal insider selling tend to be concentrated in high growth firms. The clear take-away from these two figures is that failure to control for firm growth in these settings is likely to result in upward biased estimates of discretionary accruals and a bias in favor of finding earnings management.

[Insert Figure 2 here]

2.3 Alternative Jones-type model discretionary accrual specifications

The two most popular models for estimating the discretionary component of accruals are the cross-sectional Jones model (Jones, 1991) and modified-Jones model (Dechow, Sloan, and Sweeney, 1995). The quarterly equivalents of these two models for current or working capital accruals \((WCA_{i,t})\) are specified below:

**Quarterly Jones Model:**

\[
WCA_{i,t} = \beta_0 + \beta_1 Q_{1,i,t} + \beta_2 Q_{2,i,t} + \beta_3 Q_{3,i,t} + \beta_4 Q_{4,i,t} + \beta_5 \Delta SALES_{i,t} + \\
\beta_6 WCA_{i,t-4} + \epsilon_{i,t}.
\]  

(5)

In this expression, subscript \(i\) denotes the firm and \(t\) denotes the calendar quarter. \(Q_{1,i,t}\) to \(Q_{4,i,t}\) are fiscal quarter dummies that allow for possible fiscal quarter effects in accruals, \(\Delta SALES_{i,t}\) is the quarterly change in sales measured relative to the previous quarter’s sales, and \(WCA_{i,t-4}\) is the working capital accruals from the same fiscal quarter in the preceding year. All independent variables except the intercept
term are scaled by lagged total assets. Using Compustat data, the regressions are run by calendar quarter for the cross-section of all firms belonging to the same industry as the sample firm (i.e., same two-digit SIC code). The Jones model discretionary accruals are calculated as the residuals $\varepsilon_{i,t}$ from equation (5).  

Quarterly Modified-Jones Model—Common Specification [Mod-Jones(C)]:

Mod-Jones(C) model discretionary accruals are calculated as the residuals $\xi_{i,t}$ from the following model. We examine the most common way of estimating the modified-Jones model that treats all credit sales in the event period and the estimation period as discretionary for both the treatment and control firms included in the regression [we refer to this as Mod-Jones(C)].

$$WCA_{i,t} = \lambda_0 + \lambda_1 Q_{1,i,t} + \lambda_2 Q_{2,i,t} + \lambda_3 Q_{3,i,t} + \lambda_4 Q_{4,i,t} +$$

$$\lambda_5(DSALES_{i,t} - \Delta R_{i,t}) + \lambda_6 WCA_{i,t-4} + \xi_{i,t}$$

where all notation have the same meaning as described above. $\Delta AR_{i,t}$ is measured over adjacent quarters.

Models that adjust for accruals’ role in noise reduction and timely loss recognition:

Ball and Shivakumar (2006) posit that accruals serve two major purposes: (1) ameliorating transitory shocks to operating cash flows (CFO); and (2) promoting efficient contracting by providing timely loss recognition. They demonstrate how explicitly recognizing these two roles results in formulation of non-linear discretionary accruals models that offer substantial specification improvement over existing models. Models that adjust for accruals’ noise reduction and timely loss recognition roles

---

6 Throughout this paper we adopt the one-step approach to estimating discretionary accruals that is the dominant approach in the literature subsequent to the Kothari, Leone, and Wasley (2005) paper. Under this approach, both treatment and control (benchmark) firm observations are included in the estimating equation used to determine non-discretionary accruals. This is in contrast to the two-step approach that uses only observations from the non-event or control firm sample to estimate parameters for determining non-discretionary accruals. These parameter estimates are then combined with observed values of the economic determinants of accruals for treatment firms in the event period to form expected (non-discretionary) accruals. The difference between the actual and expected accruals is used as the proxy for abnormal or discretionary accruals for treatment firms in the event period that is hypothesized to give rise to earnings management. The choice between the two approaches has little effect on Type I error rates, but the one-step approach can be slightly less powerful when there is clustering of data in calendar time or within industry.

7 We also estimate discretionary accruals using the original specification of the modified-Jones model proposed by Dechow, Sloan and Sweeney (1995), which treats all credit sales in the event period (but not in the estimation period or benchmark sample) as discretionary. For brevity, we do not table these results, but they are available from the authors upon request.

8 The Mod-Jones(C) model assumes nondiscretionary accruals [the fitted part of equation (6)] are related only to cash sales for all sample and benchmark firms included in the regression.
explain substantially more cross-sectional variation in accruals than equivalent linear models. This has important implications for assessing the power of tests (and Type I error rates) in detecting earnings management as we demonstrate below.

Ball and Shivakumar (2006) note that one reason why transitory operating cash flows occur is because firms’ operating activities cause working capital items like inventory, receivables, and payables to vary over time. Working capital accruals adjust operating cash flow to produce an earnings number that is less noisy in measuring periodic performance and more efficient for contracting with lenders, managers, and others. Quarterly CFO measures are particularly noisy for businesses with strong seasonality. Quarterly accruals for inventories, receivables, and payables represent non-discretionary adjustments to reduce the transitory fluctuations in CFO that naturally occur in these seasonal businesses. Thus, adding contemporaneous CFO in discretionary accruals models greatly enhances standard Jones-type models’ ability to capture the true dynamics of the accrual process in quarterly settings (Jeter and Shivakumar, 1999).

Ball and Shivakumar (2006) note that another way that accrual accounting functions is to provide recognition of unrealized gains and losses. Timely gain and loss recognition occurs around the time of revision in expectations about future cash flows, which likely occurs prior to the actual realization of the cash flows, thus requiring an accrual. Because the recognition of gains and losses is asymmetric (Basu, 1997), the relation between accruals and cash flows cannot be linear. This implies that standard linear forms of Jones-type models like those examined above could be misspecified for the purpose of estimating discretionary accruals. Accordingly, we adopt the Ball and Shivakumar’s proposed adjustments to the standard Jones model to capture the noise reduction and asymmetric loss recognition properties of accruals in quarterly settings as follows:

\[ WCA_{i,t} = \beta_0 + \beta_1 Q_{1,i,t} + \beta_2 Q_{2,i,t} + \beta_3 Q_{3,i,t} + \beta_4 Q_{4,i,t} + \beta_5 \Delta SALES_{i,t} + \beta_6 WCA_{i,t-4} + \beta_7 CFO_{i,t} + \beta_8 DCFO_{i,t} + \beta_9 DCFO \times CFO_{i,t} + \varepsilon_{i,t} \]  

(7)

where \( CFO_{i,t} \) is operating cash flows for firm \( i \) in quarter \( t \) and \( DCFO_{i,t} \) is a dummy variable set to 1 if \( CFO_{i,t} < 0 \) and zero otherwise. All other variables are as previously defined in equation (5). Henceforth,
we refer to this specification as Jones + CFO. We supplement the modified-Jones model in a similar fashion and refer to this specification as Mod-Jones(C) + CFO.

Following Ball and Shivakumar (2006), models with CFO terms are estimated by pooling all firm-quarters for each industry. This is unlike models without CFO terms that are estimated for each industry-quarter. The difference arises due to a greater number of terms in the former case and the limited number of observations within many industry-quarters.

2.4 Adjustments for performance and firm growth

The residuals from the above alternative specifications adjusted for like residuals from firms matched on ROA and/or sales growth (SG) form the basis for our subsequent specification and power of tests. For ROA adjustment, we choose the matching firm that is from the same two-digit industry with the closest ROA during quarter $t-4$. For ROA + SG adjustment we arrange all same-industry firms during quarter $t-4$ into five ROA quintiles and choose the matching firm that has the closest SG from quarter $t-4$ to $t$ in the relevant quintile. We calculate ROA as the net income divided by total assets, and SG as the sales during quarter $t$ divided by sales during quarter $t-4$ minus one. Only employee growth is calculated using Compustat annual data, and it is over a one-year period ending the fiscal year before the current quarter $t$. All accrual measures and partitioning variables are winsorized at the 1% and 99% levels.

Virtually all of the studies that test for earnings management that are detailed in Appendix 1 make no explicit adjustment for accruals’ role in noise reduction and timely loss recognition. Thus, we begin by calculating the Jones and Mod-Jones(C) model abnormal accruals from equations (5) and (6) without adjustment for CFO. We then performance adjust these discretionary accrual estimates by subtracting the discretionary accruals of the ROA matching firm as described above. Finally, we calculate Jones model and Mod-Jones(C) model discretionary accruals and adjust for both performance (ROA) and sales growth (SG) by subtracting the Jones or Mod-Jones(C) model residuals of the ROA + SG matched firm from the same model residuals of the treatment firm. This provides six discretionary accrual estimates: (1) Jones model, (2) Jones model with ROA matching, (3) Jones model with ROA + SG matching, (4) Mod-
Jones(C) model, (5) Mod-Jones(C) model with ROA matching, and (6) Mod-Jones(C) model with ROA + SG matching.

Table 1 provides summary results for tests of upward earnings management for each of these six discretionary accrual estimates around the three events-related partitioning variables enumerated above—stock splits, SEOs, and stock acquisitions. Each cell reports the average discretionary accrual estimate stated as a percentage of the beginning-of-quarter total assets and related t-statistic. As shown, both baseline models [Jones and Mod-Jones(C)] yield highly significant positive abnormal accruals for all three events, with values ranging from 0.188% to 0.616% of total assets. ROA matching generally reduces the average abnormal accrual, but for stock splits and SEOs the magnitudes remain highly significant for both models. Matching on both ROA and SG results in much lower abnormal accrual estimates for the stock split and stock acquisitions samples and none of the mean values are significantly different from zero. Only the SEO sample produces ROA + SG matched abnormal accruals that are significantly positive for both model specifications. Thus, consistent with the prior findings of Teoh, Welch, and Wong (1998) and Rangan (1998), there does appear to be upward earnings management associated with SEOs, but the degree of upward management appears to be considerably less than previously documented (especially when using the Mod-Jones(C) model).

[Insert Table 1 here]

Across all three samples, the bias in test results is most severe for the Mod-Jones(C) and Mod-Jones(C) + ROA matching estimates, which are the most popular models used in prior research (see Appendix 1.1). For the stock split and stock acquisition samples, the bias for these two discretionary accruals models is particularly acute. The bottom panel of Table 1 shows why. As shown there and as demonstrated in Figure 1, these two samples are heavily populated with high growth firms. The average SG (EG) decile rank is 7.60 (7.79) for the stock acquisition sample and is 7.25 (7.10) for the stock split

---

9 We do not report a parallel experiment with partitioning variables of stock-based compensation and abnormal insider selling. This is due to the continuous nature of these variables and the associated difficulty in identifying firm-years clearly belonging to “treatment” or “control” samples (unlike in the case of events). Later we report regression tests in which the accrual measures are related to these continuous variables to illustrate the same point.
sample. Thus, studies that fail to control for growth when testing for earnings management in these settings are likely to reject the null hypothesis of no earnings management, when, in fact, the null is true.

3. Discretionary accrual measures in the aggregate sample and across ROA and SG deciles

The previous section provides evidence on Type I error rates in identified samples where the partitioning variable is highly correlated with firm growth. In this section we demonstrate how alternative discretionary accrual measures vary across ROA and SG deciles for a broad sample of firms described below to provide a sense of the potential bias in more general settings.

3.1 A comprehensive sample of firm-quarters

Most of our tests in this paper start with a comprehensive sample of 203,090 Compustat firm-quarters that span 1991-Q1 to 2007-Q4. We require that the relevant data to calculate the accrual measures used in this study and the partitioning variables ROA and sales growth (SG) are available. Following Hribar and Collins (2002), we calculate current accruals from the cash flow statement as

\[ \text{ACC} = \text{CHGAR} + \text{CHGINV} + \text{CHGAP} + \text{CHGTAX} + \text{CHGOTH} \]

The bracketed quantities in this expression represent the changes in accounts receivable, inventories, accounts payable, taxes payable, and other items.\(^{10,11}\) We undo the year-to-date nature of these quarterly cash flow statement items and compute the quantities for the quarter under consideration. We additionally require that: (1) Total assets exceed $10 million in 2007 dollars; (2) The firm is not in the financial industry (which excludes two-digit SIC codes between 60 and 69); (3) The CRSP share code is 10 or 11 (which excludes ADRs, REITs, units, certificates, and trusts); (4) There are at least 20 firms in the included two-digit SIC code during a given calendar quarter; and (5) None of the accrual measures (normalized by total assets) exceeds one.

\(^{10}\) Notice a positive (negative) value of \text{CHGAR} and \text{CHGINV} represents a decrease (increase) in accounts receivable and inventories, while a positive (negative) value of \text{CHGAP}, \text{CHGTAX}, and \text{CHGOTH} represents an increase (decrease) in accounts payable, taxes payable, and other items. These variables carry names of RECCHY, INVCHY, APALCHY, TAXCHY, and AOLOCHY in the current version of Compustat. We recode missing values of RECCHY, INVCHY, APALCHY, and TAXCHY as zero if there is a nonmissing value of AOLOCHY. Conversely, if AOLOCHY is missing but the other items are not missing, then we recode AOLOCHY as zero. In other tests, we obtain CFO by undoing the year-to-date nature of the Compustat variable OANCFY.

\(^{11}\) Unlike the other four items, \text{CHGOTH} is not all current accruals. It includes current items such as deferred revenues and expenses, but can also include gains (losses) on sales of fixed assets, asset impairment charges, foreign currency translation gains (losses), and restructuring charges. We include this item as part of current accruals because often items missing values of other items may be included here. (See previous footnote.)
3.2 Distributional statistics for accruals in the aggregate sample

The first three columns of Table 2 show the distributional statistics of various accrual measures for the aggregate sample. Current accruals calculated from the statement of cash flows scaled by lagged assets have mean and median values of 0.43% and 0.32%. The positive mean and median values are consistent with positive firm growth over time. Current accruals also show considerable cross-sectional variation, with a standard deviation of 4.54%. Some of this variation is explained by Jones and Mod-Jones(C) models, but a large part remains unexplained. The residual accruals from the Jones and Mod-Jones(C) models have standard deviation of 3.69% and 3.70%, respectively, which average 83% of the 4.54% standard deviation of raw accruals. By construction, the residual accruals from the Jones and Mod-Jones(C) models have mean values close to zero and a symmetric distribution around zero (unlike raw accruals). Matching on either ROA or SG increase the cross-sectional standard deviation of the resultant accrual difference measures by a factor of about $\sqrt{2}$. Finally, although not shown in Table 2, the average adjusted-$R^2$ of Jones and Mod-Jones(C) model regressions in equations (5) and (6) is on the order of 0.26.

The next six columns of Table 2 provide distributional statistics for the high (low) sales growth quintiles of firm-quarter observations for the six alternative model specifications. Breaking the overall sample down in this way clearly demonstrates the systematic bias that results in abnormal accrual estimates across all four models that fail to control for firm growth. The high sales growth quintile exhibits positive bias in abnormal accrual estimates ranging from 0.33% to 0.78% of total assets when one fails to match on SG. Conversely, the low sales growth quintile exhibits negative bias in abnormal accruals with estimates ranging from -0.45% to -0.88% of total assets. The bias is most severe for the Mod-Jones(C) model, which is the most popular model found in the literature (see Appendix A1.1). Importantly, both models with ROA + SG matching exhibit no systematic bias in abnormal accruals.

---

12 This is explained as follows. Suppose the Jones model (or modified-Jones model) residuals for sample firm and matching firm are denoted by $\epsilon_{it}$ and $\epsilon_{it,m}$. The matching procedure calculates discretionary accruals as $\epsilon_{it} - \epsilon_{it,m}$. In a random sample, on average, the standard deviation of the two residuals are approximately equal, so the standard deviation of the difference can be written as the standard deviation of either term multiplied by $\sqrt{2(1 - \rho)}$, where $\rho$ is the correlation between the two residuals. The typical value of $\rho$ is quite small (see Table 3).
estimates for either the high or low growth quintiles. The average abnormal accruals range from -0.03% to 0.02% (-0.08% to -0.04%) of total assets for the high (low) growth quintile.

[Insert Table 2 here]

3.3 Correlations between discretionary accrual measures, ROA, and SG

Table 3 shows the Spearman correlation matrix for various accrual measures, ROA, and SG. The correlations between the different accrual measures are quite high, as expected. More importantly, the correlation between ROA and SG is rather low at 0.03. Thus, ROA matching is unlikely to control for the effects of firm growth on discretionary accrual estimates. This is demonstrated by the significantly positive correlations between ROA-adjusted discretionary accrual estimates and SG, which are as high as 0.15 for the Mod-Jones(C) model. Matching by ROA or ROA + SG results in zero correlation between various discretionary accruals measures and ROA, as expected.

3.4 Discretionary accruals across ROA and SG deciles

We next examine various discretionary accrual estimates across stratified random samples formed by ROA and SG deciles. We present several three-dimensional plots where the horizontal axes are ROA and SG decile ranks and the vertical axis is one of the discretionary accrual estimates. Panel A of Figure 3 plots Jones model residuals on the vertical axis, Panel B plots Jones model residuals with ROA matching, and Panel C plots Jones model residuals with ROA + SG matching. Panels D, E, and F plot the corresponding measures for the Mod-Jones(C) model. The grid points in each plot represent the median accrual values within the intersections of ROA and SG deciles.13

Consistent with the correlation analysis in Table 3, Panel A of Figure 3 shows that the median Jones model residuals vary considerably more across SG deciles than across ROA deciles. This pattern persists with ROA matching in Panel B, and it is exacerbated with Mod-Jones(C) model discretionary accrual estimates shown in Panels D and E. Averaged across all ROA deciles, the Jones model residuals increase from -0.34% of lagged assets in the lowest SG decile to 0.24% of lagged assets in the highest SG decile. The corresponding values are -0.38% and 0.27% for Jones model with ROA matching, -0.74% and

13 Figure 3 plots were made in Matlab and involved some use of cubic smoothing splines. This does not alter any of the inferences. Unsmoothed plots are available from the authors on request.
0.60% for Mod-Jones(C) model, and -0.83% and 0.76% for Mod-Jones(C) model with ROA matching. Thus, across the aggregate sample of Compustat firm-quarters our results show that Mod-Jones(C) model discretionary accrual estimates with ROA matching will produce the most biased results with quarterly data if the event or firm characteristic hypothesized to give rise to earnings management is positively or negatively correlated with sales growth.

In contrast to the large variation in discretionary accrual estimates across SG deciles, Panels A and D show only minor variation in Jones model and Mod-Jones(C) model residuals across ROA deciles. Averaged across all SG deciles, the median accruals in the top and bottom ROA deciles equal 0.13% and 0.03% with Jones model, and 0.03% and 0.16% with Mod-Jones(C) model. These results differ from those documented by Kothari, Leone, and Wasley (2005), who examine total accruals calculated using the balance sheet method of estimating accruals with annual data. They find that Jones model and modified-Jones model residuals differ considerably between lowest and highest E/P quartile partitions (-3.23% and 0.28% in the first case and -3.85% and 0.31% in the second case – see their Table 1), so they recommend ROA matching. However, their recommendation does not automatically apply to discretionary accrual models using quarterly data. In summary, our analysis shows that the residuals from Jones and Mod-Jones(C) models using quarterly data vary considerably more with firm growth (SG) than with performance (ROA).

[Insert Figure 3 here]

To control for the effects of firm growth on accruals, we adjust Jones model or Mod-Jones(C) model residuals by subtracting the residual accruals of same-industry firms matched by both ROA and SG as described in Section 2.4 and Table 2. The corresponding results are shown in Panels C and F of Figure 3. As expected, the resulting measures do not show any systematic variation across either ROA or SG deciles in the aggregate sample. It follows that in event studies or cross-sectional studies, the use of Jones-type models with ROA + SG matching corrects for the nondiscretionary part of accruals associated with both performance and growth.
4. Specification tests (Type I errors) of discretionary accrual models using quarterly data

4.1 Results for highly concentrated samples of high (low) growth firms

In this section we replicate a typical research design employed to detect earnings management around specific corporate events in order to test the specification (Type I error rates) of six discretionary accrual measures: (1) Jones model, (2) Jones with ROA matching, (3) Jones with ROA + SG matching, (4) Mod-Jones(C), (5) Mod-Jones(C) with ROA matching, and (6) Mod-Jones(C) with ROA + SG matching. Instead of real events we select a random sample of 200 observations taken from either the aggregate sample of Compustat firm-quarters or from the bottom (Low) or top (High) quintile of firm-quarters ranked by SG, EG, or ROA. The choice of 200 observations is somewhat arbitrary. While it is less than the typical sample size in most event studies, having all these observations from the top or bottom quintile of firm characteristics makes it comparable to a bigger sample with more modest concentration in extreme quintiles. Later we repeat the specification tests for these accrual measures with larger samples, but a lower concentration of observations in extreme deciles.

We repeat the above sampling procedure 250 times with replacement. Since the firm-quarters are selected at random, there is no reason to believe systematic earnings management is present in these samples. Thus, the null hypothesis of no earnings management is assumed to be true. Using an $\alpha$-level of 5%, we measure the percentage of the 250 trials that the null hypothesis of no earnings management (zero abnormal accruals) is rejected in favor of the alternate hypothesis of either positive or negative abnormal accruals using a $t$-test of means. With 250 replications, there is a 95% probability that the measured rejection rate will lie between 2.4% and 8.0% if the discretionary accrual measure is not misspecified and the null is true.

Table 4 presents the simulation results using accruals taken from the cash flow statement. Panel A (B) shows the rejection rates against the alternative hypothesis that discretionary or abnormal accruals are negative (positive). The first column provides rejection frequencies for samples drawn from the aggregate sample, the next two columns present results for samples drawn from the bottom and top quintile of firm-quarters ranked on SG, the following two columns present results for samples drawn from the bottom and
top quintile of firm-quarters ranked on EG, and the last two columns present results for samples drawn from the bottom and top quintile of firm-quarters ranked on ROA. The key findings are as follows.

[Insert Table 4 here]

1. For samples drawn from the aggregate set of Compustat firm-quarters, in 11 out of 12 cases the rejection rates lie between the bounds of 2.4% and 8.0%. The one violation is for the Mod-Jones(C) model without matching firm adjustment. Notice one modest violation in 12 experiments (a rejection rate of 2.0% compared to lower bound of 2.4%) may be expected by random chance when the bounds are set with 95% confidence level. Overall, we interpret the evidence with the aggregate set of Compustat firm-quarters as fairly neutral. This is not surprising because roughly half of each draw of 200 observations should have ROA and SG values above or below the corresponding mean value for all same-industry firms. Thus, any performance or growth-related bias in discretionary accruals estimates will likely cancel out in samples selected across ROA or SG deciles. Overall, these tests serve as an important validation-check of our simulation procedures and provide a benchmark for comparing samples drawn from top and bottom quintiles of ROA or SG distributions.

2. Type I error occurs when the null hypothesis is wrongly rejected as indicated by a rejection rate higher than 8.0%. When the null hypothesis of no earnings management is true, a test is misspecified when the rejection rate is lower than 2.4% or higher than 8.0%. Clearly, rejection rates higher than 8.0% are the more serious case as the researcher erroneously concludes in favor of earnings management when there is none. As shown in Table 4, Jones and Mod-Jones(C) models suffer from large Type I errors in samples selected from low and high SG quintiles of Compustat firm-quarters. In low SG sample the researcher erroneously concludes in favor of downward earnings management, and in high SG sample erroneously concludes in favor of upward earnings management. Looking across models, the Mod-Jones(C) model is associated with the highest Type I errors, with rejection rates as high as 89.2% (86.0%) in the low (high) SG partitions. The rejection rates are the lowest for Jones model, 41.2% (33.6%) in the low (high) SG partition, still very high in absolute terms. This

---

14 We draw the following analogy. Suppose a binomial variable equals Y with probability 0.95 and N with probability 0.05. Then the probability that all 12 draws of this variable will have the value Y equals $0.95^{12} = 0.54$. Thus, there is a 46% chance that at least one of the draws will have the value N.
general pattern of rejection rates for Jones and Mod-Jones(C) model results hold for all subsequent
tests in this paper.

3. ROA matching moderates the over-rejection rates, but they remain high in absolute terms. When the
alternative hypothesis is negative discretionary accruals, the Jones and Mod-Jones(C) models with
ROA matching give rejection rates of 30.0% and 66.8% in the bottom SG quintile. Similarly, when
the alternative hypothesis is positive discretionary accruals, the two models with ROA matching give
rejection rates of 21.2% and 58.0% in the top SG quintile. Thus, similar to the evidence based on the
event studies in Table 1 and the three-dimensional plots in Figure 3, tests of earnings management
based on performance-adjusted (ROA matched) Jones and Mod-Jones(C) models are highly
misspecified when samples have extreme growth characteristics.\footnote{Notice ROA matching slightly increases the magnitudes of modified-Jones model residuals in high and low SG
partitions in Figure 3 (see Panels B and E), but it decreases the rejection rates in both panels of Table 4. This
highlights the limitation of examining only the rejection rates in simulation models. The rejection rates depend on
the magnitudes of biases as well as their standard deviations. Table 2 shows that any kind of matching increases the
cross-sectional standard deviation of modified-Jones model residuals by a factor of about $\sqrt{2}$. Thus, ROA-matching
depends on the magnitudes of biases as well as their standard deviations. Table 2 shows that any kind of matching increases the
rejection rates in Table 4 even though it increases (or leaves unchanged) the magnitudes of biases in
Figure 3. We later revisit this issue by providing quantitative estimates of biases within select partitions in Figure 8.

4. As one might expect, both Jones and Mod-Jones(C) models with ROA + SG matching yield
reasonably well-specified results in all SG and ROA partitions under either alternative hypothesis.
Rejections rates are in the range of 2.4% to 8.0% for samples drawn from extreme SG quintiles and in
the range of 4.4% to 8.0% for samples drawn from extreme ROA quintiles. In none of the 16 tests
involving extreme SG or extreme ROA samples does the rejection rate lie outside the 5% bounds of
2.4% to 8.0%. We find rejection rates in the range of 2.0% to 12.4% in eight tests involving extreme
employee growth (EG) quintiles with ROA + SG matching, compared to 95% confidence bounds of
2.4% to 8.0%, indicating slight misspecification within these partitions.\footnote{When we match on EG, all rejection rates are within the 95% confidence bounds of 2.4% to 8.0%.

Overall, we conclude that
ROA + SG matching is critical to obtaining unbiased results from Jones-type discretionary accrual
models using quarterly data in concentrated samples of high (low) growth firms.
4.2 Results for samples with varying proportions of high growth firms

The specification test results in Table 4 show the Type I error rates when a relatively small sample (200 observations) is drawn entirely from firm-quarters in the high (low) growth quintiles. Essentially, this assumes that the partitioning variable used to identify cases of earnings management has a 100% overlap with firms with extreme growth. In general, however, the partitioning variable in most studies of earnings management rarely coincides perfectly with firm growth. Rather, the samples often are only partially over-represented by high (low) growth firms. Thus, the degree to which firm growth may confound test results varies depending on the event chosen. For example, the histograms in Figure 1 suggest that samples used in studies that test for earnings management around stock acquisitions are more likely to be over-represented by high growth firms than are studies that test for earnings management around stock splits.

To assess how varying the proportion of high growth firms in a sample can impact Type I error rates in sample sizes more commonly used in testing for earnings management, we conduct the following simulation. We begin by taking 250 stratified random samples (with replacement) of 1,000 firm-quarters, and for each of the 250 trials we vary the proportion of firms drawn from the high sales growth quintile.\(^{17}\) We estimate the mean abnormal accrual ($\mu$) for each sample conditional on each of the six models shown in Table 4. Using a significance level of 5%, we measure the percentage of the 250 trials that the null hypothesis of $\mu = 0$ is rejected in favor of the alternative hypothesis of $\mu > 0$ using a one-tailed $t$-test of the mean. Figure 4 shows the rejection rates (Type I error rates) across samples with an increasing proportion of firms-quarters drawn from the high SG quintile.\(^{18}\) The origin on the $x$-axis (0) represents random samples of 1,000 firm-quarters with no excess proportion of firm-quarters drawn from the high SG quintile. Thus, 20% of the firm-quarters in each of these random samples would be expected to come from the high SG quintile. The 30% point on the $x$-axis represents samples that have 2.5 times the normal representation of high SG quintile firms. Thus, these samples would have roughly 50% of the 1,000

\(^{17}\) Note that 1,000 firm-quarters is roughly comparable to sample sizes used in prior studies to test for earnings management around stock splits, SEOs and stock acquisitions. See Figure 1.

\(^{18}\) The plots are virtually identical when varying the proportion of the sample from the low SG quintile and testing the alternative hypothesis of $\mu < 0$. 
observations coming from the high SG quintile while the remaining 50% are evenly distributed across the lower four SG quintiles. Note that samples comprised of 50% high quintile SG firms roughly mirror the proportion of high SG firms found in SEO and stock acquisition samples reported in Figure 1.

The results show that all six models yield rejection rates of right around 5% when there is no excess representation of high SG firm-quarters in the sample (i.e., 0 point on x-axis). As the excess proportion of high SG firm-quarters increases, Type I error rates for the Jones and the Mod-Jones(C) models, with and without adjustment for ROA, increase rather dramatically. For example, in samples with 30% excess representation of firm-quarters from the high SG quintile (i.e., samples with 50% of the observations coming from the high SG quintile), a true null hypothesis of no earnings management ($\mu = 0$) is rejected roughly 28% and 73% of the time when using the Jones and Mod-Jones(C) models, respectively. Adjusting these models by matching on ROA reduces the Type I error rates, but they are still excessive at 13% and 48%, respectively. In comparison, Jones and Mod-Jones(C) models with ROA + SG matching yield well specified tests (rejection rates hover around 5%) across all levels of excess representation of high SG firms in the sample. The overall results show that with realistic representation of firm-quarters across SG quintiles that mirror the distribution of common events and with realistic sample sizes of 1,000 observations the Jones and Mod-Jones(C) models without ROA + SG adjustment are considerably misspecified.

[Insert Figure 4 here]

Figure 5 repeats the analyses in Figure 4 for the versions of the Jones and Mod-Jones(C) models that adjust for accruals’ role in reducing the noise in quarterly operating cash flows due to seasonality and for non-linearities due to asymmetric timeliness of loss recognition (Ball and Shivakumar, 2006). Although not widely adopted in the literature, we present the results for these models for completeness and to demonstrate that the increased power that results from using these models (demonstrated in Section 5 below) potentially comes at the expense of high Type I error rates. Specifically, Figure 5 shows that rejection frequencies for samples with excess representation of high growth firms are well above the 5%
nominal rate for all versions of the models that fail to control for firm growth. However, as before, the Jones and the Mod-Jones(C) models with CFO adjustment yield well specified tests (rejection frequencies that hover around 5%) when matching on both ROA and SG.

[Insert Figure 5 here]

Finally, although not tabulated or plotted, we replicated all our specification tests of Table 4 and Figures 4 and 5 starting with raw accruals calculated using the balance sheet method for calculating accruals. Hribar and Collins (2002) document that this method is subject to severe biases when there are nonoperating events such as acquisitions, divestitures, and foreign currency translations during the event period. Despite their warning, this method continues to be used in many studies. We find that without matching for firm growth the various discretionary accrual measures reported in this paper are subject to even greater misspecification when raw accruals are calculated using the balance sheet method.

5. Power of tests based on alternative discretionary accrual measures

5.1 Power of tests in random samples with no over-representation of high growth firms

In this section we address two distinct questions: (1) How does controlling for firm growth affect the power of tests to detect earnings management; and (2) How does adjustment of Jones-type models for accruals’ noise reduction and timely loss recognition roles affect the power of tests? Strictly speaking, power tests should be carried out on sampling distributions where all competing measures are known to have similar Type I errors. From Table 4 and Figure 4, we know that sampling distribution drawn from the aggregate sample of Compustat firm-quarters is the only one we have examined that meets this requirement. So we report power test results based on samples drawn from this aggregate sample. We artificially add a seed of 0.25% of lagged total assets to raw working capital accruals of each randomly picked firm-quarter and compute the following four discretionary accrual measures without CFO adjustment: Jones with ROA matching, Jones with ROA + SG matching, Mod-Jones(C) with ROA matching, and Mod-Jones(C) with ROA + SG matching. We next repeat the process using versions of

---

20 Under this method, the raw accruals are calculated as $\Delta CA - \Delta CL - \Delta CASH + \Delta STDEBT$, where $\Delta CA$ is the change in current assets during the quarter, $\Delta CL$ is the change in current liabilities, $\Delta CASH$ in the change in cash and cash equivalents, and $\Delta STDEBT$ is the current maturities of long-term debt and other short-term debt included in current liabilities.
these models with CFO adjustment (i.e., versions of models that adjust for accruals’ noise reduction role and asymmetric timely loss recognition) as outlined above. As before, our inferences are based on a $t$-test for mean discretionary accruals and a one-tailed significance level of 5%. For brevity, we report only tests of the null hypothesis of zero discretionary accruals against the alternative hypothesis of positive discretionary accruals, which is the more common alternative hypothesis in the finance and accounting literature. For any given seed level, the probability of rejecting the null hypothesis depends on the sample size. Therefore, we run 250 simulations for each measure with sample sizes ranging between 200 and 2000 firm-quarters in increments of 200 observations. The results are summarized in Figure 6.

[Insert Figure 6 here]

First, keeping aside the issue of CFO adjustment, there is no reason to believe that any of the four competing Jones-type model specifications considered in this figure is more powerful than the others when samples are drawn from the broad cross-section of Compustat firm-quarters. The key determinant of power is the cross-sectional standard deviation of an accrual measure. Because each measure calculates the difference between Jones or Mod-Jones(C) model residuals of a sample firm and a matching firm from the same industry, and because Table 2 shows that Jones and Mod-Jones(C) model residuals have nearly identical standard deviations, all four measures should have comparable power. This prediction is confirmed in the lower grouping of plots depicted in Figure 6. The differences in power across the alternative models are minor and may be attributed to random simulation errors. Averaged across all models without CFO adjustments, sample sizes of 1200, 1600, and 2000 firm-quarters lead to average rejection rates of 42%, 51%, and 61%, respectively, across these varying sample sizes with a seed level of +0.25% of lagged assets.

Next, we address the issue of CFO adjustment. The upper grouping of plots in Figure 6 shows that the same four models but with CFO adjustments yield average rejection rates of 58%, 70%, and 78% across samples sizes of 1200, 1600, and 2000 firm-quarters, respectively. Thus, samples of 1200 to 2000 or more observations, which are common in many research settings, have roughly 60% to 80% probability of detecting earnings manipulation of 0.25% of lagged assets with quarterly data when CFO adjustments are made to control for the accruals’ noise reduction and timely loss recognition roles. As
before, there is no significant difference across models depending on Jones versus Mod-Jones(C) specifications or ROA versus ROA + SG matching. The main factor that improves the model power is the inclusion of CFO adjustments for accruals’ noise reduction role and non-linearities due to asymmetric timely loss recognition as suggested by Jeter and Shivakumar (1999) and Ball and Shivakumar (2006).

5.2 Sample distribution and mean versus median tests

All tests of discretionary accrual models reported so far have used a t-test of mean abnormal accruals or accrual differences. This test is commonly used in the literature, and it underlies all cross-sectional regressions that include discretionary accruals as the dependent variable of interest. It is also the primary test employed by Kothari, Leone, and Wasley (2005) in their examination of discretionary accrual models. The t-test is a parametric test, and generally speaking parametric tests are more powerful than nonparametric tests when the underlying variable is normally distributed. However, this generalization breaks down when the underlying variable is not normally distributed. In such cases nonparametric tests such as the Wilcoxon signed-rank test of medians can be more powerful than the t-test for mean (Blair and Higgins, 1985).

We examine the skewness and kurtosis of different discretionary accrual measures using the aggregate sample of 203,090 firm-quarters to measure the departures from normality. For the Jones model without (with) CFO adjustment the skewness equals -0.07 (-1.03) and the excess kurtosis (which subtracts three from the scaled fourth moment of distribution) equals 2.06 (5.33). The greater departure from normality with CFO adjustment indicates that this procedure works better for some industries than others. The skewness of both measures is corrected when we subtract the corresponding value for an

---

21 Skewness of a variable with values $x_i, i = 1, 2, ..., n$ is defined as $\sum_{i=1}^{n} \frac{(x_i - \mu)^3}{n\sigma^3}$, where $\mu$ is the population mean and $\sigma$ is the standard deviation. Kurtosis is defined as $\sum_{i=1}^{n} \frac{(x_i - \mu)^4}{n\sigma^4} - 3$. (It is sometimes known as excess kurtosis due to “minus 3”.) For a normal distribution, skewness and kurtosis should both equal zero, while winsorizing should reduce the kurtosis to slightly less than zero. A negative (positive) skewness indicates a distribution tilted to the left (right), and a negative (positive) kurtosis indicates thinner (thicker) tails than a normal distribution. Given the very large number of observations in the aggregate sample, tests of normality always reject it. So we focus on the magnitudes of skewness and kurtosis to measure the economic magnitude of departure from normality.

22 For example, the CFO adjustment should work better for businesses with greater seasonality, higher discretion over cash versus credit sales, and frequent loss recognition. In support of this conjecture we first find that, averaged across 38 industries (i.e., two-digit SIC codes), equation (5) gives an adjusted-$R^2$ of 0.20 for Jones model without CFO adjustment and equation (7) gives an adjusted-$R^2$ of 0.42 for Jones model with CFO adjustment. The incremental $R^2$ is 0.22. (The difference between average adjusted-$R^2$ values of 0.20 in this section and 0.26 in
ROA or ROA + SG matching firm, but the kurtosis remains significantly different from zero. With ROA matching the kurtosis of Jones model discretionary accruals without (with) CFO adjustment equals 1.13 (3.50). With ROA + SG matching it equals 1.17 (3.51). The evidence is quite similar when we examine the variants of the Mod-Jones(C) model.

Panel B of Figure 6 shows the power of different discretionary accrual models to detect artificially induced earnings management of 0.25%, but using the Wilcoxon signed rank test for median instead of the t-test for mean in Panel A. We find that the median test is more powerful for all discretionary accrual models. However, the increase in power is greater for models that include CFO controls than for models that do not include CFO controls. This is not surprising in view of the above evidence that CFO adjustment results in greater departures from normality for all discretionary accrual measures. More specifically, without CFO controls, the Jones and Mod-Jones(C) models, with ROA or ROA + SG matching, have rejection rates of 54%, 64%, and 74% in sample sizes of 1200, 1600, and 2000 firm-quarters when we use the Wilcoxon signed rank test for median. These rejection rates are 12% to 13% higher than corresponding rejection rates with the t-test for mean. In comparison, with CFO controls, the same models have rejection rates of 78%, 89%, and 95% with the median test, which are 17% to 20% higher than for the mean test (despite the already higher baseline rejection rates with t-test for mean). In absolute terms, including CFO terms in Jones-type models combined with ROA + SG matching enables a researcher to detect earnings management of 0.25% of total assets with a probability hovering around 95% when the sample size is around 2000 observations.23

---

23 Section 3.2 can be attributed to running the regressions by industry versus industry-quarter. The latter gives a better fit.) There is considerable variation in incremental $R^2$ across industries. The three lowest incremental $R^2$ values are 0.09, 0.10, and 0.11 (for chemical and allied products, communications, and eating and drinking establishments), and the three highest incremental $R^2$ values are 0.45, 0.42, and 0.35 (for wholesale trade – durable goods, motion pictures, and automotive dealers and gasoline service stations). These differences can create additional heterogeneity and departures from normality in CFO-adjusted discretionary accruals in samples drawn from the aggregate dataset of Compustat firm-quarters.

23 In untabulated results, we also find that the median tests is well-specified in the presence of ROA + SG matching.
5.3 Simulations with over-representation of high growth firms and varying amounts of earnings management

We next examine rejection rates for samples with excess representation of firms from the high sales growth quintile. In these simulations, we vary the seed from -0.50% to +0.50% of lagged total assets in increments of 0.125%. The rejection rates are shown in Figure 7 for the alternative hypothesis of positive abnormal accruals for 250 replications of sample sizes of 1000 firm-quarters of which 500 are drawn from the top SG quintile and the remaining 500 observations are drawn randomly from the remaining SG quintiles. Recall that this incidence of over-representation of high SG firm-quarters is roughly similar to what we find for the SEO and stock acquisition samples depicted in Figure 1.

Panel A of Figure 7 shows the rejection rates as a function of seed size for the following models without CFO adjustment: Jones, Jones with ROA matching, and Jones with ROA + SG matching. Panel B shows the corresponding results for Mod-Jones(C) model without CFO adjustment. Finally, Panels C and D show the rejection rates for the Jones and Mod-Jones(C) models but with CFO adjustment for accruals’ noise reduction role and asymmetric loss recognition across varying earnings management seeds.

[Insert Figure 7 here]

In stratified random samples considered in Figure 7 (i.e., samples with over-representation of high SG firms) the dual issues of model power and specification cannot be separated. Consider, for example, the rejection rates for the Mod-Jones(C) model without CFO adjustment but with ROA matching (middle line in Panel B). Even when earnings management is not present (seed of 0.00%), this popular way of measuring abnormal accruals will reject the null hypothesis of zero discretionary accruals 48% of the time. It is therefore not surprising that this model seemingly detects earnings management of 0.25% of total assets with a probability of 91% in samples over-represented by high growth firms. Panel D shows that the rejection rates for Mod-Jones(C) models with CFO adjustment. With no earnings management (seed of 0.00%), the rejection rates increase to 70% without ROA matching and to 58% with ROA matching. Indeed, even when negative earnings management is seeded in the data, these models

---

24 For parsimony, we do not show results when the samples are over-represented by firms from the low SG quintile and the alternate hypothesis is downward earnings management. The findings are similar to those reported below.
with CFO adjustment tend to reject the null hypothesis of no earnings management in favor of the alternative hypothesis of positive earnings management. Clearly, rejection rates in excess of 90% with a seed size of 0.25% for commonly used Mod-Jones(C) specifications (both without and with CFO adjustments) do not represent model power but model power combined with model misspecification. The misspecification is reduced to some extent by using versions of models that adjust for performance (ROA matching), but remains high. The solid lines in the various panels of Figure 7 demonstrate that the critical ingredient for unbiased inference is the SG adjustment. All four panels of Figure 7 show that the rejection rates remain close to the theoretical value of 5% when one adjusts for ROA + SG and there is no earnings management. This is true regardless of whether one employs the Jones or Mod-Jones(C) model and whether one does or does not adjust for CFO.

5.4 A closer look at biases in discretionary accruals estimates from alternative models when samples are over-represented by high growth firms.

While the dual issues of model power and specification cannot easily be separated by looking at rejection rates across varying levels of earnings management, one can gain a better sense of how well the alternative models perform on these two dimensions by examining the means and standard deviations of various discretionary accrual estimates compared to known amounts of seeded earnings management. Figure 8 presents the actual amount of seeded earnings management of -0.5%, -0.25%, 0.00%, +0.25%, and +0.5% of lagged assets and the average discretionary accruals estimates across the 250 simulation trials for each of these conditions for alternative models outlined above. Panels A and B of this figure correspond to the samples and measures presented in Panels A and B of Figure 7. Both Jones and Mod-Jones(C) model accruals without any matching or with ROA matching are upward biased. For example, for a seed level of -0.25% added to raw accruals, the mean discretionary accruals estimate is -0.09% for the Jones model, and -0.10% for the Jones model with ROA matching. The Jones model with ROA + SG matching yields an average discretionary accruals estimate of -0.24%, very close to the -0.25% seed. For a seed level of +0.25%, the corresponding estimates for these three models are 0.37%, 0.36%, and 0.23%.
As expected, the biases are more acute with Mod-Jones(C) model. For a seed level of -0.25%, the mean discretionary accruals estimate is actually positive at 0.09% of lagged assets for the Mod-Jones(C) model and 0.08% of lagged assets for the Mod-Jones(C) model with ROA matching. For the Mod-Jones(C) model with ROA + SG matching, the bias is much smaller with an average discretionary accruals estimate of -0.21% of lagged assets. For a seed level of +0.25% of lagged assets, the corresponding discretionary accruals estimates across these three models are 0.55%, 0.54%, and 0.25%. Thus, matching on firm growth (SG) produces discretionary accrual estimates very close to the seeded values for both Jones and Mod-Jones(C) models, while failing to control for firm growth produces upward biased estimates.

In untabulated results we look at standard deviations of discretionary accrual estimates across the 250 trials and find them to be approximately equal for all models involving any type of matching firm. Of course, not using any matching firm produces standard deviations of estimates that are smaller by a factor of $\sqrt{2}$ as discussed earlier. In untabulated results we also find that the mean discretionary accrual estimates are relatively unaffected by CFO adjustment while standard deviations decrease significantly.

[Insert Figure 8 Here]

Overall, there are three takeaways from this analysis. First, failing to control for firm growth in tests of earnings management yields false detection rates (power) when the samples under consideration include an over-representation of high (low) growth firms and the alternative hypothesis is upward (downward) earnings management. Second, adjusting Jones and Mod-Jones(C) model estimates for both performance (ROA) and growth (SG) yields (1) well specified tests under a true null of no earnings management, and (2) consistent and reliable detection rates for modest levels of earnings management in samples randomly selected across growth partitions as well as samples that are correlated with growth.

---

25 Notice that with ROA + SG matching also we notice the slight upward biases for negative seed level and slight downward biases with positive seed level. This is the result of running Jones or Mod-Jones(C) model regressions shown by Equations (5) and (6) in Section 2.3 over the combined population of sample (or treatment) firms and same-industry and same-quarter firms. This washes away some of the effect, but by equal amounts for all variations of Jones and Mod-Jones(C) models for a given seed level. Further, these biases are also proportional to the seed level and average between one-tenth and one-twentieth of the seed level. If more precise measurements are required, then we recommend running the regressions only for same-industry firms and using the regression coefficients to separate out the non-discretionary and discretionary components of sample firm accruals.
Given that firm growth is often correlated with partitioning variables in tests of earnings management, it seems clear that researchers should control for both ROA and SG going forward. Third, inclusion of CFO terms as suggested by Ball and Shivakumar (2006) greatly improves the detection power of all discretionary accrual measures estimated with quarterly data that start with Jones or Mod-Jones(C) models. This improvement in model power occurs without compromising model specification (Type I error rates) when adjustments are made for both performance (ROA) and firm growth (SG). Thus, we suggest that researchers use CFO adjustment to increase the power of their tests, especially in quarterly settings where there are likely to be transitory shocks to operating cash flows due to strong seasonality in the business environment of sample firms.

6. Does adjusting for firm growth throw the baby out with the bathwater?

The simulations presented in the preceding sections infuse earnings management seeds into aggregate working capital accruals. A natural question arises as to whether one “throws the baby out with the bathwater” if firms manage earnings through revenue manipulation and abnormal accruals are estimated using the ROA + SG matching procedure. Essentially, the concern is that if a part of the treatment firm’s sales growth is due to revenue manipulation, then one will end up choosing a control firm matched on sales growth that is too high. As a result, all or part of the earnings management that is accomplished through revenue manipulation may be negated when the matching firm’s Jones-type-model residuals are subtracted from the residuals of the corresponding sample firm. This section reports simulation results designed to address the baby and bathwater concern.\(^{26}\) We show below that the resultant bias in discretionary accrual estimates from SG matching is very modest in most realistic settings and much less serious than biases that result from failure to control for firm growth.

To provide a benchmark for the amount of revenue manipulation seeded into the data, we use descriptive statistics on annual sales growth of a sample of firms reporting restatements from 1997 to 2002 identified in a recent study by Badertscher, Collins, and Lys (2012).\(^{27}\) For firms reporting a

\(^{26}\) For an alternative approach to isolating discretionary and non-discretionary revenues in tests of earnings management see Stubben (2010).

\(^{27}\) The sample comes from a study of restatements published by the Government Accountability Office (GAO 2002). We thank Brad Badertscher for providing us with these descriptive statistics.
difference between originally reported and restated sales, the mean (median) difference in annual sales growth in their study is 5.02% (5.16%). So for the simulations in this section we introduce a revenue management seed of 5.0% of four-quarter lagged sales, $S_{i,t-4}$. For benchmarking and comparability purposes, we also investigate cases where the seed is 0.0% and 2.5% of lagged sales. It follows that the manipulated or overstated sales in period $t$ becomes $S'_{i,t} = S_{i,t} + \text{seed} \times S_{i,t-4}$ and sales growth becomes $SG' = [(S'_{i,t} + \text{seed} \times S_{i,t-4}) - S_{i,t-4}] / S_{i,t-4} = SG + \text{seed}$. Assuming that all overstated sales are on credit, the accounts receivable in turn becomes $AR'_{i,t} = AR_{i,t} + \text{seed} \times S_{i,t-4}$. In all expressions a superscript $'$ attached to any quantity denotes a manipulated value. Finally, the effect of this amount of sales overstatement $(S'_{i,t} - S_{i,t})$ on bottom-line earnings is calculated as $(1 - \tau) \cdot (S'_{i,t} - S_{i,t}) \cdot GM_{i,t}$, where $\tau$ is the marginal corporate tax rate (35%) and $GM_{i,t}$ is the average gross margin for all firms with the same 2-digit SIC code during the same quarter.\textsuperscript{28} These amounts normalized by lagged assets are presented in column (3) of Table 5 and provide the benchmark for assessing the degree of bias using alternative methods for calculating discretionary accruals that result from revenue overstatement.\textsuperscript{29,30}

Table 5 presents our test results of whether ROA + SG matching throws the baby out with the bathwater when the source of earnings management is revenue manipulation. Panel A presents variations of the Jones model with three different matching procedures, and Panel B presents the corresponding variations of the Mod-Jones(C) model. The model presented in column (4) uses regressors that include the inflated revenue amounts and matches only on ROA. This form of the model represents the most popular form used in the literature to date. The model presented in columns (5) uses regressors that include the inflated revenue amounts and matches on $ROA + SG'$, the inflated sales growth amount. The model presented in columns (6) uses regressors that include the inflated revenue amounts and matches on $ROA + SG$, the sales growth without the manipulated portion. Although this is unobservable to the researcher, it provides a useful benchmark for evaluating bias in discretionary accruals estimates that result from

\textsuperscript{28} We assume that when revenues are overstated, then the cost of sales is also overstated by $(1 - GM_{i,t}) \cdot (S'_{i,t} - S_{i,t})$.

\textsuperscript{29} As always, we normalize all quantities appearing in Jones-type models (in particular, the adjacent-quarter change in sales and accounts receivable and current accruals) by lagged assets.

\textsuperscript{30} We denote the adjacent quarter change in sales, which is the primary explanatory variable in the Jones model, as $\Delta S'_{i,t} = S'_{i,t} - S_{i,t-1}$. 

matching on inflated sales numbers, which is the difference in the bias in model (6) versus model (5) presented in column (10) of Table 5.

For all tests reported in this table we draw 1,000 samples of 1,000 firm-quarters of which 500 observations are drawn from the top SG quintile and the remaining 500 observations are drawn randomly from the remaining SG quintiles.\footnote{Whereas all tests elsewhere in this paper employ 250 replications, in this table we employ 1,000 replications. This is because the baby with the bathwater issue turns out to be of a smaller magnitude in many cases and requires more precise measurement.} Recall that this incidence of over-representation of high SG firm-quarters is roughly similar to what we find for the SEO and stock acquisition samples depicted in Figure 1. Regarding the extent of revenue manipulation, we first consider the most extreme case where 100% of the sample firms overstate sales revenue, which is presented in Subpanels A1 and B1. More realistic cases are presented in the remaining subpanels and are based on studies of restatements by the United States Government Accountability Office (GAO) issued in 2002 (2006). These studies indicate that 37.9% (20.1%) of restatements involve some type of revenue manipulation. Therefore, in Subpanels A2 and B2 (A3 and B3) we report simulations in which the source of earnings management is revenue overstatement for 40% (20%) of the observations. For the remaining 60% (80%) of observations in these subpanels, we assume that the source of earnings management is expense understatement in an amount equivalent to the accruals overstatement resulting from sales overstatement in the other 40% (20%) of the observations. Column (3) shows the magnitude of the resultant earnings effect stated as a percentage of lagged assets. The table legend shows the remaining simulation details.

6.1. Variations of Jones model (Panel A of Table 5)

Recall that the variations of Jones model effectively treat all earnings manipulation through revenues as nondiscretionary because the observed change in sales ($\Delta S'_{it}$) is the primary regressor, which effectively removes any revenue manipulation imbedded in $\Delta S'_{it}$ from the discretionary accruals estimate. Thus, the discretionary portion of accruals that is used to test for earnings management is understated (i.e., biased downward towards zero when we insert a positive seed). This will become evident in the tabled results discussed below.
Row 1 in Subpanel A1 uses a seed of 0.0%. In other words, there is no earnings management. We investigate this situation to illustrate the bias in discretionary accruals estimates that results when there is an over-representation of high growth firms in the sample and one does not control for firm growth. Note that the Jones model with ROA matching yields discretionary accruals of 0.118% of lagged assets when there is no earnings management seeded into the data. In comparison, estimates of discretionary accruals based on ROA + SG′ or ROA + SG matching procedures yield unbiased estimates of discretionary accruals (bias of less than 1 basis point – see columns 8 and 9).

Row 3 in Subpanel A1 presents results for an extreme case scenario where every observation is seeded with excess sales revenue equal to 5.0% of $S_{t,t-4}$ and, as a result, earnings are overstated by 0.316% of lagged assets. Jones model with ROA matching (column 4) gives discretionary accruals of 0.332% of lagged assets, not much different from the induced amount of earnings management, which is 0.316% of lagged assets. This result is deceiving, however, because it results from a downward bias from using an overstated $\Delta S_{t,t}'$ regressor (which means that the effects of upward revenue manipulation are removed from the discretionary accruals estimate) offset by an upward bias due to failing to control for the effects of firm growth on discretionary accrual estimates. These offsetting effects net out to an upward bias of less than 2 basis points. One should not erroneously conclude from this that Jones model with ROA matching gives unbiased estimates of discretionary accruals. It is simply a chance result of two opposite biases cancelling each other. The net result is a deceptively low total bias which is reported in Column (7).

Next we consider the Jones model with ROA + SG′ matching (column 5). This model corrects the upward bias that results from failing to control for overrepresentation of high growth firms in the sample, but introduces a downward bias due to using an overstated $\Delta S_{t,t}'$ regressor and matching control firms with treatment firms based on inflated SG′ values. This method results in a total bias of -0.156% of lagged total assets reported in Column (8). But the bias introduced by matching on SG′ relative to the conceptually correct but un-implementable ROA + SG matching yields a very small difference of only 0.034% of lagged assets reported in Column (10). Thus, most of the downward bias in the Jones with
More realistic scenarios are presented in Subpanels A2 and A3 where the source of earnings management through revenue manipulation only occurs for some of the observations and expense understatement occurs for the remaining observations. We start with Row 6 where for 40% of observations sales are overstated by 5.0% of $S_{t-4}$, and for the remaining 60% of observations expenses are understated to create an equivalent amount of earnings management of 0.316% of lagged assets.

Relative to Subpanel A1, the Jones model with ROA matching leads to a larger upward bias in discretionary accruals of 0.096% of lagged assets (see column 7). This larger bias is because revenue manipulation is washed away for only 40% of the sample due to using observed (overstated) change in adjacent-quarter sales. This leaves a greater net upward bias due to failing to control for firm growth. The Jones with $ROA + SG'$ matching gives a smaller bias of -0.058% of lagged assets, and the difference between $ROA + SG'$ and $ROA + SG$ matching is a negligible 0.014% of lagged assets. Once again, the bias introduced by matching on inflated $SG'$ is minimal.

Row 9 shows the results when revenue manipulation is the source of earnings management for 20% of observations. In this case, the bias in Jones model discretionary accruals with ROA matching increases to 0.123%, the bias with $ROA + SG'$ matching decreases to -0.025%, and the difference between $ROA + SG'$ matching and $ROA + SG$ matching is close to zero.

Overall, with a realistic mix of revenue and expense manipulation as the two sources of earnings management, we find that Jones model with $ROA + SG'$ matching, an implementable approach, stands up well relative to other approaches for estimating discretionary accruals when sales revenue is manipulated by 5% of four-quarter lagged sales. If the revenue manipulation is only by half of this amount (i.e., the 2.5% seed cases in Rows 2, 5, and 8), this approach does much better for reasons explained in a footnote to Table 5. The bias in discretionary accrual estimates due to the effect of matching on $SG'$ is negligible, often less than 5 basis points in realistic settings (see column 10). Next, we show that variations of Mod-Jones(C) model give results that are even more favorable to the $ROA + SG'$ matching procedure.

[Insert Table 5 Here]
6.2. Variations of Mod-Jones(C) model (Panel B of Table 5)

Subsequent to Dechow, Sloan, and Sweeney (1995), it is well recognized that if revenue manipulation is the likely source of earnings management, then Mod-Jones(C) model provides less biased estimates of discretionary accruals compared to variations of the Jones model. This is because Mod-Jones(C) model includes \((\Delta S_{it} - \Delta AR_{it}^\prime)\) as a regressor. When revenue is manipulated, both terms in this expression are overstated by the same amount, so the difference between them is unaffected by revenue manipulation and the downward bias in discretionary accruals discussed in the previous section is eliminated. However, the upward bias due to failure to control for firm growth is exacerbated with Mod-Jones(C) model and ROA matching. This is because the entire change in accounts receivable is treated as discretionary accruals, which impacts high growth firms more than average growth firms. This bias depends mainly on how the sample is distributed across SG quintiles and it is largely independent of the source and magnitude of earnings management as evident from Column (7) across all nine rows of Panel B. It is also substantial in magnitude, averaging around 0.350% of lagged assets for our sample distribution.

The cases where revenues are managed upward by five percent for 100%, 40%, and 20% of the sample observations are presented in rows 12, 15, and 18. Mod-Jones(C) with \(ROA + SG^\prime\) matching gives reasonably unbiased estimates of discretionary accruals that differ from known discretionary accruals by between -0.031% and 0.024% of lagged assets (depending on the mix of sales overstatement and expense understatement). In comparison, Mod-Jones(C) with \(ROA + SG\) matching gives discretionary accruals that are upward biased by amounts ranging between 0.040% and 0.050% of lagged assets in each case. The difference between discretionary accruals obtained by these two matching procedures ranges between 0.015% and 0.073% (see column 10) and it is much smaller than the upward bias of around 0.350% of lagged assets created by ignoring sales growth matching altogether. This finding alleviates the concern that in cases of revenue overstatement, sales growth matching throws out a large part of discretionary accruals (the baby) with non-discretionary accruals (the bathwater). Even in the worst case scenario considered in Row 12 (where 100% of the observations are subject to revenue manipulation) the bias due to matching on an inflated revenue amount is around one-fifth of the bias due to failing to control for firm
growth. Finally, with half the revenue manipulation (i.e., 2.5% seed in Rows 11, 14, and 17) the bias due to matching on an inflated revenue amount becomes less than a third of the bias reported in rows 12, 15, and 18, while the bias due to failing to control for firm growth remains comparable.

We draw two conclusions from this discussion. First, if revenue manipulation is a serious concern in samples with an over-representation of growth firms, then Mod-Jones(C) model is a better starting point as recognized in the previous literature. Second, if the residuals from this model are not adjusted further by subtracting the corresponding residuals of a performance and sales growth matching firm, then the estimated discretionary accruals are likely to contain a strong upward bias in samples that are over-represented by high growth firms. In contrast, performance and sales growth matching produces well specified results, even if sales growth has been overstated due to revenue manipulation.

7. Comparison of ROA + SG matching methodology with reversal methodology

Dechow et al. (2012) propose a new methodology of detecting earnings management that exploits the reversal property of discretionary accruals. The following equation sets forth the specification for this methodology:

\[ WCA_{i,t} = a + b \ PART_{i,t} + c \ PARTR1_{i,t} + d \ PARTR2_{i,t} + \]
\[ Usual\ Jones\ - type\ model\ terms + \varepsilon_{i,t} \]  

In this framework, \( PART \) is a dummy variable that takes the value one for a period during which the accruals are managed, and zero otherwise. \( PARTR1 \) and \( PARTR2 \) are dummy variables that take the value one during first and second reversal periods, and zero otherwise. If the accruals are expected to reverse only during the first period, then the second reversal term is dropped. Using a pooled sample of firm-years, Dechow et al. test whether the condition \( b - (c + d) = 0 \) can be rejected in favor of \( b - (c + d) > 0 \) for upward earnings management and \( b - (c + d) < 0 \) for downward earnings management.\(^{32}\)

\(^{32}\) It is important to note that Dechow et al (2012) conduct their tests of the reversal methodology only in annual settings and they do not assert that the methodology would be equally effective in quarterly settings. We include a brief comparison of the reversal and matching procedure in a quarterly setting here for completeness. A more comprehensive comparison of reversal and matching methodologies is the subject of on-going research by Collins, Pungaliya and Vijh (2011).
Dechow et al. suggest that the reversal methodology has two advantages over typical cross-sectional matching methodology. First, the reversals methodology corrects model misspecification related to a variety of firm characteristics without having to identify the source of misspecification. This is based on the assumption that factors associated with the non-discretionary part of accruals during the earnings management and reversal periods remain constant and approximately cancel each other. Second, they suggest that the reversals methodology is more powerful than procedures that use matching to control for factors correlated with the partitioning variable. This is based on the assumption that the earnings management and reversal periods can be identified reasonably accurately, in which case \( b - (c + d) \) is approximately twice the value of \( b \) (the numerator of test statistic) while the standard error of this difference (the denominator) is approximately \( \sqrt{3} \) times the standard error of \( b \). By comparison, the numerator for the matching procedure, \( b \), is unchanged and the denominator increases by a factor of \( \sqrt{2} \) in cross-sectional methodology as outlined earlier in this paper (the result of differencing Jones-type model residuals).

Applying the reversal methodology in quarterly settings faces several challenges. First, specifying the number of quarterly periods over which the hypothesized earnings management will reverse is problematic. Accelerated revenue recognition may be expected to reverse in one or two subsequent quarters, while capitalizing a cost that should be expensed or taking excessive asset write-downs are likely to reverse over a much longer horizon. Faced with this uncertainty, a researcher may err on the side of including too many periods, which lowers the power of the reversal methodology because the standard error of the test statistic increases as a function of \( \sqrt{T} \) where \( T \) is the combined number of earnings management and reversal periods. In addition, including too many periods in the reversal horizon will contribute to greater misspecification (Type I error) because there is no guarantee that factors contributing to non-discretionary accruals in the earnings management period and the reversal periods will cancel out. On the other hand, including too short of reversal period fails to fully exploit the advantages of the reversal methodology by lowering the numerator of the test statistic below what it would otherwise be. Finally, instances of earnings management in quarterly settings typically span multiple quarters (GAO...
Thus, originating earnings management and reversals of previous periods’ earnings management will tend to cancel out, which greatly complicates the specification for testing.

In sum, firms have considerable discretion over how they go about managing accruals during a chosen quarter and these choices have a substantial impact on the subsequent reversal process. Further, this discretion is likely to differ across industries, which makes it difficult to know how many quarters to include as reversal quarters. Seasonality considerations further complicate how the reversal process will unfold. A full investigation of specification and power issues of reversals methodology using quarterly data is beyond the scope of this paper. Here we describe a simple experiment to compare the power of Jones and Mod-Jones(C) models with ROA + SG matching with the power of Dechow et al. (2012) reversal methodology employing the same models. As explained earlier, power tests are best carried out over the neutral sample of all 203,090 firm-quarters for which misspecification issues arising from firm characteristics like firm growth are of minimal concern. From this comprehensive neutral sample we randomly select sample sizes of N = 600, 1,000, or 2,000 firm-quarters and inflate their raw accruals by 0.25%. Following the evidence in Baber, Kang, and Li (2011), we assume reversals take place over the subsequent quarters with the indicated frequencies: one quarter (43%), two quarters (29%), three quarters (21%), and four quarters (7%). Further, if reversals take place over n quarters, we assume that $1/n$ of the total reversal occurs each quarter from 1 to n, following the earnings management quarter. Finally, we assume that the following proportion of the original earnings management reverses during the specified reversal horizon: 100%, 50%, and 30%. We merge these seeded firm-quarters and reversal quarters with the remaining firm-quarters and carry out the following regression:

$$WCA_{i,t} = a + b\ Part_{i,t} + c\ PartR1_{i,t} + d\ PartR2_{i,t} + e\ PartR3_{i,t} + f\ PartR4_{i,t}$$

$$+ Usual\ Jones - type\ model\ terms + \varepsilon_{i,t}$$ (9)

Baber, Kang, and Li (2011) model the reversal process of quarterly accruals and show that if period t discretionary accruals ($d_t$) reverse fully in period $t + n$, where $n \geq 1$, then the minimum value of k'th order autocorrelation $\rho_k$, is achieved when $k = n$. Using this proposition, they measure the length of reversal process as the lag at which the quarterly accruals have the most negative autocorrelation. Examining a large sample of firms, they find that in only 41% of all cases does the most negative autocorrelation occur during the first four quarters (i.e., $n \leq 4$). Restricting their analysis to the first four quarters, the most negative autocorrelation occurs at the first lag in 43% of cases, second lag in 29% of cases, third lag in 21% of cases, and fourth lag in 7% of cases.
We repeat the procedure 250 times and record the frequency with which the null hypothesis of $b - (c + d + e + f) = 0$ can be rejected in favor of the alternate hypothesis of $b - (c + d + e + f) > 0$ with 5% significance level.

The results are shown in Table 6 and are briefly summarized here. Across all three sample sizes, the reversal methodology with quarterly data has lower power than Jones and Mod-Jones(C) models with ROA + SG matching without CFO adjustment. The differences in power become more dramatic as the sample size increases and as the proportion of original earnings management that reverses over the subsequent four quarters decreases. For example, for a sample size of $N = 2000$ the Jones model with ROA + SG matching has a rejection rate of 63.2% with mean test and 75.6% with median test. In comparison, the reversal methodology with Jones model has a rejection rate of 39.6% with 100% reversal, 24.0% with 50% reversal, and 19.2% with 30% reversal. Similar differences are seen for Mod-Jones(C) model.

[Insert Table 6 Here]

8. Conclusions

Numerous studies in accounting and finance investigate potential earnings management in a variety of settings. Typically, researchers test whether some measure of discretionary accruals averaged across a sample of firms is significantly different from zero in the predicted direction. The choice of the discretionary accrual measure thus becomes critically important. Following the evidence of Jones (1991) and Dechow, Sloan, and Sweeny (1995), residuals from Jones or modified-Jones model have been the popular starting point in many studies. More recently, following the evidence of Kothari, Leone, and Wasley (2005) that accruals are correlated with firm performance, these residuals are performance-adjusted by subtracting similar residuals for ROA matching firms from the same industry.

Dechow, Kothari, and Watts (1998) demonstrate that high sales growth firms require legitimate higher investments in working capital to deal with higher customer demand. Their model implies that the growth-related change in accruals should be treated as non-discretionary because this component of accruals is predictable and common across growth firms. Thus, in the absence of controls for firm growth
and performance, Jones-type discretionary accrual estimates will be confounded with the growth- and performance-related component of accruals.

McNichols (2000) is among the first to recognize the confounding effects of firm growth on Jones-type model discretionary accrual estimates. She warns of possible misspecification of these models, but stops short of identifying settings where this is most problematic or demonstrating the seriousness of the problem. Despite the warnings issued by McNichols, most earnings management studies over the past decade follow the guidance provided in Kothari, Leone, and Wasley (2005) and use performance (ROA)-matched Jones-type model discretionary accrual estimates in tests of earnings management, implicitly assuming that any distortion due to firm growth is minimal.

In this paper we demonstrate the rather severe misspecification (in terms of Type I error rates) that exists in tests of earnings management in quarterly settings that use Jones-type model discretionary accrual estimates, even after performance matching, and we show how matching on both performance and growth measures results in well-specified tests. We extend the analysis in McNichols (2000) in several ways. First, we identify multiple partitioning variables used in prior earnings management research (stock splits, SEOs, stock acquisitions, equity-based compensation, and insider trading) and demonstrate how these partitioning variables are correlated with firm growth measures. We show that the resulting measurement error can lead to over rejection of the null hypothesis of no earnings management in these settings.

Next, using stratified random samples of firms with no known earnings management, we show that the traditional discretionary accrual measures based on Jones or modified-Jones models with ROA matching are highly misspecified in both high growth and low growth subsamples of firm-quarters. The modified-Jones model as commonly estimated in the literature with ROA matching is particularly misspecified, with rejection rates of the null hypothesis of no earnings management as high as 67% and 58% (compared to theoretical 5%) in low SG and high SG quintiles.

Finally, using simulations we demonstrate that Jones-type model discretionary accrual estimates adjusted for accruals’ noise reduction role and asymmetric timely loss recognition and matched on both performance (ROA) and sales growth (SG) yield well specified tests with reasonable power (70% or
greater for parametric $t$-test of mean and 90% or greater for non-parametric Wilcoxon signed rank test of median) of detecting modest amounts (0.25% of total assets) of earnings management in sample sizes commonly found in the literature. Moreover, matching on SG does not introduce serious bias in discretionary accrual estimates when earnings management is accomplished through revenue manipulation. Our findings thus suggest that, going forward, researchers should adjust for both performance and firm growth when testing for earnings management, particularly in settings where the partitioning variable deemed to give rise to earnings management is likely to be correlated with firm growth. In addition, adjustments for accruals’ noise reduction role and for asymmetric timely loss recognition appear warranted, particularly in quarterly settings where seasonality is likely to affect the dynamics of the accrual process.
### Appendix 1

**A survey of research methodologies of detecting earnings management and empirical findings in previous literature**

The sample of 70 articles on earnings management is identified by a visual examination of all research articles referenced in Dechow, Ge, and Schrand (2009). We scan each article to obtain the summary information reported in this appendix.

#### A1.1 Summary information

<table>
<thead>
<tr>
<th>Description</th>
<th>Annual vs. Quarterly Accrual Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annual data or quarterly data</td>
<td>Annual: 48</td>
</tr>
<tr>
<td>Raw accrual measure</td>
<td>Balance sheet approach: 33</td>
</tr>
<tr>
<td></td>
<td>Cash flow statement: 9</td>
</tr>
<tr>
<td></td>
<td>Both: 6</td>
</tr>
<tr>
<td>Discretionary accrual model</td>
<td>Jones: 21</td>
</tr>
<tr>
<td></td>
<td>Modified Jones: 33</td>
</tr>
<tr>
<td></td>
<td>ROA Adjustment: 8</td>
</tr>
<tr>
<td>Cross-sectional or time-series application of Jones-type models</td>
<td>Cross-sectional application: 37</td>
</tr>
<tr>
<td></td>
<td>Time-series application: 8</td>
</tr>
<tr>
<td></td>
<td>Both: 3</td>
</tr>
<tr>
<td>The partitioning variable is likely to be correlated with firm growth</td>
<td>Yes: 30</td>
</tr>
<tr>
<td></td>
<td>No: 17</td>
</tr>
<tr>
<td></td>
<td>Unclear: 1</td>
</tr>
<tr>
<td>Article concludes in favor of earnings management</td>
<td>Yes: 42</td>
</tr>
<tr>
<td></td>
<td>No: 4</td>
</tr>
<tr>
<td></td>
<td>Mixed: 2</td>
</tr>
<tr>
<td>Correlation between discretionary accrual model and conclusion in favor of earnings management</td>
<td>Jones model and evidence of earnings management: 19</td>
</tr>
</tbody>
</table>
### A1.2 Detailed information on studies that use quarterly data.

<table>
<thead>
<tr>
<th>Article</th>
<th>Topic (partitioning variable)</th>
<th>Is partitioning variable likely correlated with firm growth?</th>
<th>Sample Description</th>
<th>Methodology for computing discretionary accruals</th>
<th>Quarterly, Time-series (TS) or cross-sectional (CS)</th>
<th>Evidence of earnings mgmt.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abarbanell and Lehavy (JAR, 2003)</td>
<td>Analyst recommendations</td>
<td>Yes</td>
<td>22,173 firm quarters between 1985-1998</td>
<td>X</td>
<td>BS</td>
<td>Yes</td>
</tr>
<tr>
<td>Baber, Chen, and Kang (RAS, 2006)</td>
<td>Supplementary financial disclosure</td>
<td>No</td>
<td>10,248 firm quarters between 1993-1997</td>
<td>X</td>
<td>CF</td>
<td>Yes</td>
</tr>
<tr>
<td>Baker, Collins, and Reitenga (CAR, 2009)</td>
<td>CEO stock option grants</td>
<td>Yes</td>
<td>21,388 firm quarters between 1993-2003</td>
<td>X</td>
<td>X</td>
<td>BS?</td>
</tr>
<tr>
<td>Balsam, Bartov, and Marquardt (JAR, 2002)</td>
<td>Supplementary financial disclosure in 10-Q</td>
<td>No</td>
<td>613 firm quarters during 1996-1998</td>
<td>X</td>
<td>Both</td>
<td>Yes</td>
</tr>
<tr>
<td>Bartov, Givoly, and Hayn (JAE, 2002)</td>
<td>Meeting or beating earnings expectations</td>
<td>Yes</td>
<td>65,000 firm quarters between 1983-1997</td>
<td>X</td>
<td>BS</td>
<td>Yes</td>
</tr>
<tr>
<td>Brown and Pinello (JAR, 2007)</td>
<td>Financial reporting process (interim vs. annual)</td>
<td>No</td>
<td>29,684 firm quarters between 1993-2005</td>
<td>X</td>
<td>CF</td>
<td>Yes</td>
</tr>
<tr>
<td>Coles, Hertzel, and Kalpathy (JAE, 2006)</td>
<td>Stock option reissues</td>
<td>Yes</td>
<td>159 firms during 2001-2002</td>
<td>X</td>
<td>BS</td>
<td>Yes</td>
</tr>
<tr>
<td>Das, Shroff, and Zhang (CAR, 2009)</td>
<td>Fourth quarter earnings reversal</td>
<td>Yes</td>
<td>71,963 observations during 1988-2004</td>
<td>X</td>
<td>X</td>
<td>CF</td>
</tr>
<tr>
<td>Erickson and Wang (JAE 1999)</td>
<td>Earnings mgmt (EM) by stock acquirers</td>
<td>Yes</td>
<td>55 stock acquirers between 1985-1990</td>
<td>X</td>
<td>Both</td>
<td>Yes</td>
</tr>
<tr>
<td>Gong, Louis, and Sun (JF 2008a)</td>
<td>Stock repurchases</td>
<td>Yes</td>
<td>1,720 repurchases between 1984-2002</td>
<td>X</td>
<td>X</td>
<td>BS</td>
</tr>
<tr>
<td>Author(s)</td>
<td>Topic</td>
<td>Data Provided</td>
<td>Sample Size</td>
<td>Data Frequency</td>
<td>Data Source</td>
<td>Required Calculation</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>------------------------------------------</td>
<td>---------------</td>
<td>------------------------------------------------------------------------------</td>
<td>----------------</td>
<td>--------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td>Gong, Louis, and Sun (JAE 2008b)</td>
<td>Post merger lawsuits</td>
<td>Yes</td>
<td>103 litigated stock acquisitions between 1996-2002</td>
<td>X</td>
<td>X</td>
<td>BS</td>
</tr>
<tr>
<td>Han and Wang (AR, 1998)</td>
<td>Political costs and Oil companies</td>
<td>Yes</td>
<td>76 firms in 1990</td>
<td>X</td>
<td></td>
<td>BS</td>
</tr>
<tr>
<td>Jo and Kim (JFE, 2007)</td>
<td>Disclosure frequency (SEO sample)</td>
<td>Yes</td>
<td>1,950 SEOs between 1990-1997</td>
<td>X</td>
<td>X</td>
<td>CF</td>
</tr>
<tr>
<td>Jo, Kim, and Park (RAS, 2007)</td>
<td>Underwriter choice (SEO sample)</td>
<td>Yes</td>
<td>1,950 SEOs between 1990-1997</td>
<td>X</td>
<td>X</td>
<td>CF</td>
</tr>
<tr>
<td>Kim and Park (JFQA, 2005)</td>
<td>Seasoned equity offerings</td>
<td>Yes</td>
<td>1040 SEOs from 1989-2000</td>
<td>X</td>
<td>X</td>
<td>CF</td>
</tr>
<tr>
<td>Louis and Robinson (JAE, 2005)</td>
<td>Stock splits</td>
<td>Yes</td>
<td>2,271 stock splits between 1990-2002</td>
<td>X</td>
<td>X</td>
<td>BS</td>
</tr>
<tr>
<td>Rangan (JFE, 1998)</td>
<td>Seasoned equity offerings</td>
<td>Yes</td>
<td>712 SEOs between 1987-1990</td>
<td>X</td>
<td></td>
<td>BS</td>
</tr>
<tr>
<td>Shivakumar (JAE, 2000)</td>
<td>Seasoned equity offerings</td>
<td>Yes</td>
<td>2,995 SEOs between 1983-1992</td>
<td>X</td>
<td></td>
<td>Both</td>
</tr>
</tbody>
</table>

1 BS denotes that raw accruals are calculated from the balance sheet, and CF denotes that the raw accruals are calculated from the cash flow statement.
Appendix 2: Literature on earnings management detected using discretionary accrual models

This is necessarily a partial list

A2.1 Event studies of earnings management around stock splits

Louis and Robinson (2005)

A2.2 Event studies of earnings management around IPOs and SEOs


A2.3 Event studies of earnings management around stock acquisitions


A2.4 Event studies of earnings management around stock repurchases

Hribar, Jenkins, and Johnson (2006), Gong, Louis, and Sun (2008b)

A2.5 Event studies of earnings management to maintain dividend payment

Daniel, Denis, and Naveen (2008)

A2.6 Cross-sectional relation between earnings management and performance-based executive compensation


A2.7 Cross-sectional relation between earnings management and option grants, option exercises, option repricings, and stock trading

References


Blair, Clifford, and James Higgins, 1985, Comparison of the power of the paired sample t test to that of Wilcoxon’s signed-ranks test under various population shapes, *Psychological Bulletin* 97, 119-128.


Dechow, Patricia, Weili Ge, and Catherine Schrand, 2009, Understanding earnings quality: A review of the proxies, their determinants and their consequences, working paper, SSRN.


Figure 1: Event distribution across sales growth deciles (top panel) and employee growth deciles (bottom panel). We start with a comprehensive sample of 203,090 firm-quarters during 1991-Q1 to 2007-Q4 from the Compustat and CRSP databases as described in Section 3.1. We merge this sample with samples of firms that announced stock splits, SEOs, and stock acquisitions. Stock splits are identified from the CRSP database using distribution code of 5523 and a positive split factor, and SEOs and stock acquisitions are identified from the SDC database. We require that the event announcement date and the quarterly earnings announcement date are available. The final samples include 2,646 stock splits, 2,951 SEOs, and 1,193 stock acquisitions. The sales growth is calculated as the sales during the quarter with the earnings announcement date preceding the event date divided by the sales during the same quarter of the previous year, minus one. The corresponding decile ranks are calculated each quarter using the data for all 203,090 firm-quarters. Since the employee numbers are only available annually, we follow a parallel procedure with a comprehensive sample of firm-years to calculate employee growth deciles. The average sales growth decile ranks equal 7.25, 7.04, and 7.60 for event firms making stock splits, SEOs, and stock acquisitions. The corresponding values for employee growth decile ranks equal 7.10, 7.20, and 7.79.
Figure 2: Stock based compensation (left bars) and abnormal insider selling (right bars) across sales growth deciles (top panel) and employee growth deciles (bottom panel). We start with a comprehensive sample of 41,383 firm-years during 1991 to 2007 from the Compustat and CRSP databases. From this we select a subset of firm-years for which stock based compensation data are available from ExecuComp (1992 to 2007) or insider buying and selling data are available from Thomson Financial (1991 to 2007). The insider trading data pass through several filters commonly employed in previous literature (form type 4, cleanse code R and H, transaction code P and S, and acquisition and disposal of at least 100 shares). Sales growth is calculated as sales during the year ending before the current year divided by sales during the previous year, minus one. Employee growth is similarly calculated. The corresponding decile ranks are calculated each year using the data for all firm-years. Stock based compensation is calculated as the Black-Scholes value of stock option grants plus the market value of restricted stock divided by total compensation and multiplied by 100, and total compensation is defined as the value of stock options and restricted stock plus salary and bonus. Following Beneish and Vargus (2002), firm-years characterized by abnormal insider selling are identified as follows. First, we sum the total sales and the total purchases of shares by the top five executives, calculate the difference, and divide by the total shares outstanding. Second, we check whether this scaled difference is greater than the corresponding median value for all firm-years with the same market value decile rank. The left axis in both panels shows the median value of stock based compensation, and the right axis shows the percent of all cases for which there was abnormal insider selling.
Figure 3. Discretionary accrual measures across ROA and SG deciles for the aggregate sample of Compustat firm-quarters. Panels A, B, and C plot Jones model residuals, Jones model residuals with ROA matching, and Jones model residuals with ROA + SG matching. Panels D, E, and F do the same for Mod-Jones(C) model. The calculation of these discretionary accrual measures is described in Section 2.3 and Table 2. The sample includes the 203,090 Compustat firm-quarters as described in Section 3.1 and Table 2. Each calendar quarter, we divide this sample into one hundred subsets formed by the ROA and SG decile ranks and calculate the median discretionary accruals for each subset. ROA equals net income divided by the total assets as of quarter \( t-4 \) (where quarter \( t \) is the current quarter of discretionary accruals), and SG equals the sales during quarter \( t \) divided by the sales during quarter \( t-4 \) minus one.
Figure 3. continued.

Panel A: Jones model discretionary accruals, comprehensive sample

Panel B: Jones model discretionary accruals with ROA matching, comprehensive sample

Panel C: Jones model discretionary accruals with ROA+SG matching, comprehensive sample
Figure 3. continued.

Panel D: Mod-Jones(C) model discretionary accruals, comprehensive sample

Panel E: Mod-Jones(C) model discretionary accruals with ROA matching, comprehensive sample

Panel F: Mod-Jones(C) model discretionary accruals with ROA+SG matching, comprehensive sample
Figure 4. Specification tests of discretionary accrual measures using quarterly data as an increasing proportion of the sample is drawn from the top quintile of sales growth. This figure provides specification tests similar to Table 4, but with one major difference. Whereas Table 4 examines a sample size of 200 firm-quarters from the top SG quintile, this figure examines a sample size of 1,000 firm-quarters distributed as follows. First, we pull a certain proportion of the sample from the top quintile of SG as noted on the x-axis. Second, we pull the remaining sample randomly from all SG quintiles. The vertical axis shows the percentage of 250 such samples where the null hypothesis of zero discretionary accruals is rejected at the 5% level using one-tailed t-test for mean. The aggregate sample of 203,090 Compustat firm-quarters is described in Section 3.1 and Table 2. The calculation of Jones and Mod-Jones(C) models, and the partitioning variable are described in Section 2.3 and Table 2.

In each panel, the solid line represents the raw unmatched residuals from a Jones-type model, the dotted line represents the ROA matched residuals, and the dashed line represents the ROA+SG matched residuals.

Panel A: Jones model variants

Panel B: Mod-Jones(C) model variants
Figure 5. Specification tests of discretionary accrual measures using quarterly data as an increasing proportion of the sample is drawn from the top quintile of sales growth: Models with CFO adjustment (see Ball and Shivakumar, 2006). This figure is similar to Figure 4, except that all Jones-type models include CFO adjustment as described in Section 2.3 and Table 2. Specifically, we examine a sample size of 1,000 firm-quarters distributed as follows. First, we pull a certain proportion of the sample from the top quintile of SG as noted on the x-axis. Second, we pull the remaining sample randomly from all SG quintiles. The vertical axis shows the percentage of 250 such samples where the null hypothesis of zero discretionary accruals is rejected at the 5% level using one-tailed t-test for mean. The aggregate sample of 203,090 Compustat firm-quarters is described in Section 3.1 and Table 2.

In each panel, the solid line represents the raw unmatched residuals from a Jones-type model, the dotted line represents the ROA matched residuals, and the dashed line represents the ROA+SG matched residuals.

Panel A: **Jones model variants, with CFO adjustment**

Panel B: **Mod-Jones(C) model variants, with CFO adjustment**
Figure 6. Power of discretionary accrual measures to reject the null hypothesis of no earnings management in favor of the alternate hypothesis of positive earnings management in samples drawn from the aggregate dataset of all Compustat firm-quarters. The aggregate sample includes all 203,090 Compustat firm-quarters described in Section 3.1 and Table 2. The figure shows the percentage of 250 random samples of between 200 and 2,000 firm-quarters each where the null hypothesis of zero discretionary accrual is rejected at the 5% significance level using one-tailed t-test for mean. The higher the rejection rate, the more powerful the discretionary accrual measure in detecting earnings management. Panel A reports model power using t-test for mean, and Panel B reports model power using Wilcoxon signed rank test for median. For each sample firm-quarter we increase the raw current accrual computed using the cash flow method by 0.25% of total assets. All discretionary accrual measures with and without CFO adjustment are described in Section 2.3 and Table 2.

Panel A: Jones and Mod-Jones(C) model variants, rejection rates based on t-test for mean
Figure 6. continued.

Panel B: Jones and Mod-Jones(C) model variants, rejection rates based on Wilcoxon signed rank test for median

Seed = 0.25%
Figure 7. Power of discretionary accrual measures with and without SG and CFO adjustment in samples of 1,000 firm quarters, of which the first 500 are drawn randomly from the top SG quintile and the next 500 are drawn from all SG quintiles. The null hypothesis of zero discretionary accruals is tested against the alternate hypothesis of positive discretionary accruals in all panels of this figure. Panel A shows the rejection rates as a function of the seed size for the following models without CFO adjustment: Jones, Jones + ROA, Jones + ROA + SG. Panel B shows the corresponding results for Mod-Jones(C) model without CFO adjustment. Finally, Panels C and D show the rejection rates for Jones and Mod-Jones(C) models, but with CFO adjustment. The aggregate sample includes all 203,090 Compustat firm-quarters divided into five SG quintiles as described in Table 2. The figure shows the percentage of 250 random samples where the null hypothesis of zero discretionary accrual is rejected at the 5% significance level using one-tailed t-test for mean. The seed size varies from -0.500% to +0.500% in steps of 0.125%. Thus, rejection rates higher than 5% for a seed size of 0.00% represent model misspecification rather than model power. Discretionary accrual measures with and without CFO adjustment are described in Section 2.3 and Table 2.
Figure 7. continued.

Panel A: Jones model variants, without CFO adjustment

Panel B: Mod-Jones(C) model variants, without CFO adjustment
Figure 7. Continued.

Panel C: Jones model variants, with CFO adjustment

Panel D: Mod-Jones(C) model variants, with CFO adjustment
Figure 8. Quantitative biases in discretionary accrual measures with ROA matching versus ROA + SG matching in samples of 1,000 firm quarters, of which the first 500 are drawn randomly from the top SG quintile and the next 500 are drawn from all SG quintiles. Panels A and B of this figure show the average biases underlying the power curves for various discretionary accrual measures shown in Panels A and B of Figure 7. These biases are calculated by first averaging the discretionary accrual measures across the 1,000 firm quarters in one random draw, and then averaging the averages across 250 draws. Discretionary accrual measures are described in Section 2.3 and Table 2.

Panel A: Comparison of seeded accruals with measured discretionary accruals - Variations of Jones model

Panel B: Comparison of seeded accruals with measured discretionary accruals - Variations of Mod-Jones(C) model
Table 1
Biases in discretionary accruals estimated using quarterly data and without sales growth (SG) matching before select events

We start with a comprehensive sample of 203,090 firm-quarters during 1991 to 2007 from the Compustat and CRSP databases as described below in Section 3.1 and Table 2 and merge it with samples of firms that announced stock splits, SEOs, and stock acquisitions. Stock splits are identified from the CRSP database using distribution code 5523 and a positive split factor, and SEOs and stock acquisitions are identified from the SDC database. We calculate several accrual measures using Compustat data for quarter \( t \), which is the fiscal quarter with an earnings announcement date immediately preceding the event date. Raw accruals are calculated using the cash flow statement as described in Section 3.1 and Table 2, and discretionary accruals are first calculated as Jones or Mod-Jones(C) model residuals as described in Section 2.3 and Table 2. ROA matching or ROA + SG matching discretionary accruals are next calculated as the difference between Jones model or Mod-Jones(C) model residuals for a sample firm and its matching firm. For ROA matching we choose the matching firm that is the same-industry firm with the closest ROA during quarter \( t-4 \). For ROA + SG matching accruals we arrange all same-industry firms during quarter \( t-4 \) into five ROA quintiles and choose the matching firm that has the closest SG from quarter \( t-4 \) to \( t \) in the relevant quintile. We calculate ROA as net income divided by total assets, and SG as sales during quarter \( t \) divided by sales during quarter \( t-4 \) minus one. This table reports the mean values of various discretionary accrual measures as well as their \( t \)-statistics (in parentheses). It also reports the mean ROA, SG, and EG decile ranks for the three event samples using the distribution of all Compustat firms during quarters \( t \). EG is calculated using Compustat annual data, and it is the employee growth rate over a one-year period ending the fiscal year before the event date. The notations *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels in one-tailed tests.

<table>
<thead>
<tr>
<th>Description</th>
<th>Stock splits</th>
<th>SEOs</th>
<th>Stock acquisitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jones model</td>
<td>0.188</td>
<td>0.276</td>
<td>0.281</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(2.58)***</td>
<td>(4.19)***</td>
<td>(2.48)***</td>
</tr>
<tr>
<td>Jones model with ROA matching</td>
<td>0.192</td>
<td>0.204</td>
<td>0.069</td>
</tr>
<tr>
<td></td>
<td>(1.84)**</td>
<td>(2.08)**</td>
<td>(0.40)</td>
</tr>
<tr>
<td>Jones model with ROA + SG matching</td>
<td>0.032</td>
<td>0.202</td>
<td>-0.027</td>
</tr>
<tr>
<td></td>
<td>(0.30)</td>
<td>(2.07)**</td>
<td>(-0.16)</td>
</tr>
<tr>
<td>Mod-Jones(C) model</td>
<td>0.467</td>
<td>0.554</td>
<td>0.616</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(6.47)***</td>
<td>(8.25)***</td>
<td>(5.56)***</td>
</tr>
<tr>
<td>Mod-Jones(C) model with ROA matching</td>
<td>0.406</td>
<td>0.471</td>
<td>0.341</td>
</tr>
<tr>
<td></td>
<td>(3.93)***</td>
<td>(4.75)***</td>
<td>(2.02)**</td>
</tr>
<tr>
<td>Mod-Jones(C) model with ROA + SG matching</td>
<td>0.068</td>
<td>0.314</td>
<td>-0.053</td>
</tr>
<tr>
<td></td>
<td>(0.65)</td>
<td>(3.24)***</td>
<td>(-0.31)</td>
</tr>
<tr>
<td>ROA decile rank</td>
<td>7.28</td>
<td>5.34</td>
<td>6.19</td>
</tr>
<tr>
<td>SG decile rank</td>
<td>7.25</td>
<td>7.04</td>
<td>7.60</td>
</tr>
<tr>
<td>EG decile rank</td>
<td>7.10</td>
<td>7.20</td>
<td>7.79</td>
</tr>
<tr>
<td>N</td>
<td>2,646</td>
<td>2,951</td>
<td>1,193</td>
</tr>
</tbody>
</table>

1 All accruals are reported in percent form. Thus, an accrual of 0.123 means 0.123% of lagged assets (and not 0.123 times lagged assets).
Table 2

Descriptive statistics for various accrual measures and partitioning variables used in the simulation exercise

The sample and methodology are described in Sections 3.1 and 2.3 and reproduced in this table. The sample consists of all Compustat firm-quarters during 1991-Q1 to 2007-Q4 for which the relevant data to calculate the accrual measures and the partitioning variables reported in this table are available. We additionally require that: (1) Total assets exceed $10 million in 2007 dollars; (2) The firm is not in the financial industry (which excludes two-digit SIC codes between 60 and 69); (3) The CRSP share code is 10 or 11 (which excludes ADRs, REITs, units, certificates, and trusts); (4) There are at least 20 firms in the included two-digit SIC code during a given calendar quarter; and (5) None of the accrual measures (normalized by total assets) exceeds one. The final sample consists of 203,090 firm-quarters. The calculation of the various accrual measures follows several steps. First, we compute current accruals as $- (CHGAR+CHGINV+CHGAP+CHGTAX+CHGOTH)$, where the bracketed quantities represent the change in accounts receivable (item RECCHY), inventories (item INVCHY), accounts payable (item APALCHY), taxes (item TAXCHY), and other items (item ALOLOCHY). All quantities are obtained from the cash flow statement. We undo the year-to-date nature of these quarterly cash flow statement items and compute the quantities for the current quarter. In addition, we recode missing values of RECCHY, INVCHY, APALCHY, and TAXCHY as zero if there is a nonmissing value of ALOLOCHY. Conversely, if ALOLOCHY is missing but the other items are not missing, then we recode ALOLOCHY as zero. Second, we carry out the following cross-sectional regression:

$$WCA_{t,i} = \beta_0 + \beta_1 Q_{1,i,t} + \beta_2 Q_{2,i,t} + \beta_3 Q_{3,i,t} + \beta_4 Q_{4,i,t} + \beta_5 \Delta \text{SALES}_{i,t} + \beta_6 WCA_{t-4,i} + \epsilon_{i,t}. \quad (T2.1)$$

In this expression, subscript $i$ denotes the firm and $t$ denotes the quarter. $Q_{1,i,t} - Q_{4,i,t}$ are the fiscal quarter dummies, $\Delta \text{SALES}_{i,t}$ is the quarterly change in sales, and $WCA_{t-4,i}$ is the current accrual from the same quarter in the preceding year. The residuals $\epsilon_{i,t}$ from Model (T2.1) constitute the Jones model discretionary accruals. We estimate the following cross-sectional regression for the Mod-Jones(C) model:

$$WCA_{t,i} = \lambda_0 + \lambda_1 Q_{1,i,t} + \lambda_2 Q_{2,i,t} + \lambda_3 Q_{3,i,t} + \lambda_4 Q_{4,i,t} + \lambda_5 (\Delta \text{SALES}_{i,t} - \Delta AR_{i,t}) + \lambda_6 WCA_{t-4,i} + \xi_{i,t}. \quad (T2.2)$$

The residuals $\xi_{i,t}$ from Model (T2.2) constitute the Mod-Jones(C) model discretionary accruals. All variables are scaled by lagged total assets, and the regressions are run by the calendar quarter across all same-industry firms (i.e., with the same two-digit SIC code as the sample firm). Fifth, we calculate the ROA matching or ROA+SG matching discretionary accruals as the difference between Jones model or Mod-Jones(C) model residuals for a sample firm and its matching firm. For ROA matching we choose the matching firm that is the same-industry firm with the closest ROA during quarter $t-4$. For ROA+SG matching we arrange all same-industry firms during quarter $t-4$ into five ROA quintiles and choose the matching firm that has the closest SG from quarter $t-4$ to $t$ in the relevant quintile. We calculate ROA as the net income divided by total assets, and SG as the sales during quarter $t$ divided by sales during quarter $t-4$ minus one. Only employee growth is calculated using Compustat annual data, and it is over a one-year period ending the fiscal year before the current quarter $t$. All accrual measures and partitioning variables are winsorized at the 1% and 99% levels. This table presents descriptive statistics only for the aggregate sample of 203,090 firm-quarters.
Table 2. continued. Descriptive statistics for various accrual measures and partitioning variables used in the simulation exercise.

<table>
<thead>
<tr>
<th>Description</th>
<th>Aggregate sample</th>
<th>High sales growth quintile</th>
<th>Low sales growth quintile</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N = 203,090</td>
<td>N = 40,604</td>
<td>N = 40,592</td>
</tr>
<tr>
<td>Current accruals¹</td>
<td>0.43  0.32   4.54</td>
<td>1.20  0.82    5.10</td>
<td>-0.43  -0.15  4.97</td>
</tr>
<tr>
<td>Jones model</td>
<td>0.00 -0.00  3.69</td>
<td>0.35  0.23    4.20</td>
<td>-0.45 -0.27  4.14</td>
</tr>
<tr>
<td>Jones model with ROA matching</td>
<td>-0.00  0.00  5.26</td>
<td>0.33  0.26    5.81</td>
<td>-0.45 -0.32  5.70</td>
</tr>
<tr>
<td>Jones model with ROA+SG matching</td>
<td>-0.01  0.00  5.22</td>
<td>-0.03 -0.01  5.95</td>
<td>-0.04 -0.02  5.83</td>
</tr>
<tr>
<td>Mod-Jones(C) model</td>
<td>0.00 -0.02  3.70</td>
<td>0.78  0.50    4.20</td>
<td>-0.88 -0.59  4.13</td>
</tr>
<tr>
<td>Mod-Jones(C) model with ROA matching</td>
<td>-0.00  0.00  5.27</td>
<td>0.78  0.63    5.81</td>
<td>-0.87 -0.64  5.71</td>
</tr>
<tr>
<td>Mod-Jones(C) model with ROA+SG matching</td>
<td>-0.00  0.00  5.16</td>
<td>0.02  0.03    5.91</td>
<td>-0.08 -0.06  5.80</td>
</tr>
<tr>
<td>ROA</td>
<td>-0.56  0.90  6.23</td>
<td>-2.26  0.45  8.05</td>
<td>-1.74  0.40  7.50</td>
</tr>
<tr>
<td>Employee growth</td>
<td>9.78  3.45  34.92</td>
<td>33.08 20.00  50.67</td>
<td>-4.89  -5.48  30.57</td>
</tr>
</tbody>
</table>

¹ All accruals are reported in percent form. Thus, an accrual of 0.123 means 0.123% of lagged assets.
Table 3

Correlation matrix between accrual measures and matching variables used in the simulation exercise

The sample of 203,090 Compustat firm-quarters and the calculation of various accrual measures is described in Table 2. That table also gives the definition of ROA, sales growth, and employee growth. The reported numbers are all Spearman rank correlations.

<table>
<thead>
<tr>
<th></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>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Current accruals</td>
<td>1.00</td>
<td>0.83</td>
<td>0.57</td>
<td>0.83</td>
<td>0.57</td>
<td>0.56</td>
<td>0.04</td>
<td>0.14</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>2 Jones model</td>
<td></td>
<td>1.00</td>
<td>0.67</td>
<td>0.67</td>
<td>0.93</td>
<td>0.63</td>
<td>0.63</td>
<td>0.01</td>
<td>0.07</td>
<td>0.05</td>
</tr>
<tr>
<td>3 Jones with ROA matching</td>
<td></td>
<td></td>
<td>1.00</td>
<td>0.49</td>
<td>0.63</td>
<td>0.93</td>
<td>0.46</td>
<td>-0.00</td>
<td>0.05</td>
<td>0.04</td>
</tr>
<tr>
<td>4 Jones with ROA+SG matching</td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
<td>0.62</td>
<td>0.46</td>
<td>0.94</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
</tr>
<tr>
<td>5 Mod-Jones(C) model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
<td>0.67</td>
<td>0.66</td>
<td>0.03</td>
<td>0.15</td>
<td>0.09</td>
</tr>
<tr>
<td>6 Mod-Jones(C) with ROA matching</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
<td>0.49</td>
<td>-0.00</td>
<td>0.10</td>
<td>0.06</td>
</tr>
<tr>
<td>7 Mod-Jones(C) with ROA+SG matching</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
<td>0.00</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>8 ROA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
<td>0.03</td>
<td>0.19</td>
</tr>
<tr>
<td>9 Sales growth</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
<td>0.47</td>
</tr>
<tr>
<td>10 Employee growth</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
</tbody>
</table>
Table 4
Specification tests of discretionary accrual measures using quarterly data

This table reports the percentage of 250 samples of 200 firms each where the null hypothesis of zero discretionary accrual is rejected at the 5% level using one-tailed t-test for mean. These samples are drawn at random from the universe of 203,090 Compustat firm-quarters as described in Table 2. That table also describes the calculation of various accrual measures and partitioning variables. The low and high partitions of any partitioning variable represent the lowest and highest quintiles of the aggregate sample of firm-quarters. We calculate that if the rejection frequency within any one run of 250 samples is below 2.4% or above 8.0%, then it is statistically significantly different from the model rejection frequency of 5% at the 5% confidence level in a two-tailed frequency test.

<table>
<thead>
<tr>
<th></th>
<th>All firms</th>
<th>Sales growth</th>
<th>Employee growth</th>
<th>ROA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Jones model</td>
<td>2.8</td>
<td>41.2</td>
<td>0.0</td>
<td>30.0</td>
</tr>
<tr>
<td>Jones with ROA matching</td>
<td>4.8</td>
<td>30.0</td>
<td>0.8</td>
<td>18.4</td>
</tr>
<tr>
<td>Jones with ROA+SG matching</td>
<td>3.6</td>
<td>7.6</td>
<td>3.6</td>
<td>9.2</td>
</tr>
<tr>
<td>Mod-Jones(C) model</td>
<td>5.2</td>
<td>89.2</td>
<td>0.0</td>
<td>60.0</td>
</tr>
<tr>
<td>Mod-Jones(C) with ROA matching</td>
<td>5.2</td>
<td>66.8</td>
<td>0.0</td>
<td>36.0</td>
</tr>
<tr>
<td>Mod-Jones(C) with ROA+SG matching</td>
<td>4.4</td>
<td>8.0</td>
<td>2.4</td>
<td>9.6</td>
</tr>
</tbody>
</table>

Figures in **bold (bold italic)** signify rejection rates that significantly exceed (fall below) the 5% significance level of the test and indicate that such tests are biased against (in favor) of accepting the null hypothesis.

<table>
<thead>
<tr>
<th></th>
<th>All firms</th>
<th>Sales growth</th>
<th>Employee growth</th>
<th>ROA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Jones model</td>
<td>3.2</td>
<td>0.0</td>
<td>33.6</td>
<td>0.0</td>
</tr>
<tr>
<td>ROA 24.0</td>
<td>0.0</td>
<td>4.0</td>
<td>15.6</td>
<td>9.2</td>
</tr>
<tr>
<td>ROA 2.4</td>
<td>2.8</td>
<td>6.0</td>
<td>10.8</td>
<td>6.4</td>
</tr>
<tr>
<td>Mod-Jones(C) model</td>
<td>2.0</td>
<td>0.0</td>
<td>86.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Mod-Jones(C) with ROA matching</td>
<td>4.0</td>
<td>0.0</td>
<td>58.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Mod-Jones(C) with ROA+SG matching</td>
<td>5.2</td>
<td>3.2</td>
<td>7.8</td>
<td>2.0</td>
</tr>
</tbody>
</table>

Panel A: \( H_1: \) Discretionary accruals < 0

Panel B: \( H_1: \) Discretionary accruals > 0

1 All accruals are reported in percent form. Thus, an accrual of 0.123 means 0.123% of lagged assets.

1
Table 5

**Simulation test results when earnings are managed through revenue manipulation**

This table compares the biases in discretionary accruals estimates resulting from Jones-type models+ ROA matching with the biases resulting from Jones-type models with ROA + (overstated or true) SG matching when the source of earnings management is a mix of sales (revenue) overstatement and expense underestimation. Panel A shows the variations of Jones model, and Panel B shows the variations of Mod-Jones(C) model. In all tests we draw samples of 1000 observations at random such that 50% of observations are from the top SG quintile and the remaining 50% are from the other four SG quintiles. Given space constraints, we denote sales for firm $i$ during quarter $t$ by $S_{i,t}$. Column (2) shows the fixed amount of sales overstatement as a percent of $S_{i,t-4}$, and the subpanel titles show whether this overstatement is for 100%, 40%, or 20% of observations. Thus, in the third row of subpanels A1 and B1, $S'_{i,t} = S_{i,t} + 0.05 \cdot S_{i,t-4}$, $SG'_{i,t} = \frac{S'_{i,t} - S_{i,t-4}}{S_{i,t-4}}$, and $AR'_{i,t} = AR_{i,t} + 0.05 \cdot S_{i,t-4}$ for 100% of observations (i.e., all overstated sales are on credit). A superscript ' attached to any quantity denotes a manipulated value. The resulting accruals overstatement for each observation subject to sales overstatement is calculated as $(1 - \tau)(S'_{i,t} - S_{i,t}) \cdot GM_{i,t}$, where $\tau$ is the marginal corporate tax rate (35%) and $GM_{i,t}$ is the average gross margin for all firms with the same 2-digit SIC code during the same quarter. For the remaining 0%, 60%, or 80% observations the source of earnings management is through expense underestimation. For these observations the accruals are directly overstated by the same amount as the average of all other observations with induced sales overstatement, but without the corresponding sales overstatement. The resulting average accrual overstatement for all observations is shown in Column (3). Given that a researcher observes only $S'_{i,t}$, Columns (4) to (6) report the results of Jones model or Mod-Jones(C) model using this estimate of current sales and one of the three matching procedures described below. First, the models are estimated as follows:

**Jones Model:**

$$WCA_{i,t} = \beta_0 + \beta_1 Q_{1,i,t} + \beta_2 Q_{2,i,t} + \beta_3 Q_{3,i,t} + \beta_4 Q_{4,i,t} + \beta_5 S'_{i,t} + \beta_6 WCA_{i,t-4} + \epsilon_{i,t}.$$ 

**Mod-Jones(C) Model:**

$$WCA_{i,t} = \lambda_0 + \lambda_1 Q_{1,i,t} + \lambda_2 Q_{2,i,t} + \lambda_3 Q_{3,i,t} + \lambda_4 Q_{4,i,t} + \lambda_5 (\Delta S'_{i,t} - \Delta AR'_{i,t}) + \lambda_6 WCA_{i,t-4} + \xi_{i,t}.$$ 

where $\Delta S'_{i,t} = S'_{i,t} - S_{i,t-1}$ and $\Delta AR'_{i,t} = AR'_{i,t} - AR_{i,t-1}$. The remaining variables on right side are defined in Table 2 and Section 2.3. All variables are normalized by lagged assets. Second, the Jones or Mod-Jones(C) model residuals are adjusted by the corresponding residuals for ROA matching, $ROA + SG'$ (overstated sales growth) matching, and $ROA + SG$ (true but unobservable sales growth) matching firms. The detailed matching procedure is also described in Table 2. Columns (7) to (9) report the biases in discretionary accrual measures reported in Columns (4) to (6) by using the true accrual overstatement in Column (3) as the benchmark. The last Column (10) reports the difference between discretionary accruals calculated using $ROA + SG$ (true sales growth) matching and $ROA + SG'$ (overstated sales growth) matching. Thus, this column addresses the question of whether our matching procedure throws the baby out with the bathwater (which is the primary focus of our analysis). All results are based on 1000 simulation runs and the precision of discretionary accrual estimates is the order of 0.005% to 0.010%.
Table 5 continued …

**Panel A: Variations of Jones model**

<table>
<thead>
<tr>
<th>Row number</th>
<th>Seeding process</th>
<th>Discretionary accrual measures</th>
<th>Biases</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Induced sales overstatement, i.e., $S'<em>{lt} - S</em>{lt}$ as a percent of $S_{lt-4}$, which also equals $SG'<em>{lt} - SG</em>{lt}$</td>
<td>Jones with overstated regressor $\Delta S'_{lt}$ and ROA + SG' (overstated sales growth) matching</td>
<td>Difference between (4) and (3)</td>
</tr>
<tr>
<td></td>
<td>Resulting accruals overstatement as a percent of assets, i.e., $WCA'<em>{lt} - WCA</em>{lt}$</td>
<td>Jones with overstated regressor $\Delta S'_{lt}$ and ROA + SG (true sales growth) matching</td>
<td>Difference between (6) and (5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Difference between (6) and (3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Difference between (9) and (7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Difference between (8) and (7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Difference between (10) and (7)</td>
</tr>
<tr>
<td>1.</td>
<td>0.0 0.000</td>
<td>0.118 -0.009 -0.009 0.118 -0.009</td>
<td>0.000</td>
</tr>
<tr>
<td>2.</td>
<td>2.5 0.158</td>
<td>0.250 0.109 0.111</td>
<td>0.002</td>
</tr>
<tr>
<td>3.</td>
<td>5.0 0.316</td>
<td>0.332 0.160 0.194</td>
<td>0.034</td>
</tr>
</tbody>
</table>

**Subpanel A1:** For 100% of firm-quarters $S_{lt}$ is overstated as shown (i.e., source of earnings manipulation is through sales overstatement)

Subpanel A2: For 40% of firm-quarters selected at random $S_{lt}$ is overstated as shown (i.e., source of earnings manipulation is through sales overstatement), for the remaining 60% of firm-quarters $S_{lt}$ is not overstated but $WCA_{lt}$ is directly overstated by 0.316%, 0.158%, or 0.000% of assets (i.e., source of earnings manipulation is through expense understatement)

Subpanel A3: For 20% of firm-quarters selected at random $S_{lt}$ is overstated as shown (i.e., source of earnings manipulation is through sales overstatement), for the remaining 80% of firm-quarters $S_{lt}$ is not overstated but $WCA_{lt}$ is directly overstated by 0.316%, 0.158%, or 0.000% of assets (i.e., source of earnings manipulation is through expense understatement)
Table 5 continued …

**Panel B: Variations of Mod-Jones(C) model**

<table>
<thead>
<tr>
<th>Row number</th>
<th>Seeding process</th>
<th>Discretionary accrual measures</th>
<th>Biases</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Induced sales overstatement, i.e., $S'<em>{it-4} - S</em>{it}$ as a percent of $S_{it}$, which also equals $SG'<em>{it} \times SG</em>{it}$</td>
<td>Mod-Jones(C) with overstated regressor $(\Delta S'<em>{it} - \Delta AR'</em>{it})$ and ROA matching</td>
<td>Difference between (4) and (3)</td>
</tr>
<tr>
<td></td>
<td>Resulting accruals overstatement as a percent of assets, i.e., $WCA'<em>{it} - WCA</em>{it}$</td>
<td>Mod-Jones(C) with overstated regressor $(\Delta S'<em>{it} - \Delta AR'</em>{it})$ and $ROA + SG'_{it}$ (overstated sales growth) matching</td>
<td>Difference between (5) and (3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mod-Jones(C) with overstated regressor $(\Delta S'<em>{it} - \Delta AR'</em>{it})$ and $ROA + SG_{it}$ (true sales growth) matching</td>
<td>Difference between (6) and (3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Difference between (6) and (5)</td>
</tr>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>10.</td>
<td>0.0</td>
<td>0.000</td>
<td>0.298</td>
</tr>
<tr>
<td>11.</td>
<td>2.5</td>
<td>0.158</td>
<td>0.507</td>
</tr>
<tr>
<td>12.</td>
<td>5.0</td>
<td>0.316</td>
<td>0.655</td>
</tr>
<tr>
<td>Subpanel B1: For 100% of firm-quarters $S_{it}$ is overstated as shown (i.e., source of earnings manipulation is through sales overstatement)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13.</td>
<td>0.0</td>
<td>0.000</td>
<td>0.298</td>
</tr>
<tr>
<td>14.</td>
<td>2.5</td>
<td>0.158</td>
<td>0.506</td>
</tr>
<tr>
<td>15.</td>
<td>5.0</td>
<td>0.316</td>
<td>0.654</td>
</tr>
<tr>
<td>Subpanel B2: For 40% of firm-quarters selected at random $S_{it}$ is overstated as shown (i.e., source of earnings manipulation is through sales overstatement), for the remaining 60% of firm-quarters $S_{it}$ is not overstated but $WCA_{it}$ is directly overstated by 0.316%, 0.158%, or 0.000% of assets (i.e., source of earnings manipulation is through expense understatement)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16.</td>
<td>0.0</td>
<td>0.000</td>
<td>0.298</td>
</tr>
<tr>
<td>17.</td>
<td>2.5</td>
<td>0.158</td>
<td>0.506</td>
</tr>
<tr>
<td>18.</td>
<td>5.0</td>
<td>0.316</td>
<td>0.654</td>
</tr>
<tr>
<td>Subpanel B3: For 20% of firm-quarters selected at random $S_{it}$ is overstated as shown (i.e., source of earnings manipulation is through sales overstatement), for the remaining 80% of firm-quarters $S_{it}$ is not overstated but $WCA_{it}$ is directly overstated by 0.316%, 0.158%, or 0.000% of assets (i.e., source of earnings manipulation is through expense understatement)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes:

1. The numbers reported in the second row of Column (10) are always less than half of the numbers reported in the third row of the same column in all subpanels even though the induced sales overstatement is exactly 2.5% in the second row and 5.0% in the third row. This is a discreteness issue. In many cases with 2.5% sales overstatement in the second row the matching procedure ends up picking the same firm as with no sales overstatement in the first row. Hence, there is no difference between $ROA + SG$ and $ROA + SG'$ matching (given the sparse population of firms within any quarter for a given 2-digit SIC code).

2. Although each of the two terms in the regressor $(\Delta S'_{it} - \Delta AR'_{it})$ for Mod-Jones(C) is overstated by 0.05 $\times S_{it-4}$, the difference is not overstated.
Table 6
Comparison of power of ROA + SG matching methodologies and Dechow et al. (2011) reversal methodology
in quarterly setting

This table reports the percentage of 250 samples of N = 600, 1000, or 2000 firms each where the null hypothesis of zero discretionary accrual is rejected at the 5% level using one-tailed t-test for mean and Wilcoxon signed-rank test for median. These samples are drawn at random from the universe of 203,090 Compustat firm-quarters as described in Table 2. That table also describes the calculation of Jones and Mod-Jones(C) methodologies with ROA + SG matching. The t-test for mean and the Wilcoxon signed rank test for median are further described in Figure 6. The Dechow et al. (2011) reversals methodology is described as follows. We pick N = 600, 1000, or 2000 firm-quarters at random from the universe of 203,090 firm-quarters and increase their raw accruals by 0.25%. Following the evidence in Baber, Kang, and Li (2011), we assume reversals take place over the subsequent quarters with the indicated frequencies: one quarter (43%), two quarters (29%), three quarters (21%), and four quarters (7%). Further, if reversals take place over n quarters, we assume that 1/nth of the total reversal occurs each quarter from 1 to n, following the earnings management quarter. Finally, we assume that the following proportion of the original earnings management reverses during the specified reversal horizon: 100%, 50%, and 30%. We mix these seeded quarters and reversal quarters with the remaining sample of firm-quarters and carry out the following regression:

\[ WCA_{it} = a + b \text{PART}_{it} + c \text{PARTR1}_{it} + d \text{PARTR2}_{it} + e \text{PARTR3}_{it} + f \text{PARTR4}_{it} + \text{usual Jones - type model terms} + \epsilon_{it} \]  

\[ \text{(T5.1)} \]

\( \text{PART} \) is a dummy variable that takes the value one for the earnings management quarter and zero otherwise. Similarly, \( \text{PARTR1} - \text{PARTR4} \) are dummy variables that take the value one for each of the next four quarters and zero otherwise. The subscripts \( i \) and \( t \) denote the firm and the quarter. The usual model terms are the same as included in equations (T2.1) and (T2.2) of Table 2 for Jones and Mod-Jones(C) models. We repeat the procedure 250 times and record the frequency with which the null hypothesis of \( b - (c + d + e + f) = 0 \) can be rejected in favor of the alternate hypothesis of \( b - (c + d + e + f) > 0 \).

<table>
<thead>
<tr>
<th>Model/methodology</th>
<th>N = 600</th>
<th>N = 1000</th>
<th>N = 2000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jones with ROA+SG matching – mean test</td>
<td>33.6</td>
<td>35.6</td>
<td>63.2</td>
</tr>
<tr>
<td>Jones with ROA+SG matching – median test</td>
<td>37.6</td>
<td>42.8</td>
<td>75.6</td>
</tr>
<tr>
<td>Jones with reversals methodology – 100% reversal</td>
<td>18.0</td>
<td>28.0</td>
<td>39.6</td>
</tr>
<tr>
<td>Jones with reversals methodology – 50% reversal</td>
<td>11.2</td>
<td>16.8</td>
<td>24.0</td>
</tr>
<tr>
<td>Jones with reversals methodology – 30% reversal</td>
<td>8.8</td>
<td>13.6</td>
<td>19.2</td>
</tr>
<tr>
<td>Mod-Jones(C) with ROA+SG matching – mean test</td>
<td>32.0</td>
<td>37.2</td>
<td>63.2</td>
</tr>
<tr>
<td>Mod-Jones(C) with ROA+SG matching – median test</td>
<td>38.8</td>
<td>44.4</td>
<td>75.6</td>
</tr>
<tr>
<td>Mod-Jones(C) with reversals methodology – 100% reversal</td>
<td>22.0</td>
<td>28.8</td>
<td>43.6</td>
</tr>
<tr>
<td>Mod-Jones(C) with reversals methodology – 50% reversal</td>
<td>12.4</td>
<td>20.8</td>
<td>27.6</td>
</tr>
<tr>
<td>Mod-Jones(C) with reversals methodology – 30% reversal</td>
<td>10.8</td>
<td>16.8</td>
<td>23.6</td>
</tr>
</tbody>
</table>