Detecting Earnings Management: A New Approach*

Patricia M. Dechow
The Haas School of Business
University of California, Berkeley
Berkeley, CA 94705
Patricia_dechow@haas.berkeley.edu

Amy P. Hutton
Carroll School of Management
Boston College
Chestnut Hill, MA 02467
Amy.Hutton.1@bc.edu

Jung Hoon Kim
The College of Business Administration
Florida International University
Miami, FL 33199
junghoon.kim@fiu.edu

and

Richard G. Sloan
The Haas School of Business
University of California, Berkeley
Berkeley, CA 94705
Richard_Sloan@haas.berkeley.edu

This Version: October 2011

*We are grateful for the comments of an anonymous referee, Phil Berger (the Editor), Frank Ecker, Jennifer Francis, Joseph Gerakos, Maureen McNichols, Per Olsson, Katherine Schipper and workshop participants at the 2011 Journal of Accounting Research Conference, the 2011 American Accounting Association Annual Meetings, the University of Arizona, Brigham Young University, the University of Houston, the University of Texas at Austin, the University of Washington and UCLA.
Detecting Earnings Management:

A New Approach

This paper provides a new approach to testing for accrual-based earnings management. Our approach exploits the inherent property of accrual accounting that any accrual-based earnings management in one period must reverse in another period. If the researcher has priors concerning the timing of the reversal, incorporating these priors can significantly improve the power and specification of tests for earnings management. Our results indicate that incorporating reversals can increase test power by over 40% and also provides a robust solution for mitigating model misspecification arising from correlated omitted variables.
1. Introduction

Earnings management is an important accounting issue for academics and practitioners alike.\textsuperscript{1} A large body of academic research examines the causes and consequences of earnings management.\textsuperscript{2} A major limitation of this research is that existing techniques for measuring earnings management lack power and are often misspecified. The most common techniques for measuring earnings management attempt to isolate the ‘discretionary’ portion of the accrual component of earnings. The limitations of such techniques are enumerated in Dechow, Sloan and Sweeney [1995]. The techniques lack power for earnings management of plausible magnitudes because of the poor ability of the models to isolate discretionary accruals. Moreover, tests using these techniques are misspecified due to correlated omitted variables in samples with extreme financial performance, a situation that is not uncommon in tests for earnings management. Few improvements have been made since Dechow, Sloan and Sweeney [1995] (DSS hereafter). Alternative techniques have been proposed for identifying discretionary accruals (e.g., Dechow and Dichev [2002]), but whether they represent improvements is questionable (e.g., Wysocki [2009]). Performance matching procedures have been adopted to mitigate misspecification (e.g., Kothari, Leone and Wasley [2005]), but cause substantial reductions in test power and are only effective when the matching procedure employs the relevant omitted variable.

In this paper, we propose a new approach for the detection of earnings management that simultaneously improves test power and specification. Our approach exploits an inherent characteristic of accrual-based earnings management that has gone largely ignored in previous research. Specifically, we recognize that any accrual-based earnings management in one period must reverse in another period.\textsuperscript{3} If the researcher has reasonable priors concerning the period(s) in which the hypothesized earnings management is expected to reverse, the power and specification of tests for earnings management can be significantly improved by incorporating

\textsuperscript{1} Dechow and Skinner [2000] review the earnings management literature, discuss the prevalence of earnings management and provide both academic and practitioner perspectives on earnings management.

\textsuperscript{2} Dechow, Ge and Schrand [2010] provide a recent review of this research.

\textsuperscript{3} A growing body of research examines the properties and pricing of accrual reversals (e.g., Defond and Park [2001], Allen, Larson and Sloan [2010], Baber, Kang and Li [2011], Fedyk, Singer and Sougiannis [2011]). To our knowledge, ours is the first study to formally develop and evaluate techniques for measuring earnings management that incorporate accrual reversals.
these reversals. For example, if the researcher correctly identifies the periods in which earnings management originates and reverses, incorporating reversals essentially doubles the amount of variation in discretionary accruals that is attributable to the hypothesized earnings management. Our calibrations suggest doing so increases test power by around 40% in typical earnings management studies.\(^4\)

In addition to improving test power, incorporating reversals in tests of earnings management also mitigates misspecification arising from correlated omitted variables. For example, firm size has been identified as potentially important correlated omitted variable in tests for earnings management (e.g., Ecker, Francis, Olsson and Schipper [2011]). In particular, large firms tend to have lower nondiscretionary accruals because they have lower growth prospects. Thus, a researcher testing for evidence of downward earnings management in a sample of firms that are large relative to the control firms may incorrectly conclude that earnings are being managed downward. This problem is mitigated by incorporating reversals, since the researcher would also test for evidence of higher accruals arising from the reversal of the earnings management in an adjacent period. But because firm size is a persistent economic characteristic, the related accruals would also be lower in adjacent periods. This would inform the researcher that the lower accruals are not attributable to earnings management. Note that the researcher does not have to identify the relevant omitted variables for tests incorporating reversals to mitigate the associated misspecification. As long as the omitted variables do not completely reverse in the same period as the earnings management, test specification is improved.\(^5\) We show that incorporating accrual reversals provides a robust solution to mitigating misspecification across a variety of economic characteristics.

Our paper proceeds as follows. First we develop an econometric framework to summarize common tests for earnings management and highlight their associated problems. Next, we introduce a flexible procedure for addressing these problems by incorporating

---

\(^4\) Intuitively speaking, incorporating earnings management reversals in tests of earnings management is equivalent to doubling the number of observations for which earnings management occurs.

\(^5\) Note that if the correlated omitted variable is highly persistent, then incorporating reversals completely eliminates the associated correlated omitted variable problem. On the other hand, if the correlated omitted variable is completely transitory, incorporating reversals only partially eliminates the associated correlated omitted variable problem. It is only in cases where the correlated omitted variable is highly negatively serially correlated that incorporating reversals would not improve test specification. For further details, see section 2.4.
researchers’ priors concerning the reversal of discretionary accruals in tests for earnings management. This procedure requires the researcher to identify the period(s) in which accruals are predicted to be managed and the period(s) in which this accruals management is predicted to reverse. A standard test for joint significance is then used to test for earnings management. Our procedure can be adapted to all common models of nondiscretionary accruals.

We next evaluate the power and specification of tests incorporating accrual reversals relative to the traditional \( t \)-test on discretionary accruals and to the Kothari, Leone and Wasley’s [2005] \( t \)-test on performance-matched discretionary accruals. Following DSS, we use four sets of analyses to evaluate the competing tests for earnings management. First, we evaluate the specification of the tests in samples of historical data using randomly assigned earnings management years. We show that all tests, including our new tests, are reasonably well specified in random samples. Second, we conduct simulations using archival data and seeded earnings management to examine how incorporating discretionary accrual reversals enhances the power of tests for earnings management. The simulations indicate that if the researcher’s priors about the reversal year are as accurate as the priors about the earnings management year, then incorporating the reversal increases test power by over 40%. These simulations also show that test power increases even in cases where the researcher is less than half as accurate in identifying the reversal year relative to the earnings management year. Third, we evaluate the power of the tests in a sample where the SEC alleges firms have overstated earnings. We find that incorporating accrual reversals in the period(s) following the alleged overstatements substantially increases the power of tests for earnings management in this sample. The gains in power dwarf any differences arising from choosing between different models of nondiscretionary accruals. Fourth, we evaluate the specification of tests in samples of historical data with extreme economic characteristics. We show that standard \( t \)-tests are highly misspecified and that performance matching only mitigates misspecification when the matching procedure employs the relevant omitted variable. For example, the commonly followed Kothari, Leone and Wasley’s [2005] procedure of matching on return on assets (ROA hereafter) mitigates misspecification in samples with extreme ROA but exaggerates misspecification in samples with extreme firm size. In contrast, our new tests, which incorporate accrual reversals, are robust in mitigating misspecification across a range of economic characteristics.
Overall, our results suggest that our new approach for detecting earnings management leads to substantial improvements in both test power and specification. We therefore encourage subsequent earnings management research to consider this approach. Our reversal framework should also be useful for practitioners interested in establishing the existence of earnings management in historical data. Examples include regulators tasked with enforcing accounting principles, investors evaluating the quality of managements’ past financial reports and class action lawyers establishing cases of fraud on the market.

The remainder of our paper is organized as follows. Section 2 reviews and evaluates existing techniques for detecting earnings management and motivates our extension to incorporate accrual reversals. Section 3 describes the earnings management models and tests that we consider in this paper. Section 4 describes our research design, section 5 presents our results and section 6 concludes.

2. Review and Motivation

2.1 Statistical Framework

This section develops a statistical framework for summarizing common tests of earnings management and identifying potential misspecifications associated with these tests. It builds on the framework introduced by McNichols and Wilson [1988] and extended by DSS. Accrual-based tests of earnings management are based on the following linear model:

$$ DA_{i,t} = a + bPART_{i,t} + \varepsilon_{i,t} $$

(1)

where

DA = discretionary accruals;
PART = a dummy variable that is set to 1 in periods during which a hypothesized determinant of earnings management is present and 0 otherwise; and
\( \varepsilon \) = the impact of other determinants of earnings management.

Invoking the standard OLS assumptions, the OLS estimator of \( b \), denoted \( \hat{b} \), is the best linear unbiased estimator with standard error:
\[
\text{SE}(\hat{b}) = \frac{s_e}{\sqrt{(n-1)s_{\text{PART}}}}
\]

where

\(n\) = number of observations;

\(s_e\) = standard error of the regression (residual sum of squares divided by n-2); and

\(s_{\text{PART}}\) = sample standard deviation of PART.

The ratio of \(\hat{b}\) to \(\text{SE}(\hat{b})\) has a \(t\)-distribution with \(n-2\) degrees of freedom. The null hypothesis of no earnings management is rejected if \(\hat{b}\) has the hypothesized sign and is statistically significant at conventional levels. Consequently, the expected \(t\)-statistic and hence the power of the associated test for earnings management is increasing in:

- \(\hat{b}\), the magnitude of the earnings management;
- \(n\), the number of observations; and
- \(s_{\text{PART}}\), the standard deviation of PART. Note that if we let \(\rho\) denote the proportion of the observations for which PART=1, then \(s_{\text{PART}} = \sqrt{\rho(1-\rho)}\), so \(s_{\text{PART}}\) is greatest when \(\rho = 0.5\) and gradually declines to zero as \(\rho\) approaches either zero or one.

while power is decreasing in:

- \(s_e\), the standard error of the regression (residual sum of squares divided by n-2).

Unfortunately, the researcher does not directly observe discretionary accruals, and so must use a discretionary accrual proxy that measures discretionary accruals (DA) with error:

\[
DAP_{i,t} = (DA_{i,t} - \mu_{i,t}) + \eta_{i,t}
\]

where

\(\mu\) = discretionary accruals that are unintentionally removed from DAP;

\(\eta\) = nondiscretionary accruals that are unintentionally left in DAP.

To understand the resulting misspecification, we first substitute DAP for DA in equation 1:

\[
DAP_{i,t} = a + b\text{PART}_{i,t} + (\mu_{i,t} + \eta_{i,t} + \epsilon_{i,t})
\]

(1)’
The OLS estimator of \( b \) obtained from regressing DAP on PART, denoted \( \hat{b} \), is misspecified by the omission from the regression of \((-\mu + \eta)\). In particular, \( \hat{b} \) is a biased estimator of \( b \), with bias given by:

\[
E(\hat{b}) - b = \beta(-\mu + \eta)(\text{PART})
\]

where

\( \beta(-\mu + \eta)(\text{PART}) \) = regression coefficient in a regression of \((-\mu+\eta)\) on PART; and

\( E(.) \) = the expectations operator.

Also, the OLS standard error for \( \hat{b} \), is given by:

\[
\text{SE}(\hat{b}) = \text{SE}(\hat{b})(1-r^2(-\mu + \eta)(\text{PART}))/(1-r^2(DAP)(-\mu + \eta)(\text{PART}))
\]

where

\( r^2(-\mu + \eta)(\text{PART}) \) = the r-squared from a regression of \((-\mu+\eta)\) on PART; and

\( r^2(DAP)(\text{DAP})(-\mu + \eta)(\text{PART}) \) = the r-squared from a regression of DAP on the component of \((-\mu+\eta)\) that is orthogonal to PART.

The above expressions highlight three distinct types of misspecification that can arise in the estimation of (1)’:

**Problem 1. Bias and loss of power caused by the omission of \(-\mu\) from DAP.** Recall that \( \mu \) represents discretionary accruals that have been unintentionally removed from DAP. The presence of \( \mu \) causes bias in \( \hat{b} \) that is bounded by \(-b \leq \beta(-\mu)(\text{PART}) \leq 0\). This means that \( \hat{b} \) is biased toward zero, with the limiting case where we unintentionally remove all discretionary accruals resulting in \( \hat{b} = 0 \). This bias reduces the likelihood of rejecting the null hypothesis of no earnings management when it is false (i.e., increased type II error rate). Intuitively, removing some of the discretionary accruals results in a less powerful test (we have ‘thrown the

---

7 The inequality assumes \( b>0 \). For \( b<0 \), \( 0 \leq \beta(-\mu)(\text{PART}) \leq -b \).
Problem 2. Bias and misspecification caused by the inclusion of correlated $\eta$ in DAP.

Recall that $\eta$ represents nondiscretionary accruals that have been unintentionally left in DAP. The presence of $\eta$ biases $\tilde{b}$ so long as $\eta$ is correlated with PART. In particular, $\tilde{b}$ will not equal zero even when $b=0$. This increases the likelihood of rejecting the null hypothesis of no earnings management even when it is true (i.e., excessive type I error rate). Intuitively, we mistakenly infer the presence of earnings management because of nondiscretionary accruals that happen to be correlated with PART.

Problem 3. Inefficiency caused by the inclusion of uncorrelated $\eta$ in DAP. If nondiscretionary accruals are left in DAP, but they are uncorrelated with PART, $\tilde{b}$ is unbiased. However, $\text{SE}(\tilde{b}) = \text{SE}(\hat{b} \mid 1-r^2_{(DAP)(\eta)})$. So the standard error of the estimated coefficient is increasing in the proportion of the variation in DAP that is attributable to $\eta$. This reduces the likelihood of rejecting the null hypothesis of no earnings management when it is false (i.e., increased type II error rate). Intuitively, failing to extract nondiscretionary accruals from DAP results in a less powerful test even when these nondiscretionary accruals are uncorrelated with PART.

Balancing these competing sources of misspecification presents a trade-off. Incorporating every conceivable determinant of nondiscretionary accruals is likely to exacerbate Problem 1. But incorporating too few determinants exacerbates Problem 2 and Problem 3.

2.2 Overview of Discretionary Accrual Models

Following Healy [1985], most discretionary accrual models start with working capital accruals as their base measure of accruals. Early research then simply employs the levels (e.g., Healy [1985]) or changes (e.g., DeAngelo [1986]) in working capital accruals as discretionary accrual proxies, implicitly assuming that nondiscretionary accruals are constant. This assumption is unlikely to be empirically descriptive, because nondiscretionary accruals are expected to change with firms’ underlying business activities (e.g., Kaplan [1985], McNichols
Subsequently, more sophisticated models that attempt to explicitly model nondiscretionary accruals have been developed, enabling total accruals to be decomposed into discretionary and nondiscretionary components. The most popular models are attributable to Jones [1991], DSS, Dechow and Dichev [2002] and McNichols [2002]. We will describe these models in more detail in the next section. Such models typically require at least one parameter to be estimated, and were originally implemented through the use of a firm-specific ‘estimation period,’ during which no systematic earnings management was hypothesized. Starting with Defond and Jiambalvo [1994], researchers have generally employed cross-sectional and panel estimation of these models.

Concerns that these models fail to capture all nondiscretionary accruals have also led researchers to supplement the models with performance-matching procedures. Kothari, Leone and Wasley [2005] propose a popular matching procedure that entails subtracting estimates of discretionary accruals from Jones type models using control firms matched by industry and ROA in either the current or the previous period.

2.3 Limitations of Existing Models

While various models described above have been used extensively in the literature to test for earnings management, their effectiveness is known to be limited. DSS provide the first comprehensive assessment of the specification and power of commonly used test statistics across the measures of discretionary accruals generated by several of these models. They conclude that: (i) all of the models generate well-specified test statistics when applied to random samples; (ii) all models generate tests of low power for earnings management of economically plausible magnitudes (e.g. one to five percent of total assets); and (iii) all models are misspecified when applied to samples of firms with extreme financial performance. McNichols [2000] reiterates point (iii) and shows that all models are particularly misspecified for samples with extreme forecasts of long-term earnings growth.

Kothari, Leone and Wasley [2005] propose the performance matching procedure mentioned earlier to mitigate performance-related misspecification. Their results indicate that
performance matching is no panacea. First, their performance matching procedure rarely eliminates misspecification and sometimes exaggerates misspecification. For example, their results indicate that performance matching on ROA mitigates misspecification for samples with extreme earnings-to-price and book-to-market, but can exaggerate misspecification in samples with extreme size and operating cash flows. Second, their results highlight the low power of existing tests for earnings management and show that performance matching exacerbates this problem. For example, their simulations show that using random samples of 100 firm-years, seeded earnings management equal to 1% of total assets and a 5% test level results in rejection rates of only 20 percent with no performance matching and a paltry 14 percent with performance matching.

These results highlight two key limitations of performance matching procedures. First, performance matching is only effective in mitigating misspecification when the researcher matches on the relevant correlated omitted variable. Second, performance matching reduces test power by increasing the standard error of the test statistic. We can use the framework developed in section 2.1 to formalize these problems. Using the subscript j for the matched control firm, the resulting performance matched discretionary accrual proxy is:

\[ DAP_{i,t} - DAP_{j,t} = (DA_{i,t} - DA_{j,t}) - (\mu_{i,t} - \mu_{j,t}) + (\eta_{i,t} - \eta_{j,t}) + (\epsilon_{i,t} - \epsilon_{j,t}) \]

If we have chosen a perfect match, then \((\eta_{i,t} - \eta_{j,t})=0\). But even in this case, we have introduced two new problems. First, it is possible that \(DA_{i,t}\) and \(DA_{j,t}\) will be positively correlated. This seems particularly likely when matching on ROA, because DA is a component of ROA. This will generate a special case of Problem 1, removing discretionary accruals and reducing the power of the test in the presence of earnings management. Second, assuming the \(\epsilon\)’s are independently and identically distributed, the new standard error of the regression will be \(\sqrt{2}s_\epsilon\), causing the \(t\)-statistic to be reduced accordingly. This problem is similar to Problem 3, in that it introduces additional uncorrelated noise into the residual leading to a less powerful test.
2.4 Incorporating Discretionary Accrual Reversals

We introduce a new approach for the detection of earnings management that has the potential to simultaneously improve test power and mitigate misspecification. Our approach exploits an inherent property of discretionary accruals. Discretionary accruals are made with the purpose of shifting earnings between reporting periods. The accrual accounting process requires misstatements in one period to be reversed in another period. For example, if a firm overstates its receivables in one period, the overstatement must be reversed in the subsequent period when it is determined that the associated cash flows will not be received.

Nondiscretionary accruals, in contrast, are tied to the operations of the underlying business (e.g., McNichols [2000]). At an aggregate level, they will tend to originate during periods when the business is either growing (e.g., purchasing inventory in anticipation of future sales growth) or making strategic changes to its operating and investing decisions (e.g., granting more generous credit terms). Since businesses operate as going concerns, their operating characteristics tend to persist. As such, the associated nondiscretionary accruals should also tend to persist. In other words, while specific nondiscretionary accruals must reverse, reversing accruals are likely to be replaced by originating accruals in businesses that are going concerns (e.g., replacement of inventory as it is sold) such that nondiscretionary accruals will tend to persist at the aggregate level.

Because discretionary accruals should reverse while nondiscretionary accruals should persist, we can test for earnings management not only by testing for the presence of discretionary accruals in the earnings management period, but also by testing for the reversal of those accruals in an adjacent period. Incorporating reversals should both increase test power and eliminate misspecification associated with correlated nondiscretionary accruals included in DAP. We can use the framework developed in section 2.1 to understand the intuition behind these improvements. To do so, we introduce two new earnings management partitioning variables:

\[ \text{PARTR} = \text{a dummy variable that equals 1 in periods during which the earnings management is hypothesized to reverse; and} \]

\[ \text{PART'} = \text{PART – PARTR (i.e., PART'} = 1 \text{ for earnings management years, -1 for reversal} \]
years and 0 otherwise).

We also make the additional assumptions that (i) we correctly measure DA, (ii) we correctly identify the earnings management and accrual reversal periods; and (iii) the earnings management and reversal periods are mutually exclusive. Consider estimating equation (1) after replacing PART with PART', denoting the corresponding regression estimates by $b'$ and $\text{SE}(b')$ respectively. Recall that $s_{\text{PART}} = \sqrt{(\rho - \rho^2)}$ and we can readily determine that $s_{\text{PART}'} = \sqrt{2\rho}$. Thus, as long as $\rho$ is ‘small’, we obtain:

$$E(b') \approx E(\hat{b})$$

$$s_{\text{PART}'} \approx \sqrt{2} s_{\text{PART}}$$

$$\text{SE}(b') \approx \text{SE}(\hat{b})/\sqrt{2}$$

Intuitively, since the magnitude of the earnings management reversals must equal the magnitude of the initial earnings management, the estimated coefficient on the earnings management partitioning variable is not expected to change. At the same time, the standard deviation of the earnings management partitioning variable increases by approximately $\sqrt{2}$, causing the standard error of the estimated coefficient on the earnings management partitioning variable to be reduced by approximately $\sqrt{2}$. By reducing the standard error of the coefficient estimate, we increase test power.

One can think of this increase in test power described above as being analogous to the case of doubling the number of observations where PART=1. This brings us to the issue of why we need to assume that $\rho$ is ‘small’. Recall that $\rho$ is the proportion of the observations where earnings are managed, and since we assume that earnings management and reversal periods are mutually exclusive, $\rho$ has a maximum possible value of 0.5. But in the case where $\rho =0.5$, every observation is either an earnings management period or a reversal period. Thus, modeling reversals doesn’t improve test power, because we have already implicitly identified the reversal periods by identifying the earnings management periods. Formally, in the case where $\rho=0.5$, we obtain:
\[ E(b') = E(\hat{b})/2 \]

\[ s_{\text{PART}'} = 2s_{\text{PART}} \]

\[ \text{SE}(b') = \text{SE}(\hat{b})/2 \]

In this case, modeling reversals reduces the estimated coefficient on PART by a factor of 2, but also reduces the standard error of the coefficient estimate by a factor of 2, thus having no net impact on test power. In fact, because \( \text{PART}' = -1 \) in every observation where \( \text{PART} = 0 \), \( \text{PART}' \) and PART will be perfectly positively correlated, with \( \text{PART}' = 2(\text{PART}-1) \). For example, if earnings are managed up by 5 in half of the sample and therefore managed down by 5 in the other half, we have \( E[b'] = 5 \) and \( s_{\text{PART}'} = 1 \), while \( E[\hat{b}] = 10 \) and \( s_{\text{PART}} = 1/2 \). So in using \( \text{PART}' \), what we gain in terms of a higher \( s_{\text{PART}'} \), we lose in terms of a lower \( \hat{b} \).

The preceding discussion begs the question of \( \rho \)'s magnitude in typical earnings management studies. In many earnings management studies, earnings management is only hypothesized to occur in a small proportion of the available observations. For example, in our analysis of firms subject to enforcement actions by the SEC, to be presented later in the paper, there are 406 firm-years with earnings management allegations out of a total sample of 161,119 firm-years. Thus \( \rho = 406/161,119 = 0.003 \). To give further examples from some highly cited earnings management studies, \( \rho = 0.020 \) in Defond and Jiambalvo’s [1994] study of debt covenant violations, \( \rho = 0.012 \) in Defond and Subramanyam’s [1998] study of auditor changes and \( \rho = 0.047 \) in Ball and Shivakumar’s [2008] study of IPOs. Thus successfully incorporating reversals would have substantially increased tests power in all of these studies. Nevertheless, researchers should be aware that the improvements in test power from incorporating reversals disappear in settings where \( \rho \) approaches 0.5.

We next turn to the impact of incorporating reversals on test specification in the presence of correlated omitted nondiscretionary accruals (Problem 2). Recall from section 2.1 that the presence of correlated omitted nondiscretionary accruals, \( \eta \), in DAP biases the estimate of earnings management (\( \tilde{b} \)) as follows:
\[ E(\beta - b) = \beta_{(\eta)(\text{PART})} \]

where

\[ \beta_{(\eta)(\text{PART})} = \text{regression coefficient in a regression of } \eta \text{ on PART; and} \]
\[ E(.) = \text{the expectations operator.} \]

The impact of modeling reversals on this bias therefore hinges on the impact of substituting PART’ for PART on \( \beta_{(\eta)(\text{PART})} \). First, note that \( \beta_{(\eta)(\text{PART})} = \beta_{(\eta)(\text{PART}')} \) only in the special case that \( \eta \) completely reverses in the reversal period.\(^8\) As discussed earlier, the economic characteristics driving nondiscretionary accruals tend to persist, leading us to expect that the associated nondiscretionary accruals should also persist. So a second special case of interest is when \( \eta \) completely persists into the reversal period. In this case, \( \beta_{(\eta)(\text{PART}')} = 0 \), because for every observation where PART’=1, we now have another observation for which PART’=-1 and \( \eta \) is the same. It is therefore clear that \( \eta \) cannot be linearly related to PART’. Thus, incorporating reversals in tests of earnings management completely eliminates Problem 2 when the nondiscretionary accruals completely persist into the reversal period. More generally, incorporating accrual reversals will mitigate Problem 2 to the extent that the associated nondiscretionary accruals do not completely reverse in the reversal period. Intuitively, by modeling discretionary accrual reversals, we reduce the likelihood of mistakenly attributing earnings management to nondiscretionary accruals that are correlated with PART but don’t completely reverse.

To summarize, incorporating reversals into tests of earnings management produces two potential benefits:

(i) So long as less than half of the total available sample for firm years is hypothesized to be managing earnings, incorporating reversals increases test power.

(ii) So long as any correlated omitted nondiscretionary accruals do not reverse in the same period that the earnings management reverses, incorporating

\(^8\) Note that with respect to Problem 1, we expect a complete reversal of -\( \mu \), because these are discretionary accruals that we missed when they originated, and so we also expect to miss their reversal. Thus, incorporating reversals does not mitigate Problem 1.
reversals mitigates correlated omitted variables bias.

3. Test Design

This section describes our framework for incorporating accrual reversals in tests of earnings management and summarizes key features of the nondiscretionary accrual models employed in our tests. Our primary objective is to examine the impact of incorporating accrual reversals on existing tests for earnings management. We therefore strive to keep other features of our testing framework consistent with prior research.

3.1 Test Procedure

We implement equation (1)’ as follows:

\[ WC_{ACC_{i,t}} = a + b\text{PART}_{i,t} + \sum_{k} f_{k} X_{k,i,t} + e_{i,t} \] (3)

where

\( WC_{ACC} \) = non-cash working capital accruals;

\( \text{PART} \) = a dummy variable that is set to 1 in periods during which a hypothesized determinant of earnings management is present and 0 otherwise; and

\( X_{k} \) = controls for nondiscretionary accruals.

Note that following DSS, we use working capital accruals as our base measure of accruals and directly include controls for nondiscretionary accruals as additional explanatory variables in the earnings management regression.

To incorporate reversals, we augment (3) through the inclusion of a second partitioning variable that identifies periods in which the earnings management is hypothesized to reverse (PARTR):

\[ WC_{ACC_{i,t}} = a + b\text{PART}_{i,t} + c\text{PARTR}_{i,t} + \sum_{k} f_{k} X_{k,i,t} + e_{i,t} \] (4)
We then test the linear restriction that $b - c = 0$ to test for earnings management. The alternative hypotheses for upward (downward) earnings management are $b - c > (\leq) 0$. While the assumption that earnings management reverses in one year is reasonable for working capital accruals, it is not the only possible assumption. For example, if earnings are hypothesized to be managed upward during equity offerings, one might reasonably hypothesize that such earnings management would not reverse until after sufficient time has passed that management and investment banker lock-up agreements have expired. Thus, PARTR need not always take on the value of 1 in the period immediately following that in which PART=1.

For the purpose of conducting our evaluation of model (4), we consider three scenarios regarding the timing of the reversal of earnings management. In the first scenario, we assume that the researcher has no priors regarding the reversal of the earnings management, thus ignoring the coefficient on PARTR altogether. This scenario essentially collapses to the traditional model in equation (3). In the second scenario, we assume that all earnings management reverses in the year immediately following the earnings management year. This seems to be a plausible assumption when considering working capital accruals, since most working capital accruals are expected to reverse within a year. However, since it is also possible that managers have the incentives and the ability to delay accrual reversals beyond one year, we also consider a third scenario in which we assume that all earnings management reverses over the two years following the earnings management year. Under these latter two scenarios, if earnings are hypothesized to be managed for two or more consecutive years, we assume that the reversal starts in the first year following the last of the consecutive earnings management years.

To facilitate interpretation of the results for the second scenario, we decompose PARTR into two new partitioning variables, PARTP1 and PARTP2, where PARTP1 equals 1 in the first year following an earnings management year and 0 otherwise and PARTP2 equals 1 in the second year following an earnings management year and zero otherwise:

Note that this testing procedure differs slightly from the one described in section 2.4, where we create a single new earnings management partitioning variable, $\text{PART}'=\text{PART}-\text{PARTR}$. Using $\text{PART}'$ simplifies the analytics in section 2.4 and yields the same result when reversals are symmetric. Incorporating both PART and PARTR allows us to separately observe the estimated magnitude of the earnings management and the associated reversal to evaluate whether they make economic sense.
WC_ACC\(_{i,t}\) = \(a + b\text{PART}_{i,t} + c\text{PARTP1}_{i,t} + d\text{PARTP2}_{i,t} + \sum k f_k X_{k,i,t} + e_{i,t}\) \hspace{1cm} (5)

We then conduct a test of the linear restriction that \(b - c - d = 0\) to test for earnings management. While similar in spirit to including a single reversal variable PARTR, where PARTR=\text{PARTP1}+\text{PARTP2}, this approach allows us to separately estimate the magnitude of the reversal in each of the subsequent two periods.

### 3.2 Models of Nondiscretionary Accruals

A wide variety of nondiscretionary accrual models have been employed by previous research. We examine common variants of the most popular models, and our testing framework is easily extended to other models. Two key features distinguish each model:

(i) The measure of accruals; and

(ii) The determinants of nondiscretionary accruals, \(X_k\).

We use non-cash working capital accruals (WC_ACC) as the measure of accruals in all of our models, where

\[
WC\_\text{ACC}_{i,t} = (\Delta CA_{i,t} - \Delta CL_{i,t} - \Delta Cash_{i,t} + \Delta STD_{i,t})/A_{i,t-1}
\]

and

\[
\Delta CA = \text{the change in current assets}
\]
\[
\Delta CL = \text{the change in current liabilities}
\]
\[
\Delta Cash = \text{the change in cash}
\]
\[
\Delta STD = \text{the change in short-term debt}
\]
\[
A = \text{total assets}.
\]

Early research also subtracts depreciation expense in the definition of accruals (e.g., Healy [1985]), but this adjustment is often dropped in subsequent research on the grounds that it is related to long-term capital expenditure accruals rather than working capital accruals (e.g., Allen, Larson and Sloan [2010]).

We consider five popular models that can be summarized in terms of their different choices of nondiscretionary accrual determinants as follows:
Healy
Healy [1985] does not incorporate any determinants of nondiscretionary accruals.

Jones
Jones [1991] includes the change in revenues and the level of gross property, plant and equipment (PPE hereafter) as determinants of nondiscretionary accruals.

\[ X_{1,i,t} = \Delta REV_{i,t} = (Revenue_{i,t} - Revenue_{i,t-1}) / A_{i,t-1} \]
\[ X_{2,i,t} = PPE_{i,t} = PP&E_{i,t}/A_{i,t-1} \]

Modified Jones
DSS show that the original Jones model has low power in cases where firms manipulate revenue through the misstatement of net accounts receivable. This is because the original Jones model includes the change in credit sales as a determinant of nondiscretionary accruals, resulting in the removal of discretionary accruals (Problem 2 from section 2.1). To mitigate this problem, DSS suggest that cash revenue be used in place of reported revenue.\(^{10}\)

\[ X_{1,i,t} = \Delta REV_{i,t} - (Net Accounts Receivable_{i,t} - Net Accounts Receivable_{i,t-1}) / A_{i,t-1} \]
\[ X_{2,i,t} = PPE_{i,t} = (PP&E_{i,t}/A_{i,t-1}) \]

DD
Dechow and Dichev [2002] note that if the objective of accruals is to correct temporary matching problems with firms’ underlying cash flows, then nondiscretionary accruals should be negatively correlated with contemporaneous cash flows and positively correlated with adjacent cash flows. They therefore propose including past, present and future cash flows (CF) as additional relevant variables in explaining nondiscretionary accruals.\(^{11}\)

\[ X_{1,i,t} = CF_{i,t-1} \]

\(^{10}\) DSS suggest that this adjustment only be made in years that earnings management is hypothesized. We make the adjustment in all years for two reasons. First, the change in accounts receivable has a positive sample mean, and so only adjusting earnings management years causes the change in sales to be downward biased in earnings management years and discretionary accruals to be upward biased in earnings management years, leading to excessive rejections of the null. We confirmed this fact in unreported tests. Second, when modeling reversals, an adjustment would also be required in reversal years, making the selective adjustment of earnings management-related years cumbersome.

\(^{11}\) We note that Dechow and Dichev [2002] do not specifically propose that their model be used in tests of earnings management, but subsequent research has adopted it in this context. See McNichols [2002] and Dechow, Ge and Schrand [2010] for further details.
\[ X_{2,t} = CF_{i,t} \]
\[ X_{3,t} = CF_{i,t+1} \]

where \( CF_{i,t} \) = Earnings before Extraordinary Items \( i,t \) – DAP \( i,t \).

Wysocki [2009] reasons this model will tend to classify discretionary accruals that are made with the intention of smoothing earnings as nondiscretionary. For example, a firm with deteriorating cash flows may try to manage accruals upwards to avoid reporting deteriorating earnings. This model is therefore poorly suited to tests of earnings management where the hypothesis entails earnings smoothing.

**McNichols**

Finally, McNichols [2002] recommends that researchers combine the determinants from both the Jones and the DD models described above.

We make three additional research design choices that apply to all of the models. First, we estimate the models as a single panel that pools across all available firm-years in our sample. This approach is common in the existing literature. Another common approach is to estimate each of the models by industry and year and then conduct the earnings management tests by pooling across the model residuals (e.g., Defond and Jiambalvo [1994]). In unreported tests, we confirmed that this approach yields results that are qualitatively similar to those reported in the paper. A less common approach that was adopted by early research is to estimate each model at the firm level and then conduct statistical inference by aggregating t-statistics from the firm-specific regressions (e.g., Jones [1991]). This approach has been largely dropped by later research because it results in a considerable loss of power. The loss in power arises because firms with insufficient observations to conduct a firm-specific regression have to be dropped and because a separate set of model parameters has to be estimated for each firm. In unreported tests, we confirmed that this approach results in the loss of a substantial number of observations and a significant decline in test power.

The second choice we make in our research design is to conduct all statistical tests using the heteroscedasticity consistent covariance matrix proposed in MacKinnon and White [1985]
and commonly referred to as HC3. This approach to incorporating heteroscedasticity has been shown to be the best specified across a broad range of sample sizes (e.g., Long and Erwin [2000]). Note that because tests using HC3 appeal to asymptotic theory, all linear restrictions are tested using a chi-square ($\chi^2$) statistic. We note that most previous earnings management research employs standard OLS regression analysis, implicitly assuming that the residuals in equation (3) are independently and identically distributed. We conducted a series of diagnostic tests to identify significant violations of this assumption. While we find little evidence of systematic time series or cross-sectional (e.g., industry) dependence in model residuals, we do find significant evidence of heteroscedasticity when grouping firm-years by characteristics such as size, ROA and presence of an SEC enforcement action. Hence, we recommend the use of heteroscedasticity consistent covariance matrices in tests for earnings management.

The third choice we make in our research design is to conduct tests using performance matched discretionary accruals following the procedure described in Kothari, Leone and Wasley [2005]. The matched pair is the firm-year in the same two-digit SIC code and fiscal year with the closest ROA. We follow Kothari et al. in conducting separate tests for matching on ROA, and ROA respectively. Performance matched discretionary accruals are computed by taking the residuals from each of the models of nondiscretionary accruals (estimated excluding the earnings management partitioning variables) and subtracting the corresponding residual for the matched pair in the PART=1 year. We also follow Kothari et al. in conducting inference using a standard $t$-test against a null of zero on the resulting differenced residual. For comparative purposes, when we report these test statistics, we square the $t$-statistic to arrive at the corresponding $F$ statistic, which approximates a $\chi^2$ statistic for large sample sizes and hence is comparable to the $\chi^2$ statistics from our reversal models. We emphasize that we adopt the Kothari et al. approach for comparative purposes and because of its popularity in the existing literature. Our main purpose in doing so is to demonstrate the effectiveness of modeling accrual reversals as an alternative to Kothari et al.’s performance matching procedure in addressing misspecification due to correlated omitted variables.
4. Experimental Design

4.1 Data

Our sample consists of available firm-years from the Compustat annual files for which we can calculate WC_ACC. We therefore require positive non-missing values of the following variables: (Compustat mnemonics in brackets): receivables (rect); current assets (act); current liabilities (lct); cash and equivalent (che); short-term debt (dlc); total assets (at); sales (sale); PP&E (ppegt). We also require non-missing values of earnings before extraordinary items (ib) so that we can derive cash flows, CF, for use in the Dechow and Dichev model and cash flow performance matching tests. Annual Compustat data are pulled using the DATAFMT=STD flag to ensure that we are using the original ‘as reported’ and unrestated data. We exclude financial firms, since working capital is less meaningful for these firms, and we winsorize all financial variables at the one percent tails. Following Kothari, Leone and Wasley’s [2005], we define ROA as earnings before extraordinary items divided by lagged total assets. Our final sample consists of 209,530 firm-year observations between 1950 and 2009.

4.2 Test Procedure

We follow a similar procedure to DSS to examine the power and specification of each of the models. We first examine each model in its traditional form, and we then examine the impact of incorporating earnings management reversals. The models are evaluated in four different contexts. First, we examine model specification using randomly selected earnings management years. Second, we artificially seed earnings management and its associated reversal to evaluate the gains in test power resulting from incorporating reversals. Third, we examine the power of the models using a sample of firms identified by the SEC as having manipulated earnings. Finally, we examine model specification in situations where the earnings management years are correlated with various economic characteristics.

4.2.1 Tests where the earnings management year (i.e., PART=1) is randomly selected

To evaluate test specification in random samples of firm-years we perform the following steps for each combination of models and tests:
1. From among the 209,530 firm-years, we randomly select 100 firm-year observations.\textsuperscript{12} The 100 firm-years are designated as earnings management years (i.e., PART=1). The remaining firm-years are designated as non-earnings management years (i.e., PART=0).

2. We then determine whether data are available in the two years immediately following each earnings management year. If they are, we set PARTP1=1 and PARTP2=1 for the first and second year respectively and equal to zero otherwise.

3. We conduct a pooled regression for each model as described in the previous section using all 209,530 firm-years.

4. Steps 1 and 2 are repeated 1,000 times.

5. We record the frequency with which the null hypothesis of no earnings management is rejected at the five percent level using one-tailed tests (for the $\chi^2$ tests we use a 10% level and condition on the direction in which the linear constraint is rejected, effectively conducting a one-tailed test at the 5% level).\textsuperscript{13}

\subsection*{4.2.2 Simulation Tests with Induced Earnings Management}

The purpose of these tests is to examine the power of the models to detect earnings management in settings where we know the magnitude and timing of the earnings management and associated reversal. Our tests differ from those in previous research, such as DSS, in that we also simulate the reversal of the earnings management. Our first set of simulations examines how changing the proportion of the reversal that is correctly modeled by the researcher impacts test power. These simulations are conducted through the following 6 steps:

1. From among the 209,530 firm-years, we randomly select 100 firm-year observations. The 100 firm-years are designated as earnings management years (i.e., PART=1). The remaining firm-years are designated as non-earnings management years (i.e., PART=0).

2. For the 100 earnings management years we artificially induce earnings management by adding “discretionary accruals” equal to 1% of the beginning total assets.

3. We then determine whether data are available in the year immediately following each earnings management year. If it is, we set PARTP1 equal to one for that year. We

\textsuperscript{12} The SAS code that we use is \texttt{proc surveysselect data=data1 method=seq n=100 out=data2 reps=1000 seed=1347865}.

\textsuperscript{13} For example, if a $\chi^2$ test rejects the null hypothesis that $b-c=0$ at the 10% level and $b-c>0$, we register a rejection of the null hypothesis that $b-c\leq0$ at the 5% level.
consider 11 scenarios in which the induced earnings management in step 2 is reversed in increments of 10%, from 0% (i.e., no reversal) to 100% (i.e., complete reversal) in this subsequent year.

4. We estimate a pooled regression for each model using all 209,530 firm-years and conduct tests for earnings management.

5. Steps 1 through 4 are repeated 1,000 times.

6. We repeat steps 1 through 5 after substituting earnings management of 2% of beginning total assets at step 2.

Our second set of simulations examines test power as a function of sample size. In these simulations, we assume that the researcher correctly models the reversal of earnings management and examine how incorporating reversals impacts test power relative to traditional tests of earnings management that ignore reversals. These simulations are conducted through the following 5 steps:

1. From among the 209,530 firm-years, we randomly select 100 firm-year observations. The 100 firm-years are designated as earnings management years (i.e., PART=1). The remaining firm-years are designated as non-earnings management years (i.e., PART=0).

2. For the 100 earnings management years we artificially induce earnings management by adding “discretionary accruals” equal to 2% of the beginning total assets.

3. We then determine whether data are available in the year immediately following each earnings management year. If it is, we set PARTP1 equal to one for that year and we add a “discretionary accrual reversal” equal in magnitude but opposite in sign to the discretionary accruals in step 2 (i.e., 100% reversal).

4. We estimate a pooled regression for each model using all 209,530 firm-years and conduct tests for earnings management.

5. Steps 1 through 4 are repeated 1,000 times.

6. We repeat steps 1 through 5 varying the number of earnings management firms selected in step 1 from 100 to 1,000 in increments of 100.

4.2.3 SEC Accounting and Auditing Enforcement Release (AAER) Sample

We use the AAER sample to examine the power of the different tests and models to detect earnings management in a sample of firm-years where we have strong priors that earnings have
been managed. The advantage of these tests is that we don’t have to make assumptions about either the magnitude or timing of the earnings management and reversal. Instead, we employ a sample of firm-years examined by Dechow, Ge, Larson and Sloan [2011] in which the SEC alleges that upward earnings management has taken place. Dechow, Ge, Larson and Sloan [2011] identify the specific years in which the alleged earnings management takes place by reading the associated SEC accounting and auditing enforcement releases. We expect these cases of earnings management to be particularly egregious. Moreover, the fact that they were identified and targeted by the SEC makes it probable that any earnings management subsequently reversed. So this sample provides an ideal setting to look for both evidence of earnings management and its associated reversal. If we are unable to document evidence in this sample, then it seems unlikely that our tests for earnings management could be effective in other less extreme setting.

There are 230 firms and 406 firm-years for which the SEC makes allegations of upwardly managed earnings. We evaluate the ability of the different models to detect earnings management through the following steps:

1. We set PART=1 in the 406 firm-years in which upward earnings management is alleged to have taken place and PART=0 otherwise.
2. We set PARTP1=1 in the first year following the final earnings management year and PARTP1=0 otherwise.
3. We set PARTP2 =1 for the second year following the last earnings management year and PARTP2=0 otherwise.
4. We conduct a pooled regression for each model using all 161,119 firm-years during the 1973-2003 period spanned by the AAERs.
5. We repeat the above steps 1 through 4 for a subset of 122 of the 406 AAER firm years in which the SEC specifically alleges that a component of working capital accruals was manipulated.

### 4.2.4 Tests where the earnings management year (i.e., PART) is randomly selected from portfolios with extreme economic characteristics

To determine the specification of the models for samples where the earnings management partitioning variable is correlated with common economic characteristics, we perform the following steps for each model:
1. We rank the 209,530 firm-years into ten portfolios based on the corresponding economic characteristic, where decile 10 consists of firms with the highest values of the characteristic. We then randomly select 100 firm-years from decile 10. The 100 firm-years are designated as earnings management years (i.e., PART=1) with the subsequent two firm-years designated as reversal years (i.e., PARTP1=1 and PARTP2=1).

2. We conduct a pooled regression for each model described in the previous section using all 209,530 firm-years.

3. Steps 1 and 2 are repeated 1,000 times.

4. We record the frequency with which the null hypothesis of no earnings management is rejected at the five percent level for each earnings management test.

5. We repeat steps 1 through 4, but select 100 firm years from decile 1 (lowest values of the characteristic).

We perform these tests for a variety of economic characteristics that are commonly encountered in earnings management studies. These characteristics include ROA, sales growth, size (market capitalization), operating cash flows and the consensus analyst forecast of long-term earnings growth.14

5. Results

5.1 Descriptive Statistics

Table 1 reports descriptive statistics for working capital accruals. This table illustrates the intuition behind several of our subsequent results. Panel A indicates that working capital accruals have a positive mean, suggesting that the sample firms have grown in scale over the sample period. Panel B reports some pertinent correlations. First, the correlation between working capital accruals and earnings is 0.18, indicating that working capital accruals are an important driver of contemporaneous earnings. Second, the serial correlation in working capital accruals is weakly positive. This result tells us that working capital accruals tend to neither immediately reverse nor strongly persist ‘on average’. It is consistent with the results in Allen, Larson and Sloan [2010], which shows that working capital accruals contain both strongly

14 Consensus analyst forecasts of long-term earnings growth are obtained from I/B/E/S and are only available for a subset of 53,025 firm-years.
reversing and strongly persistent components that cancel each other ‘on average’.

Panel C reports average accruals for years t, t+1 and t+2 by deciles formed on earnings performance (i.e., ROA) in year t. There is clear evidence of a positive correlation between working capital accruals and earnings performance in all three years. Thus, there is a strong positive correlation between earnings and accruals that persists over the next 2 years. These results are consistent with the existence of nondiscretionary accruals that capture persistent economic performance. The fact that the high (low) accruals in high (low) earnings deciles tend to persist helps rule out the possibility that these accruals are due to reversing earnings management. These results also highlight the intuition for why incorporating accrual reversals into tests of earnings management helps to mitigate misspecification due to correlated omitted determinants of nondiscretionary accruals. If we only infer that earnings management is present when we see clear evidence of a corresponding accrual reversal, we would not infer that the accruals in the extreme earnings performance deciles in panel C represent earnings management, because they do not reverse.

Panel D reports average accruals for years t, t+1 and t+2 for firm-years in the AAER sample that are alleged to have managed earnings upwards in year t. Accruals are large and positive (0.077), in year t and similar in magnitude to the top earnings decile in panel C (0.081). But unlike the top decile panel C accruals that stay high in periods t+1 and t+2 (0.057 and 0.036, respectively), the accruals for the AAER sample exhibit a sharp reversal and are significantly negative in periods t+1 and t+2 (-0.047 and -0.035, respectively). Note also that the sum of the negative accruals in periods t+1 and t+2 are opposite in sign but approximately equal in magnitude to the positive accruals in period t. Thus, in a sample where we expect that earnings management is present, we see clear evidence of the predicted subsequent accrual reversal. This pattern illustrates how incorporating accrual reversals should increase the power of tests for earnings management.

A final feature to note from table 1 is that the standard deviation of accruals varies widely across sample partitions. The overall sample standard deviation from panel A is 0.124. But the standard deviation ranges from a low of 0.077 for earnings decile 5 in period t to highs of 0.213 for the lowest earnings decile in period t and 0.192 for the AAER sample in period t. The standard deviation of accruals is clearly associated with firm-year characteristics such as earnings performance and the presence of SEC enforcement actions. These significant
violations of the assumption of independently and identically distributed errors that underlie OLS regression analysis are what prompt us to control for heteroscedasticity in the estimation of our earnings management models (see section 3.2 for details).

5.2 Specification of Tests for Earnings Management in Random Samples

Table 2 Panel A reports the mean coefficients and t-statistics from the 1,000 simulations using randomly assigned earnings management years for each of the competing earnings management models and test statistics. As expected for the random assignments, the parameter estimates on the earnings management partitioning variables and associated reversal variables are all close to zero for all models. Similarly, the mean performance-matched discretionary accruals are also close to zero. The various explanatory variables in the models also take on their predicted values, consistent with previous research. The coefficients on sales growth are significantly positive, the coefficients on contemporaneous cash flows are significantly negative and the coefficients and lead/lag cash flows are significantly positive.

Panel B reports the rejection frequencies for each of the discretionary accrual models using each of the competing tests for earnings management. There are five models (Healy, Jones, Modified Jones, DD and McNichols models) and five different tests ($b=0$, $b-c=0$, $b-c-d=0$, performance matching on ROA$t$ and performance matching on ROA$t-1$). We also report one-tailed tests for both positive and negative earnings management, so panel B reports 40 sets of rejection frequencies in total. All tests are conducted using a 5% test level, and so the observed rejection frequencies should be 5% for well-specified tests.

The results indicate that all models are relatively well specified, in that their rejection frequencies are close to the specified test level. The only notable exceptions are for tests of the form $b-c-d=0$ using the Healy and DD models, where the rejection frequencies look somewhat high for positive earnings management and somewhat low for negative earnings management. We note that the average level of accruals has declined slightly over time, providing a potential explanation for these rejection rates. The fact that models controlling for sales growth do not exhibit this problem suggests that the decline in accruals reflects a corresponding decline in nondiscretionary accruals that are correlated with sales growth.

Finally, panel C of table 2 reports serial correlations for estimated discretionary accruals (i.e., the residuals) and nondiscretionary accruals (i.e., the fitted values) from each of the
nondiscretionary accrual models. Recall from table 1 that working capital accruals display weak positive serial correlation in the pooled sample. Allen, Larson and Sloan [2010] predict that nondiscretionary accruals should display positive serial correlations, since they reflect persistent economic performance, while discretionary accruals should display negative serial correlations, since they reflect reversing earnings management. The first two columns of panel C report the mean firm-specific serial correlations for discretionary and nondiscretionary accruals using each of the models. The results for the Healy, Jones and Modified Jones model are all consistent with the predictions of Allen et al. The discretionary accruals are negatively serially correlated while the nondiscretionary accruals are positively serially correlated. The serial correlations for the DD and McNichols models, in contrast, display negative serial correlations for nondiscretionary accruals. These results are consistent with Wysocki’s [2009] conclusion that by controlling for contemporaneous cash flows, these models classify some discretionary accruals as nondiscretionary. In order to gain further insights into these results, we focus on the serial correlations for firms that are alleged by the SEC to have managed earnings. The middle two columns of panel C report the mean serial correlations for firms that belong to the AAER sample. The serial correlations for these firms generally look similar to those for the full sample. But remember that these firm-specific results use all available observations for each AAER firm in our 59 year sample period, whereas earnings management is typically only alleged in one or two years. We therefore report pooled serial correlations using just the firm-years during which earnings management is alleged by the SEC (i.e., PART=1) and the immediate following year (i.e., PARTP1=1), since this is where we would expect strong evidence of reversals. Consistent with this prediction, we see much stronger negative serial correlation in discretionary accruals for all 5 of the models.

### 5.3 Power of Tests for Earnings Management using Simulations with Seeded Earnings Management

Our objectives in inducing earnings management into random samples are twofold. First, as in DSS, these tests illustrate the effectiveness of particular models in detecting earnings management of known magnitudes. Second, we can use them to illustrate the gains in power from modeling accrual reversals, including settings where the researcher has relatively weaker priors concerning the timing of earnings management reversals. For example, assume that the
researcher hypothesizes that earnings management will occur in year t and reverse in year t+1, and while the hypothesized earnings management all occurs in period t, only 50% of it actually reverses in year t+1. Is the power of the test still improved by modeling reversals in year t+1?

Figure 1 provides the results for tests where we successively induce earnings management in the magnitude of 1%, and then 2% of total assets. Recall that these tests are based on 1,000 simulations in which 100 observations out of 209,530 have induced earnings management. For each case, we report the frequency with which the null hypothesis of no earnings management is rejected in favor of the alternative of positive earnings management using a 5% test level. All tests use the Healy model and incorporate reversals in the next period (i.e., $b-c=0$). For comparative purposes, we also report rejection rates using (i) the Healy model of discretionary accruals using the standard $t$-test for earnings management (i.e., $b=0$); and (ii) performance matched discretionary accruals using the Healy model and matching on ROA. Note that the rejection rates for (i) and (ii) are not expected to change with the magnitude of the reversal, since these tests do not incorporate reversals. We use the horizontal axis to vary the proportion of the induced earnings management in period t that reverses in period t+1. On the far left, we simulate no reversal (i.e., 0%) and on the far right we simulate a complete reversal (i.e., 100%).

Figure 1A provides results where the induced earnings management is 1% of assets. The benchmark rejection frequencies are 20.7% for the standard test of $b=0$ and 15.6% for performance matching. Note that consistent with the results in Kothari, Leone and Wasley [2005], performance matching reduces test power by approximately 40%. These rejection frequencies indicate that standard tests have low power for earnings management under these conditions. If we model a reversal in period t+1 when there isn’t one, the rejection frequency drops to 14.2%. Thus, incorporating reversals that do not occur is detrimental to test power. On the other hand, if we model a reversal in period t+1 and a 100% reversal actually occurs, then the rejection frequency increases to 27.8%. Relative to the standard $t$-test, the breakeven point from which modeling a reversal starts to improve test power occurs when just over 50 percent of the induced earnings management reverses in the following period. Relative to the performance-matched test, the breakeven point is just over 10%.

---

15 The key takeaways are similar using the other models and tests, and so we omit them for brevity.
Figure 1B reports corresponding results where the induced earnings management is increased from 1% of assets to 2% of assets. The rejection frequency for the standard $t$-test increases to 46.7% and the rejection frequency from modeling 100% reversals increases to 67.0%. The breakeven point when modeling the reversal improves test power relative to the standard $t$-test is when approximately 40 percent of the accrual reverses in the next period. The rejection frequency for performance matched discretionary accruals is 35.0% and the modeling accrual reversals results in greater test power when there is at least a 20% reversal. These tests illustrate that modeling reversals has the potential to increase test power by almost fifty percent relative to the standard tests (67.0%/46.7%) and almost one hundred percent compared to performance matching (67.0%/35.0%). Moreover, note that we have assumed the researcher is one hundred percent sure about the earnings management period and have only considered uncertainty about the reversal period. It is also possible that the researcher faces uncertainty concerning the earnings management period. For example, consider earnings management around equity offerings. Management may manage earnings for several quarters prior to the offering, and then take a ‘big bath’ write-down in the fourth quarter following the offering (when manager and underwriter lock-ups typically expire). If the researcher has relatively weak priors about the timing of the originating earnings management and strong priors about its reversal, then the gains to modeling reversals can be even greater.\footnote{Other examples where the researcher may have stronger priors concerning the timing of the reversal relative to the origination of the earnings management include ‘big bath’ write-downs by an incoming CEO to reverse earnings management by a predecessor and the reversal of ‘cookie jar’ reserves, whereby earnings are managed slightly downward over an extended time period in order to provide flexibility to manage earnings upward in future challenging periods.}

The results in figure 1 illustrate that all tests of earnings management have relatively low power in settings where 100 observations experience earnings management of 1% to 2% of total assets. A common rule of thumb in hypothesis testing is that the ex ante power of a test should be at least 80% in order to have the precision to provide reliable inferences (e.g., Aberson [2010], p. 15). None of the tests in figure 1 meet this rule of thumb. Figure 2 summarizes power analysis simulations that allow us to determine the number of observations with PART=1 necessary to generate an 80% rejection rate. For these simulations, we induce earnings management equal to 2% of assets and incorporate a 100% reversal in the next period. The test incorporating reversals reaches 80% power between 100 and 200 observations. In contrast, it
takes over 300 observations for the standard $t$-test to reach 80% power. Performance matching further increases the required sample size to over 500 observations. These results highlight how incorporating reversals can provide more reliable tests of earnings management when large sample sizes are not available.

To summarize, the results in this section indicate that tests incorporating earnings management reversals do not unambiguously increase test power. If the researcher models a reversal in a period when no reversal actually occurs, test power is reduced. But so long as at least half of the earnings management reverses in the period in which it is modeled, incorporating reversals increases test power. In cases where the researcher’s priors about the timing of the reversal of the earnings management are as strong as those about its origination, test power increases by almost fifty percent relative to the standard $t$-test and almost doubles compared to performance matching. If the researcher has stronger priors about the timing of the reversal than the origination of the earnings management, increases in power can be even more substantial. Given the low power of tests for earnings management in typical research settings, such gains in power can significantly increase test reliability.

5.4 Power of Tests for Earnings Management Using the AAER Sample

Our next set of tests examines the importance of considering accrual reversals in the AAER sample, where we have strong priors that upward earnings management has taken place. Figure 3 plots mean accruals for our AAER sample centered on the alleged year of the earnings management, comparing it to a sample where the earnings management year is randomly selected. The figure indicates that accruals are unusually high for AAER firms in year 0. Interestingly, they are also high in the years immediately prior to year 0. This could indicate that these firms were engaging in less egregious earnings management in the preceding years. Perhaps the most striking feature of figure 3 is the strong accrual reversals that take place in years $t+1$ and $t+2$. It is perhaps these reversals that were catalysts for the SEC investigations. It seems intuitive that tests incorporating these reversals should provide more power in tests for earnings management.

Table 3 presents the results of formal tests for earnings management in the AAER sample. Panel A reports the coefficient estimates from the five models. Note that the underlying AAER
sample is significantly enlarged and updated relative to the original sample used in DSS (we have 406 firm-years versus 56 in their sample). The point estimates for \( b \), the coefficient in the earnings management year range from 2.8% for the McNichols model to 5.7% for the Healy model. Note that relative to the Healy model, all of the other models that contain additional controls for nondiscretionary accruals have lower point estimates for \( b \). This finding is consistent with the contention in DSS that controls for nondiscretionary accruals unintentionally eliminate discretionary accruals (Problem 1 from section 2.1). The associated \( t \)-statistics are all significant at conventional levels and range from 3.714 for the McNichols model to 6.046 for the DD model.

The point estimates for \( c \), the coefficient in the first potential ‘reversal’ year, range from -6.4% for the Jones model to -6.8% for the Modified Jones model. All of these estimates are consistent with significant accrual reversals in the first post earnings management year. Moreover, the reversals are all at least as large as the associated earnings management. The point estimates for \( d \), the coefficient in the second year after the earnings management year, range from -4.3% for the McNichols model to -5.5% for the Healy model. These estimates are consistent with continued reversals in the second post earnings management year. Note that the summed magnitude of the reversals exceeds the magnitude of the original earnings management. For example, the estimate of earnings management in the Healy model is 5.7%, while the associated reversals over the next two years sum to -6.6% + -5.5% = -12.1%. There are several possible explanations for this result. First, visual inspection of figure 3 indicates that discretionary accruals in the AAER sample are unusually high in the three years leading up the earnings management year. While year 0 is the year in which the SEC alleges that egregious earnings management has occurred, it is quite likely that these firms have been engaging in less egregious earnings management in the preceding years (e.g., Schrand and Zechman [2011]). So the reversals in years 1 and 2 may relate to the sum of the positive accruals in several prior years. Second, the SEC may be more likely to identify and litigate earnings manipulations characterized by large accrual reversals. Third, firms may be forced to liquidate operations in the immediate post AAER period due to lack of financing stemming from investor mistrust and a higher cost of
capital (e.g., McNichols and Stubben, [2008]).

Panel B of table 3 reports the associated point estimates and test statistics on earnings management using performance-matched discretionary accruals. The point estimates are very similar to those reported in panel A. The associated $t$-statistics, however, are generally lower using performance matching. Recall that performance matching results in less powerful tests because of the increased standard errors resulting from the use of a matched pair.

Finally, panel C of table 3 compares the power of the 5 different tests for earnings management using each of the 5 models. We report $\chi^2$ statistics for tests of linear constraints in the reversal models and $F$-statistics (which are just the square of the associated $t$-statistics) for the performance-matched tests. The $F$ distribution approximates the $\chi^2$ distribution for large sample sizes, so relative power of the tests can be directly evaluated from the magnitude of the associated test statistics. The tests incorporating accrual reversals have the largest test statistics for all 5 models. The performance matched tests, meanwhile, generally have the lowest test statistics. For example, using the Healy model, the test statistics range from 89.524 for tests of the linear constraint that $b-c-d=0$ to 22.184 for tests based on performance matching on ROA in period $t-1$. Modeling accrual reversals dramatically improves the power of tests for earnings management in the AAER sample.

A potential shortcoming of the tests in table 3 is that only a subset of the AAERs examined involves explicit allegations of upward earnings management via working capital accruals. Many of the AAERs involve other types of earnings management such as management of noncurrent assets and liabilities. Of course, the fact that there is no explicit allegation of working capital accrual management does not rule out the possibility that such accruals are managed. Managers will likely use within-GAAP earnings management in multiple accrual categories before resorting to the type of egregious GAAP violations targeted by the SEC. Nevertheless, we would expect to see greater evidence of working capital accrual management in cases where the AAER has explicitly targeted working capital accruals.

---

17 In unreported tests, we confirmed that the accrual reversals for the AAER sample in years $t+1$ and $t+2$ are attributable to the reversal of past asset accruals rather than the origination of new liability accruals. We also confirmed that cash flows for the AAER sample are unusually low in years $t+1$ and $t+2$. These results suggest that the year $t+1$ accrual reversals are due to the reversal of past upward management in asset accruals rather than the origination of new liability accruals or the liquidation of assets.
Confirming that this is the case would also help to corroborate that our tests are picking up earnings management.

Table 4 replicates the tests in table 3 for a subsample of 122 of the original 406 AAER firm-years in which the SEC explicitly alleges that a working capital account (e.g., accounts receivable, inventory or accounts payable) was managed. In this sample, we set PART=1 for the 122 firm years and PART=0 for all other firm years (including the 384 other AAER firm-years). The results indicate that the point estimates of the amount of earnings management and associated reversals are all greater than in table 3. For example, using the Healy model, the estimated coefficient on \( b, c \) and \( d \) are 8.4%, -8.6% and -8.9% in table 4 versus 5.7%, -6.6% and -5.5% in table 3. Note, however, that the statistical significance is lower in all tests, due to the vastly reduced sample size.

To summarize, the results in this section indicate that tests incorporating reversals of earnings management display substantially higher power in the AAER sample. The biggest increases come from modeling the reversal in the first subsequent year, with additional improvements from modeling reversals in the second year. The results indicate that the choice of nondiscretionary accrual model has relatively little impact on test power. It appears that the controls for nondiscretionary accruals also extract some discretionary accruals, such that there is little net effect on test power. Remember, however, that these models are also designed to mitigate misspecification stemming from correlated nondiscretionary accruals. Thus, the choice between competing models should also consider model specification in the absence of earnings management. Our next set of results examines misspecification in more detail.

5.5 Specification of Tests for Earnings Management in Samples with Extreme Economic Characteristics

As described earlier, a common problem that arises in tests of earnings management is the omission of determinants of nondiscretionary accruals. In particular, the hypothesized determinants of earnings management are often correlated with economic characteristics that

\(^{18}\) In unreported tests, we estimate a single panel regression incorporating separate earnings management partitioning variables for the 122 working capital related AAERs and the 384 other AAERs. Using F-tests for equality of coefficients across the two samples, we find that the coefficients on the partitioning variables are generally significantly higher for the working capital AAERs.
may influence nondiscretionary accruals. Earnings management tests that do not control for these correlated determinants of nondiscretionary accruals will be misspecified (Problem 2 in section 2.1). DSS illustrate this problem by showing that samples of firms with extreme earnings and cash flow performance have rejection rates that deviate substantially from the specified test levels. Commonly proposed solutions to this problem include adding the omitted variables to the nondiscretionary accruals models (e.g., Jones, Modified Jones, DD and McNichols models) and the performance matching procedures suggested by Kothari, Leone and Wasley [2005]. These procedures have several limitations. Most importantly, they require the researcher to know the relevant omitted variables. Yet existing research only provides a rudimentary understanding of the determinants of nondiscretionary accruals. Also, controlling for these determinants can reduce test power both through the incorrect modeling of discretionary accruals as nondiscretionary, and by increasing standard errors through the inclusion of irrelevant variables.

Incorporating accrual reversals in tests of earnings management provides an alternative approach for mitigating model misspecification. Most firms are going concerns, and so their economic characteristics tend to persist. As such, any associated nondiscretionary accruals should also persist. Earnings management, on the other hand, must ultimately reverse. If the researcher has reasonable priors concerning the timing of reversals, tests incorporating reversals should mitigate correlated omitted variable bias.

Figure 4 illustrates the intuition behind the above argument. DSS document that accruals are correlated with contemporaneous earnings performance. Figure 4 plots mean working capital accruals for firms in the highest and lowest deciles of earnings performance (i.e., ROA) in event year 0. The firms with high earnings performance have positive accruals in year 0 and the firms with low earnings performance have negative accruals in year 0. But, we also see that these patterns in accruals tend to persist in the surrounding years. This pattern can be contrasted to the accruals for the AAER earnings management sample in figure 3. Accruals are also high in year 0 for this sample, but they exhibit a strong reversal in the next two periods, which is the trademark of earnings management. As long as the omitted determinants of nondiscretionary accruals do not reverse, incorporating accrual reversals provides a general solution for mitigating associated misspecification in tests of earnings management. The two
key advantages of this approach are (i) the underlying determinants of nondiscretionary accruals do not need to be identified; and (ii) the power of the tests for earnings management increases (recall that the performance matching approach, in contrast, reduces test power).

Table 5 formally evaluates the ability of tests incorporating accrual reversals to mitigate misspecification using simulations on samples with extreme earnings performance. Note that we have selected the sample based on the same variable that Kothari, Leone and Wasley [2005] use for performance matching (i.e., earnings performance). Thus, we have simulated the exact situation where their performance matching technique should be most effective by construction. The more interesting question we address here is the effectiveness of modeling accrual reversals in this situation.

The left hand side of table 5 reports rejection frequencies using a 5% test level for the alternative hypothesis of positive earnings management. Consistent with previous research, we see that the standard test for \( b=0 \) has excessively high rejection frequencies in the highest earnings performance decile for all five models. For example, the Healy model rejects 98.8% of the time using a 5% test level. We also see that, as expected, performance matching substantially mitigates this problem. For example, the rejection frequencies for the Healy model drop to 7.8% matching on ROA\(_{t-1}\) and 6.9% matching on ROA\(_t\). More importantly, we also see that modeling reversals substantially alleviates the excessive rejection frequencies. For example, using the Healy model, the rejection frequencies drop to 27.8% when testing \( b-c=0 \) and 7.2% when testing \( b-c-d=0 \).

The right hand side of table 5 reports rejection frequencies for the alternative hypothesis of negative earnings management. Consistent with previous research, the standard test for \( b=0 \) has excessively high rejection frequencies in the lowest earnings performance sample for all four models. For example, the Healy model rejects 86.9% of the time using a 5% test level. We also see that, as expected, performance matching substantially mitigates this problem. More importantly, we again see that modeling reversals substantially alleviates the excessive rejection frequencies. For example, using the Healy model, the rejection frequencies drop to 30.3% when testing \( b-c=0 \) and 12.1% when testing \( b-c-d=0 \). Note that in this particular context, performance matching on ROA is more effective at mitigating misspecification. But remember
that we have simulated the exact situation where performance matching on ROA should work. We now move to tests where we select samples based on other economic characteristics to evaluate the extent to which these results generalize.

We select four alternative economic characteristics for our additional specification tests. We select three characteristics that are previously shown to result in misspecified tests in Kothari, Leone and Wasley [2005]. These three characteristics are sales growth, size (i.e., market capitalization at the end of the period) and cash flows (i.e., earnings before extraordinary items less working capital accruals, deflated by beginning of period total assets). The fourth characteristic is the consensus analyst forecast of long-term earnings growth. McNichols [2000] identifies this characteristic as an important correlated omitted variable in tests for earnings management.

Table 6 reports the simulated rejection frequencies using a 5% test level for samples of 100 firm-years randomly selected from the extreme deciles of each of these characteristics. It is useful to interpret these rejection frequencies in conjunction with the graphs in figure 5, which show event-time plots of mean working capital accruals for the extreme decile portfolios of each of these characteristics. Panel A (B) of table 6 reports the frequency of significantly positive (negative) rejections of the null hypothesis of no earnings management. The first column of results is for sales growth. We can see from figure 5-1 that high sales growth firms tend to have positive accruals in year 0. Thus, we should expect to see excessive rejection frequencies for tests of positive earnings management in samples of firms with high sales growth. The potential reduction in misspecification from modeling a reversal depends on the persistence of the associated accruals. Figure 5-1 indicates that high accruals in year 0 remain high (i.e., persist) over the next two years. Our reversal test that subtracts the level of accruals in year 1 from the level of accruals in year 0 should therefore reduce the likelihood that the Healy model will reject in favor of positive discretionary accruals. This intuition is confirmed for the Healy model in panel A of table 6, where the standard test for $b=0$ is rejected 88.8% of the time in the highest sales growth sample. The rejection frequencies drop to 50.4% when testing $b-c=0$ and 37.4% when testing $b-c-d=0$. We also see that performance matching does little to mitigate this misspecification, with performance matching on ROA, only lowering rejection frequencies from 88.8% to 78.8%. Incorporating reversals, however, significantly reduces misspecification. Thus,
while none of the tests substantially eliminates misspecification, incorporating reversals helps more than performance matching.

The next set of results in panel A of table 6 is for the Jones model. Not surprisingly, the reported rejection frequencies are much closer to the specified test level of 5% using the Jones model. This is because the Jones model explicitly incorporates sales growth as a determinant of nondiscretionary accruals. Results are somewhat similar for the Modified Jones and McNichols models. Finally, the rejection frequencies for tests using the DD model are similar to those for the Healy model, with the tests incorporating reversals being the least misspecified.

The next column of results in table 6 is for size. We can see from figure 5-2 that large firms tend to have lower than average accruals in year 0 and that these accruals are strongly persistent. Thus, we should expect to see rejection frequencies that are higher than the specified test levels for tests of negative earnings management in samples of large firms (reported in panel B of table 6). Moreover, because there is strong persistence in size, modeling accrual reversals in tests of earnings management should mitigate the high rejection frequencies. This intuition is confirmed for the Healy model in panel B of table 6, where the standard test for \( b=0 \) is rejected 52.9% of the time in the highest size sample. Incorporating a one period reversal, significantly reduces misspecification, with the rejection frequency falling to 4.0% for \( b-c=0 \). However, incorporating a two period reversal results in rejection frequencies falling to an excessively low rate of 0.1% for \( b-c-d=0 \). This perverse result arises because size is a strongly persistent characteristic. Thus modeling reversals over the next two periods creates the impression of positive earnings management (i.e., the sum of \(-c\) and \(-d\) is almost twice the magnitude of \( b \)). Performance matching, in contrast, does little to mitigate size-related misspecification, with performance matching on ROA, only reducing the rejection frequency to 34.3%. The rejection frequencies for the other models display a similar pattern, with the tests incorporating a one period reversal consistently proving to be the least misspecified.

The next column of results in table 6 is for cash flows. We can see from figure 5-3 that in contrast to earnings and sales growth, where both have a positive correlation with accruals, cash flows exhibit a negative correlation with accruals. Therefore, we should expect to see rejection frequencies that are higher than the specified test levels for tests of positive earnings
management in samples with low cash flows. Because there is strong mean reversion in accruals associated with extreme cash flows (accruals for low cash flow firms are 0.07 in year 0 and revert to -0.01 in year 1), modeling accrual reversals in tests of earnings management has limited effectiveness in mitigating misspecification. This intuition is confirmed for the Healy model in panel A of table 6, where the standard test for \( b=0 \) is rejected 75.1% of the time in the lowest cash flow sample. Incorporating earnings management reversals has little impact on the rejection rates. Consistent with the results in Kothari, Leone and Wasley [2005], performance matching on ROA actually makes the problem worse, increasing the rejection frequencies from 75.1% to 96.0%. These results are broadly consistent across the other accrual models. The McNichols model is the best specified of all models, because it incorporates a comprehensive set of explanatory variables. The rejection frequency for \( b=0 \) is 20.0% and this declines to 15.9% for the test \( b-c=0 \). Performance matching results in more misspecified tests, increasing rejection rates from 20.0% to 78.8% when matching on ROA. Thus, while incorporating accrual reversals results in modest improvements, performance matching on ROA exaggerates misspecification.

The final column of results in table 6 is for long-term earnings growth. We can see from figure 5-4 that high growth firms tend to have higher than average accruals in year 0, and that these accruals are weakly persistent over the next three periods. Thus, we should expect to see excessive rejection frequencies for tests of positive earnings management in samples of high growth firms that are mitigated in tests incorporating accrual reversals. This intuition is confirmed for the Healy model in panel A of table 6, where the standard test for \( b=0 \) is rejected 56.6% of the time in the high growth sample. Incorporating reversals reduces the rejection frequencies to 24.5% for \( b-c=0 \) and 21.1% for \( b-c-d=0 \). Performance matching is less effective at mitigating misspecification, with performance matching on ROA reducing the rejection frequencies to 34.0%. These results are broadly consistent using the DD model. The Jones, Modified Jones and McNichols models, in contrast, result in less excessive rejection frequencies for all tests. This is because these models control for sales growth, alleviating the correlated omitted variable problem.

We close this subsection by summarizing the key findings:
i. Misspecification due to correlated omitted variables associated with economic characteristics is a pervasive problem in tests for earnings management. There is no panacea for this problem. Performance matching only mitigates the problem when the matching procedure happens to identify the appropriate omitted variable(s). When it does not, it can even make the problem worse.

ii. Tests for earnings management that incorporate accrual reversals mitigate misspecification due to correlated omitted variables for a range of economic characteristics. This approach does not require the correlated economic characteristics be known, but just that they do not reverse in the same period as the earnings management.

iii. In the case of highly persistent correlated omitted economic characteristics, incorporating reversals over two or more subsequent periods can lead to an over-correction problem. Modeling reversals over only one period avoids this problem.

6. **Conclusions and Implications**

   A key feature of the accrual accounting process is that accrual distortions in one period must reverse in another period. In the case of working capital accruals that typically span less than a year, such reversals generally occur within a year or so of the original distortion. We provide a flexible framework for incorporating these reversals in tests for earnings management and show that existing accrual-based tests for earnings management can be significantly improved by incorporating accrual reversals.

   We close by providing some caveats and guidelines for earnings management research. Our caveats are twofold:

   i. The power of accrual-based tests for earnings management is low for earnings management of plausible magnitudes. For example, our simulations indicate that even the most powerful tests that incorporate reversals reject the null hypothesis of no earnings management less than 30% of the time with earnings management equal to 1% of total assets, a sample size of 100 firms and a test level of 5%.  

   

39
ii. Tests for earnings management are susceptible to misspecification due to the omission of correlated determinants of nondiscretionary accruals. There is no panacea for this problem, because models of nondiscretionary accruals are crude and determinants of earnings management are often correlated with economic characteristics that influence nondiscretionary accruals.

In terms of guidelines for future research, we emphasize that the appropriate choice of test for earnings management will depend on the specific research setting being considered. With this in mind, we offer the following broad guidelines:

i. If the researcher has reasonable priors concerning the timing of the reversal of earnings management, then tests incorporating these priors will generally have improved power and specification. In the absence of specific priors, a reasonable assumption is that working capital accruals will reverse in the subsequent year (e.g., Allen, Larson and Sloan, [2010]).

ii. In terms of selecting an appropriate model of nondiscretionary accruals the researcher should consider economic characteristics that are likely to be correlated with the hypothesized earnings management. For example, if the hypothesized earnings management is correlated with firm growth, use of the Jones or Modified Jones models should alleviate omitted variable bias. At the same time, the researcher should also consider the nature of the hypothesized earnings management so as to avoid reducing test power by unintentionally classifying discretionary accruals as nondiscretionary. For example, if earnings are managed through deferred revenues, both the Jones and Modified Jones models will tend to classify the earnings management as nondiscretionary. Similarly, if earnings are managed in order to smooth the underlying cash flows, the DD model will tend to classify the earnings management as nondiscretionary.

iii. We suggest caution in the use of performance matching. Performance matching is generally only effective when the correlated omitted variables are known and can therefore be used to identify an appropriate matched pair. Performance matching also entails a significant reduction in test power and so is undesirable in cases where the researcher’s loss function places a relatively high weight on type II errors.
Finally, while we follow prior research in restricting our analysis to aggregate working capital accruals, the reversal framework can be extended to individual working capital accrual accounts and to long-term accruals (e.g., PP&E and purchased goodwill). An advantage of focusing on specific accrual accounts is that researchers could directly observe and measure write-downs of previous accrual overstatements (e.g., inventory write-downs, goodwill impairments).
REFERENCES


FEDYK, T.; Z. SINGER; and T. SOUGIANNIS. ‘Does the Accrual Anomaly End When Abnormal Accruals Reverse?’ Unpublished paper, Arizona State University, McGill University and University of Illinois at Urbana-Champaign, 2011.


Table 1. Descriptive statistics on working capital accruals

Panel A. Distribution of working capital accruals

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>5% Lower Quartile</th>
<th>Median</th>
<th>Upper Quartile</th>
<th>95% Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>WC_ACC</td>
<td>0.021</td>
<td>0.124</td>
<td>-0.145</td>
<td>-0.025</td>
<td>0.010</td>
<td>0.058</td>
<td>0.218</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.698</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7.928</td>
</tr>
</tbody>
</table>

Panel B. Pearson correlations

<table>
<thead>
<tr>
<th></th>
<th>Earnings (t)</th>
<th>WC_ACC (t+1)</th>
<th>WC_ACC (t+2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WC_ACC (t)</td>
<td>0.180*</td>
<td>0.047*</td>
<td>0.040*</td>
</tr>
</tbody>
</table>

* Significant at 1% level.

Panel C. Descriptive statistics on working capital accruals for deciles formed on earnings performance

<table>
<thead>
<tr>
<th>Rank of Earnings</th>
<th>Mean of WC_ACC</th>
<th>Standard Deviation of WC_ACC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Year t</td>
<td>Year t+1</td>
</tr>
<tr>
<td>1</td>
<td>-0.041</td>
<td>-0.005</td>
</tr>
<tr>
<td>2</td>
<td>-0.012</td>
<td>-0.010</td>
</tr>
<tr>
<td>3</td>
<td>0.004</td>
<td>-0.006</td>
</tr>
<tr>
<td>4</td>
<td>0.013</td>
<td>0.005</td>
</tr>
<tr>
<td>5</td>
<td>0.018</td>
<td>0.013</td>
</tr>
<tr>
<td>6</td>
<td>0.023</td>
<td>0.017</td>
</tr>
<tr>
<td>7</td>
<td>0.031</td>
<td>0.026</td>
</tr>
<tr>
<td>8</td>
<td>0.039</td>
<td>0.036</td>
</tr>
<tr>
<td>9</td>
<td>0.051</td>
<td>0.045</td>
</tr>
<tr>
<td>10</td>
<td>0.081</td>
<td>0.057</td>
</tr>
</tbody>
</table>

Panel D. Descriptive statistics on working capital accruals for the AAER sample

<table>
<thead>
<tr>
<th></th>
<th>Mean of WC_ACC</th>
<th>Standard Deviation of WC_ACC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Year t</td>
<td>Year t+1</td>
</tr>
<tr>
<td></td>
<td>0.077</td>
<td>-0.047</td>
</tr>
</tbody>
</table>

* Variable definitions (Compustat mnemonics in brackets)

WC_ACC_i,t = (ΔCA_i,t – ΔCL_i,t – ΔCash_i,t + ΔSTD_i,t)/A_{i,t-1}; Working capital accruals

ΔCA: Change in current assets (act)
ΔCL: Change in current liabilities (lct)
ΔCash: Change in cash (che)
ΔSTD: Change in short-term debt (dlc)
A: Total assets (at)

Earnings_{i,t} = Earnings before extraordinary items_{i,t}/A_{i,t-1}

This table reports descriptive statistics on working capital accruals. Panel A reports descriptive statistics on working capital accruals for 209,530 firm year observations between 1950 and 2009. In selecting the 209,530 firm year observations, positive non-missing values are required for receivables (rect), current assets (act), current liabilities (lct), cash and equivalent (che), short-term debt (std), sales (sale) and PP&E (ppegt). It is also required that income before extraordinary items (ib) is non-missing. Characters in parentheses indicate Compustat mnemonics. Panel B reports Pearson correlations between working capital accruals in year t and earnings in year t, working capital accruals in year t+1 and working capital accruals in year t+2. The number of firm year observations used in this correlation analysis is 172,969. Panel C reports means and standard deviations of working capital accruals in year t, year t+1 and year t+2 for each earnings performance decile rank. Earnings performance is defined as earnings before extraordinary items (ib) divided by the beginning total assets (at). Earnings performance decile ranks are formed with 209,530 firm year observations between 1950 and 2009. Rank 1 is the lowest earnings performance decile and rank 10 is the highest earnings performance decile. Panel D reports means and standard deviations of working capital accruals in year t, year t+1 and year t+2 for 406 SEC Accounting and Auditing Enforcement Release (AAER) sample firm years that were subject to enforcement actions by the SEC for upward annual earnings management between 1971 and 2003.
Table 2. Descriptive statistics and rejection rates for tests of earnings management when earnings management years are randomly selected

Panel A. Mean coefficients (t-statistics) for earnings management model parameters, and mean performance matched discretionary accruals (t-statistic)

<table>
<thead>
<tr>
<th>Model Parameter Estimates</th>
<th>Mean Adjusted R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>b on PART</td>
<td>c on PARTP1</td>
</tr>
<tr>
<td>Healy</td>
<td>0.000</td>
</tr>
<tr>
<td>Jones</td>
<td>0.000</td>
</tr>
<tr>
<td>Modified Jones</td>
<td>0.000</td>
</tr>
<tr>
<td>DD</td>
<td>0.000</td>
</tr>
<tr>
<td>McNichols</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Performance Matched Discretionary Accruals

<table>
<thead>
<tr>
<th>Matched on ROA(t-1)</th>
<th>Matched on ROA(t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Healy</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(-0.009)</td>
</tr>
<tr>
<td>Jones</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(-0.007)</td>
</tr>
<tr>
<td>Modified Jones</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(-0.012)</td>
</tr>
<tr>
<td>DD</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(-0.002)</td>
</tr>
<tr>
<td>McNichols</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(-0.006)</td>
</tr>
</tbody>
</table>

Panel B. Rejection rates

<table>
<thead>
<tr>
<th>Healy</th>
<th>Significantly Positive</th>
<th>Significantly Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Level: One-tailed 5%</td>
<td>b = 0</td>
<td>4.1%</td>
</tr>
<tr>
<td></td>
<td>b – c = 0</td>
<td>6.3%</td>
</tr>
<tr>
<td></td>
<td>b – c – d = 0</td>
<td>8.8%**</td>
</tr>
<tr>
<td>Performance matched on ROA(t-1)</td>
<td>5.2%</td>
<td>5.0%</td>
</tr>
<tr>
<td>Performance matched on ROA(t)</td>
<td>5.5%</td>
<td>4.8%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Jones</th>
<th>Significantly Positive</th>
<th>Significantly Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Level: One-tailed 5%</td>
<td>b = 0</td>
<td>4.3%</td>
</tr>
<tr>
<td></td>
<td>b – c = 0</td>
<td>5.6%</td>
</tr>
<tr>
<td></td>
<td>b – c – d = 0</td>
<td>6.0%</td>
</tr>
<tr>
<td>Performance matched on ROA(t-1)</td>
<td>5.1%</td>
<td>5.5%</td>
</tr>
<tr>
<td>Performance matched on ROA(t)</td>
<td>5.2%</td>
<td>5.3%</td>
</tr>
</tbody>
</table>

Panel B continued
Table 2 Panel B continued

<table>
<thead>
<tr>
<th>Modified Jones</th>
<th>Significantly Positive</th>
<th>Significantly Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test Level: One-tailed 5%</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b = 0</td>
<td>4.1%</td>
<td>6.5%*</td>
</tr>
<tr>
<td>b – c = 0</td>
<td>6.0%</td>
<td>5.0%</td>
</tr>
<tr>
<td>b – c – d = 0</td>
<td>6.6%*</td>
<td>4.6%</td>
</tr>
<tr>
<td>Performance matched on ROA(t-1)</td>
<td>5.2%</td>
<td>5.4%</td>
</tr>
<tr>
<td>Performance matched on ROA(t)</td>
<td>5.3%</td>
<td>5.1%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DD</th>
<th>Significantly Positive</th>
<th>Significantly Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test Level: One-tailed 5%</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b = 0</td>
<td>4.7%</td>
<td>5.9%</td>
</tr>
<tr>
<td>b – c = 0</td>
<td>5.1%</td>
<td>2.2%**</td>
</tr>
<tr>
<td>b – c – d = 0</td>
<td>7.9%**</td>
<td>2.3%**</td>
</tr>
<tr>
<td>Performance matched on ROA(t-1)</td>
<td>4.7%</td>
<td>4.7%</td>
</tr>
<tr>
<td>Performance matched on ROA(t)</td>
<td>5.0%</td>
<td>4.4%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>McNichols</th>
<th>Significantly Positive</th>
<th>Significantly Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test Level: One-tailed 5%</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b = 0</td>
<td>5.1%</td>
<td>4.9%</td>
</tr>
<tr>
<td>b – c = 0</td>
<td>4.6%</td>
<td>3.3%*</td>
</tr>
<tr>
<td>b – c – d = 0</td>
<td>5.1%</td>
<td>4.3%</td>
</tr>
<tr>
<td>Performance matched on ROA(t-1)</td>
<td>4.9%</td>
<td>4.6%</td>
</tr>
<tr>
<td>Performance matched on ROA(t)</td>
<td>5.0%</td>
<td>5.0%</td>
</tr>
</tbody>
</table>

***, * Significantly different from the specified test level at 1 and 5 percent level respectively using a two-tailed binomial test

Panel C. Serial correlations in discretionary accruals and nondiscretionary accruals

<table>
<thead>
<tr>
<th></th>
<th>Mean Serial Correlations for All Available Firms</th>
<th>Mean Serial Correlations for Available AAER Firms</th>
<th>Serial Correlations for Earnings Management and Reversal Years in AAER Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Discretionary</td>
<td>Nondiscretionary</td>
<td>Discretionary</td>
</tr>
<tr>
<td>Healy</td>
<td>-0.094</td>
<td>N/A</td>
<td>-0.017</td>
</tr>
<tr>
<td>Jones</td>
<td>-0.206</td>
<td>0.200</td>
<td>-0.165</td>
</tr>
<tr>
<td>Modified Jones</td>
<td>-0.180</td>
<td>0.177</td>
<td>-0.137</td>
</tr>
<tr>
<td>DD</td>
<td>0.077</td>
<td>-0.199</td>
<td>0.065</td>
</tr>
<tr>
<td>McNichols</td>
<td>-0.080</td>
<td>-0.127</td>
<td>-0.112</td>
</tr>
</tbody>
</table>

Panel A reports mean coefficients (t-statistics) for various earnings management models, and mean performance matched discretionary accruals (t-statistics) estimated from each model when earnings management years are randomly selected. Panel B reports the rejection rates for tests of the null hypothesis of no earnings management when earnings management years are randomly selected. Panel C reports mean firm specific serial correlations of discretionary accruals and nondiscretionary accruals for (i) all available firms; (ii) AAER firms; and (iii) AAER firm-years that involve the last year of an earnings management allegation or the first year immediately following an earnings management allegation. For each reversal model, tests of earnings management are conducted based on the respective model below.

Healy: \( WC_{ACC_{it}} = a + b PART_{it} + c PARTP1_{it} + d PARTP2_{it} + e_{it} \)

where \( WC_{ACC_{it}} = (\Delta CA_{it} - \Delta CL_{it} - \Delta Cash_{it} + \Delta STD_{it})/A_{it-1} \): Working capital accruals

\( \Delta CA \): Change in current assets (act)
\( \Delta CL \): Change in current liabilities (lct)
\( \Delta Cash \): Change in cash (che)
\( \Delta STD \): Change in short-term debt (dlc)

\( A \): Total assets (at)

\( PART_{it} \): Partitioning variable that is set to 1 for randomly selected 100 firm years and 0 otherwise

\( PARTP1_{it} \): Partitioning variable that is set to 1 in the year following randomly selected 100 firm years and 0 otherwise

\( PARTP2_{it} \): Partitioning variable that is set to 1 in the second year following randomly selected 100 firm years and 0 otherwise

* Characters in brackets are Compustat mnemonics and the same definitions apply to the following models.
Jones: WC_ACC\(_{it}\) = \(a + b\text{PART}_{it} + c\text{PARTP1}_{it} + d\text{PARTP2}_{it} + f_1\Delta\text{REV}_{it} + f_2\text{PPE}_{it} + e_{it}\)

\[\Delta\text{REV}_{it} = (\text{Revenue}_{it} - \text{Revenue}_{it-1})/\text{Assets}_{i,t}; \text{ Change in revenue (sale)}\]

\[\text{PPE}_{it} = \text{PP&E}_{it}/\text{Assets}_{i,t}; \text{ Gross property, plant and equipment (ppeg)}\]

Modified Jones: WC_ACC\(_{it}\) = \(a + b\text{PART}_{it} + c\text{PARTP1}_{it} + d\text{PARTP2}_{it} + f_1\Delta\text{REV}_{it} + f_2\text{PPE}_{it} + e_{it}\)

\[\Delta\text{REV}_{it} = (\text{Revenue}_{it} - \text{Revenue}_{it-1}) - (\text{Net Accounts Receivable}_{it} - \text{Net Accounts Receivable}_{it-1})/\text{Assets}_{i,t}; \text{ Change in revenue (sale)} – \text{ Change in net accounts receivable (rect)}\]

\[\text{PPE}_{it} = \text{PP&E}_{it}/\text{Assets}_{i,t}; \text{ Gross property, plant and equipment (ppeg)}\]

DD: WC_ACC\(_{it}\) = \(a + b\text{PART}_{it} + c\text{PARTP1}_{it} + d\text{PARTP2}_{it} + f_1\Delta\text{REV}_{it} + f_2\text{PPE}_{it} + f_3\text{CF}_{it} + e_{it}\)

\[\Delta\text{REV}_{it} = \text{Change in revenue (sale)}\]

\[\text{PPE}_{it} = \text{PP&E}_{it}/\text{Assets}_{i,t}; \text{ Gross property, plant and equipment (ppeg)}\]

CF\(_{it}\) = Earnings before Extraordinary Items\(_{i,t}\) (ib) – WC_ACC\(_{it}\)

McNichols: WC_ACC\(_{it}\) = \(a + b\text{PART}_{it} + c\text{PARTP1}_{it} + d\text{PARTP2}_{it} + f_1\Delta\text{REV}_{it} + f_2\text{PPE}_{it} + f_3\text{CF}_{it} + f_4\text{PPE}_{i,t+1} + e_{it}\)

\[\Delta\text{REV}_{it} = \text{Change in revenue (sale)}\]

\[\text{PPE}_{it} = \text{PP&E}_{it}/\text{Assets}_{i,t}; \text{ Gross property, plant and equipment (ppeg)}\]

CF\(_{it}\) = Earnings before Extraordinary Items\(_{i,t}\) (ib) – WC_ACC\(_{it}\)

For each performance matching model, discretionary accruals are estimated based on the respective model below:

Healy: Discretionary accruals = WC_ACC\(_{it}\)

Jones: Discretionary accruals are residuals from WC_ACC\(_{it}\) = \(a + f_1\Delta\text{REV}_{it} + f_2\text{PPE}_{it} + e_{it}\)

\[\Delta\text{REV}_{it} = (\text{Revenue}_{it} - \text{Revenue}_{it-1})/\text{Assets}_{i,t}; \text{ Change in revenue (sale)}\]

\[\text{PPE}_{it} = \text{PP&E}_{it}/\text{Assets}_{i,t}; \text{ Gross property, plant and equipment (ppeg)}\]

Modified Jones: Discretionary accruals are residuals from WC_ACC\(_{it}\) = \(a + f_1\Delta\text{REV}_{it} + f_2\text{PPE}_{it} + e_{it}\)

\[\Delta\text{REV}_{it} = (\text{Revenue}_{it} - \text{Revenue}_{it-1}) - (\text{Net Accounts Receivable}_{it} - \text{Net Accounts Receivable}_{it-1})/\text{Assets}_{i,t}; \text{ Change in revenue (sale)} – \text{ Change in net accounts receivable (rect)}\]

\[\text{PPE}_{it} = \text{PP&E}_{it}/\text{Assets}_{i,t}; \text{ Gross property, plant and equipment (ppeg)}\]

DD: Discretionary accruals are residuals from WC_ACC\(_{it}\) = \(a + f_1\text{CF}_{i,t} + f_2\text{CF}_{i,t+1} + f_3\text{PPE}_{i,t} + e_{it}\)

\[\text{CF}_{i,t}\) = Earnings before Extraordinary Items\(_{i,t}\) (ib) – WC_ACC\(_{it}\)

McNichols: Discretionary accruals are residuals from WC_ACC\(_{it}\) = \(a + f_1\Delta\text{REV}_{it} + f_2\text{PPE}_{it} + f_3\text{CF}_{i,t} + f_4\text{CF}_{i,t+1} + f_5\text{CF}_{i,t+1} + e_{it}\)

\[\Delta\text{REV}_{it} = (\text{Revenue}_{it} - \text{Revenue}_{i,t-1})/\text{Assets}_{i,t}; \text{ Change in revenue (sale)}\]

\[\text{PPE}_{it} = \text{PP&E}_{it}/\text{Assets}_{i,t}; \text{ Gross property, plant and equipment (ppeg)}\]

\[\text{CF}_{i,t}\) = Earnings before Extraordinary Items\(_{i,t}\) (ib) – WC_ACC\(_{it}\)

The first panel of panel A reports mean coefficients and t-statistics for each model based on 1,000 pooled regressions. Using 209,530 firm year observations between 1950 and 2009, 100 firm years are randomly selected as earnings management years and PART is set to 1 for those 100 firm years (zero otherwise). Additionally, PARTP1 is set to 1 for the first year following those 100 earnings management years (zero otherwise) and PARTP2 is set to 1 for the second year following those 100 earnings management years (zero otherwise). Then, a pooled regression is run using all 209,530 observations for each model. This procedure is repeated 1,000 times. The second panel of panel A reports mean performance matched discretionary accruals and associated t-statistics for each model, based on 1,000 replications selected 100 observations for the reversal models. For performance matching on ROA\(_{t-1}\), each of 100 randomly selected observations is matched with another observation from the same two-digit SIC code with the closest return on assets in the current year. Performance matched discretionary accruals for each of 100 randomly selected observations are then computed by subtracting the matched firm’s discretionary accruals. Performance matching on ROA\(_{t-1}\) is done in a similar manner to performance matching on ROA\(_{t}\), based on the closest prior year’s return on assets. Panel B reports the rejection rates of no earnings management based on PART only (i.e., b=0), PART and PARTP1 (i.e., b-c=0), PART, PARTP1 and PARTP2 (i.e., b-c-d=0), and performance matched discretionary accruals. For reversal models, after each pooled regression is run, χ²-tests are conducted based on heteroscedasticity-consistent standard errors. This procedure is repeated 1,000 times and the percent of the 1000 regressions is determined where the χ²-statistic is significant at the five percent level (one-tailed test). Then, a binomial test is performed to determine if the percentage is significantly different from the specified test level (i.e., 5%). Note that b is the coefficient on PART, c is the coefficient on PARTP1 and d is coefficient on PARTP2. For performance matching models, t-tests are conducted to examine whether the mean performance matched discretionary accruals of 100 randomly selected observations are significant at five percent level (one-tailed test). This procedure is repeated 1000 times and the percent of the 1000 repetitions is determined where the t-statistic is significant at the five percent level (one-tailed test). Then, a binomial test is performed to determine if the percentage is significantly different from the specified test level (i.e., 5%). The first two columns of panel C report mean firm specific serial correlations of discretionary accruals and nondiscretionary accruals estimated from various earnings management test models for all available firms and AAER firms. The serial correlations of discretionary accruals and nondiscretionary accruals are Pearson correlations between year t and year t+1 for each firm. To estimate firm specific discretionary accruals and nondiscretionary accruals, only firms with six or more observations are selected. Except for the Healy model, discretionary accruals are residuals from each
earnings management test model and nondiscretionary accruals are predicted values from each earnings management test model. For the Healy model, discretionary accruals are just working capital accruals and there are no nondiscretionary accruals. For the first two columns of panel C, 11,342 firms (177,814 firm years) are used in the case of all available firms, and 194 firms (2,789 firm years) are used in the case of AAER firms. The last column of panel C documents serial correlations of discretionary accruals and nondiscretionary accruals between the last earning management years (if earnings are alleged to be managed for several consecutive years) and the first reversal years for available AAER firms. For this end, the firm specific discretionary accruals and nondiscretionary accruals are first estimated for the AAER firms with six or more observations. Then, using the discretionary accruals and nondiscretionary accruals for the last earnings management years and the first reversal years, Pearson correlations are computed across available AAER firms for each model. For the last of panel C, 151 AAER firms are used.
Table 3.  Power of the models in detecting earnings management for the AAER sample.

Panel A. Coefficients (t-statistics) for each model

<table>
<thead>
<tr>
<th></th>
<th>Healy</th>
<th>Jones</th>
<th>Modified Jones</th>
<th>DD</th>
<th>McNichols</th>
</tr>
</thead>
<tbody>
<tr>
<td>b, coefficient on PART</td>
<td>0.057</td>
<td>0.033</td>
<td>0.044</td>
<td>0.050</td>
<td>0.028</td>
</tr>
<tr>
<td></td>
<td>(5.987)</td>
<td>(3.858)</td>
<td>(4.736)</td>
<td>(6.046)</td>
<td>(3.714)</td>
</tr>
<tr>
<td>c, coefficient on PAR(T_{\text{P1}})</td>
<td>-0.066</td>
<td>-0.064</td>
<td>-0.068</td>
<td>-0.067</td>
<td>-0.065</td>
</tr>
<tr>
<td></td>
<td>(-5.666)</td>
<td>(-6.146)</td>
<td>(-6.163)</td>
<td>(-6.356)</td>
<td>(-6.760)</td>
</tr>
<tr>
<td>d, coefficient on PAR(T_{\text{P2}})</td>
<td>-0.055</td>
<td>-0.044</td>
<td>-0.050</td>
<td>-0.050</td>
<td>-0.043</td>
</tr>
<tr>
<td></td>
<td>(-4.863)</td>
<td>(-4.245)</td>
<td>(-4.552)</td>
<td>(-4.849)</td>
<td>(-4.347)</td>
</tr>
<tr>
<td>∆ Revenue</td>
<td>0.109</td>
<td>0.089</td>
<td>0.101</td>
<td>-0.001</td>
<td>-0.014</td>
</tr>
<tr>
<td></td>
<td>(73.708)</td>
<td>(52.962)</td>
<td>(74.903)</td>
<td>(31.271)</td>
<td>(34.238)</td>
</tr>
<tr>
<td>PPE</td>
<td>-0.009</td>
<td>-0.008</td>
<td>-0.014</td>
<td>-0.014</td>
<td>-0.014</td>
</tr>
<tr>
<td></td>
<td>(-10.997)</td>
<td>(-9.224)</td>
<td>(-18.133)</td>
<td>(-39.660)</td>
<td>(34.745)</td>
</tr>
<tr>
<td>Partial R(^2)</td>
<td>0.10%</td>
<td>13.20%</td>
<td>6.79%</td>
<td>19.99%</td>
<td>31.08%</td>
</tr>
</tbody>
</table>

Panel B. Performance matched discretionary accruals (t-statistic) for each model

<table>
<thead>
<tr>
<th></th>
<th>Healy</th>
<th>Jones</th>
<th>Modified Jones</th>
<th>DD</th>
<th>McNichols</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean discretionary accruals</td>
<td>0.054</td>
<td>0.036</td>
<td>0.045</td>
<td>0.053</td>
<td>0.037</td>
</tr>
<tr>
<td>performance matched on ROA(t-1)</td>
<td>(4.710)</td>
<td>(3.500)</td>
<td>(4.060)</td>
<td>(5.380)</td>
<td>(4.090)</td>
</tr>
<tr>
<td>Mean discretionary accruals</td>
<td>0.057</td>
<td>0.038</td>
<td>0.047</td>
<td>0.056</td>
<td>0.039</td>
</tr>
<tr>
<td>performance matched on ROA(t)</td>
<td>(5.100)</td>
<td>(3.740)</td>
<td>(4.290)</td>
<td>(5.790)</td>
<td>(4.320)</td>
</tr>
</tbody>
</table>

Panel C. Significance levels of test statistics (p-values) for each model

<table>
<thead>
<tr>
<th></th>
<th>Healy</th>
<th>Jones</th>
<th>Modified Jones</th>
<th>DD</th>
<th>McNichols</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\chi^2)-statistic for b = 0</td>
<td>35.849</td>
<td>14.887</td>
<td>22.428</td>
<td>36.556</td>
<td>13.793</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>(\chi^2)-statistic for b – c = 0</td>
<td>66.883</td>
<td>51.869</td>
<td>60.380</td>
<td>76.508</td>
<td>57.976</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>(\chi^2)-statistic for b – c – d = 0</td>
<td>89.524</td>
<td>68.748</td>
<td>79.861</td>
<td>97.626</td>
<td>74.961</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>F-statistic for zero discretionary accruals</td>
<td>22.184</td>
<td>12.250</td>
<td>16.484</td>
<td>28.944</td>
<td>16.728</td>
</tr>
<tr>
<td>performance matched on ROA(t-1)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>F-statistic for zero discretionary accruals</td>
<td>26.010</td>
<td>13.988</td>
<td>18.404</td>
<td>33.524</td>
<td>18.662</td>
</tr>
<tr>
<td>performance matched on ROA(t)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
</tbody>
</table>

This table compares the power of tests for earnings management in the AAER sample. Tests of earnings management are conducted using the models below:

**Healy:** \( WC_{\text{ACC}}_{\text{it}} = a + b\text{PART}_{\text{it}} + c\text{PAR}\text{\(T_{\text{P1}}\)}_{\text{it}} + d\text{PAR}\text{\(T_{\text{P2}}\)}_{\text{it}} + e_{\text{it}} \)

where

\( WC_{\text{ACC}}_{\text{it}} = (\Delta\text{CA}_{\text{it}} - \Delta\text{CL}_{\text{it}} - \Delta\text{Cash}_{\text{it}} + \Delta\text{STD}_{\text{it}})/\text{A}_{\text{it} - 1} \): Working capital accruals

\( \Delta\text{CA} \): Change in current assets (act)

\( \Delta\text{CL} \): Change in current liabilities (lct)

\( \Delta\text{Cash} \): Change in cash (che)

\( \Delta\text{STD} \): Change in short-term debt (dlc)
A: Total assets (at)
PART, Partitioning variable that is set to 1 for 406 AAER firm years in which upward earnings management is alleged to have taken place
PARTP1, Partitioning variable that is set to 1 for the first year following the final earnings management year and 0 otherwise
PARTP2, Partitioning variable that is set to 1 for the second year following the final earnings management year and 0 otherwise
* Characters in brackets are Compustat mnemonics and the same definitions apply to the following models.

Jones: WC_ACCt = a + bPARTit + cPARTP1it + dPARTP2it + f1AREVit + f2PPEit + eit

\[ \Delta \text{REV}_it = (\text{Revenue}_{it} - \text{Revenue}_{it-1}) / A_{it-1}; \text{Change in revenue (sale)} \]
PPEit = PP&Eit / Ait; Gross property, plant and equipment (ppegt)

Modified Jones: WC_ACCt = a + bPARTit + cPARTP1it + dPARTP2it + f1AREVit + f2PPEit + eit

\[ \Delta \text{REV}_it = (\text{Revenue}_{it} - \text{Revenue}_{it-1}) - (\text{Net Accounts Receivable}_{it} - \text{Net Accounts Receivable}_{it-1}) / A_{it-1}; \text{Change in revenue (sale)} \]
PPEit = PP&Eit / Ait; Gross property, plant and equipment (ppegt)

DD: WC_ACCit = a + bPARTit + cPARTP1it + dPARTP2it + f1AREVit + f2PPEit + f3CFit+1 + f4CFit + f5CFit+1 + eit

CFit = Earnings before Extraordinary Itemsit (ib) – WC_ACCit

McNichols: WC_ACCit = a + bPARTit + cPARTP1it + dPARTP2it + f1CFit+1 + f2CFit + f3CFit+1 + eit

For each performance matching model, discretionary accruals are estimated based on the respective model below;

Healy: Discretionary accruals = WC_ACCit

Jones: Discretionary accruals are residuals from WC_ACCit = a + f1AREVit + f2PPEit + eit

\[ \Delta \text{REV}_it = (\text{Revenue}_{it} - \text{Revenue}_{it-1}) / A_{it-1}; \text{Change in revenue (sale)} \]
PPEit = PP&Eit / Ait; Gross property, plant and equipment (ppegt)

Modified Jones: Discretionary accruals are residuals from WC_ACCit = a + f1AREVit + f2PPEit + eit

\[ \Delta \text{REV}_it = (\text{Revenue}_{it} - \text{Revenue}_{it-1}) - (\text{Net Accounts Receivable}_{it} - \text{Net Accounts Receivable}_{it-1}) / A_{it-1}; \text{Change in revenue (sale)} \]
PPEit = PP&Eit / Ait; Gross property, plant and equipment (ppegt)

DD: Discretionary accruals are residuals from WC_ACCit = a + f1CFit+1 + f2CFit + f3CFit+1 + eit

CFit = Earnings before Extraordinary Itemsit (ib) – WC_ACCit

McNichols: Discretionary accruals are residuals from WC_ACCit = a + f1AREVit + f2PPEit + f3CFit+1 + f4CFit + f5CFit+1 + eit

\[ \Delta \text{REV}_it = (\text{Revenue}_{it} - \text{Revenue}_{it-1}) / A_{it-1}; \text{Change in revenue (sale)} \]
PPEit = PP&Eit / Ait; Gross property, plant and equipment (ppegt)

CFit = Earnings before Extraordinary Itemsit (ib) – WC_ACCit

Panel A reports the coefficients and t-statistics for each model based on a pooled regression using 161,119 observations between 1971 and 2003. PART is set to 1 for 406 AAER firm years between 1971 and 2003 (zero otherwise), PARTP1 is set to 1 for the first year following those 406 AAER firm years (zero otherwise) and PARTP2 is set to 1 for the second year following those 406 AAER firm years (zero otherwise). If earnings are managed for several consecutive years, PARTP1 (PARTP2) is set to 1 for the first (second) year following the final year of earnings management. Panel B reports the mean performance matched discretionary accruals and associated t-statistics of 406 AAER firm years between 1971 and 2003 for each model. For performance matching on ROA(t), observations are matched with another observation from the same two-digit SIC code with the closest return on assets in the current year. Performance matched discretionary accruals are then computed by subtracting the matched firm’s discretionary accruals. Performance matching on ROA(t-1) is done in a similar manner to performance matching on ROA(t), based on the closest prior year’s return on assets. Panel C reports results of testing earnings management for 406 AAER firm years between 1971 and 2003, based on PART only (i.e., b=0), PART and PARTP1 (i.e., b=0), PART, PARTP1 and PARTP2 (i.e., b-c=0), and performance matched discretionary accruals. For reversal models, a pooled cross-sectional regression is run using all 161,119 observations and \( \chi^2 \)-tests are conducted based on heteroscedasticity-consistent standard errors. Then, \( \chi^2 \)-statistics and associated p-values are reported. Note that b is the coefficient on PART, c is the coefficient on PARTP1 and d is coefficient on PARTP2. For performance matching models, t-tests are conducted for performance matched discretionary accruals of 406 AAER firm years, and F-statistics and associated p-values are reported by squaring t-statistics to maintain comparability with \( \chi^2 \)-statistics used for reversal models.
Table 4. Power of the models in detecting earnings management for a subset of the AAER sample involving specific allegations of working capital accruals manipulations.

Panel A. Coefficients (t-statistics) for each model

<table>
<thead>
<tr>
<th></th>
<th>Healy</th>
<th>Jones</th>
<th>Modified Jones</th>
<th>DD</th>
<th>McNichols</th>
</tr>
</thead>
<tbody>
<tr>
<td>b, coefficient on PART</td>
<td>0.084</td>
<td>0.059</td>
<td>0.069</td>
<td>0.074</td>
<td>0.051</td>
</tr>
<tr>
<td></td>
<td>(4.774)</td>
<td>(3.648)</td>
<td>(4.023)</td>
<td>(4.743)</td>
<td>(3.642)</td>
</tr>
<tr>
<td>c, coefficient on PARTP1</td>
<td>-0.086</td>
<td>-0.077</td>
<td>-0.084</td>
<td>-0.085</td>
<td>-0.076</td>
</tr>
<tr>
<td></td>
<td>(-4.438)</td>
<td>(-3.995)</td>
<td>(-4.207)</td>
<td>(-4.745)</td>
<td>(-4.221)</td>
</tr>
<tr>
<td>d, coefficient on PARTP2</td>
<td>-0.089</td>
<td>-0.079</td>
<td>-0.086</td>
<td>-0.078</td>
<td>-0.070</td>
</tr>
<tr>
<td></td>
<td>(-4.930)</td>
<td>(-4.343)</td>
<td>(-4.605)</td>
<td>(-5.320)</td>
<td>(-4.745)</td>
</tr>
<tr>
<td>∆ Revenue</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.109</td>
<td>0.089</td>
<td></td>
<td>0.101</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(73.743)</td>
<td>(53.004)</td>
<td>(74.946)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.009</td>
<td>-0.008</td>
<td>-0.014</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-10.930)</td>
<td>(-9.166)</td>
<td>(18.051)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>∆ Revenue</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean discretionary accruals</td>
<td>0.067</td>
<td>0.047</td>
<td>0.055</td>
<td>0.065</td>
<td>0.046</td>
</tr>
<tr>
<td>performance matched on ROA(t-1)</td>
<td>(3.260)</td>
<td>(2.400)</td>
<td>(2.700)</td>
<td>(3.510)</td>
<td>(2.580)</td>
</tr>
<tr>
<td>Mean discretionary accruals</td>
<td>0.083</td>
<td>0.064</td>
<td>0.071</td>
<td>0.076</td>
<td>0.058</td>
</tr>
<tr>
<td>performance matched on ROA(t)</td>
<td>(3.950)</td>
<td>(3.280)</td>
<td>(3.440)</td>
<td>(4.150)</td>
<td>(3.380)</td>
</tr>
</tbody>
</table>

Panel C. Significance levels of test statistics (p-values) for each model

<table>
<thead>
<tr>
<th></th>
<th>Healy</th>
<th>Jones</th>
<th>Modified Jones</th>
<th>DD</th>
<th>McNichols</th>
</tr>
</thead>
<tbody>
<tr>
<td>χ²-statistic for b = 0</td>
<td>22.792</td>
<td>13.309</td>
<td>16.185</td>
<td>22.494</td>
<td>13.263</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>χ²-statistic for b – c = 0</td>
<td>42.195</td>
<td>29.219</td>
<td>33.800</td>
<td>44.849</td>
<td>31.020</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>χ²-statistic for b – c – d = 0</td>
<td>66.363</td>
<td>47.938</td>
<td>54.864</td>
<td>72.144</td>
<td>52.651</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>F-statistic for zero discretionary accruals performance matched on ROA(t-1)</td>
<td>10.628</td>
<td>5.760</td>
<td>7.290</td>
<td>12.320</td>
<td>6.656</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.002)</td>
<td>(0.018)</td>
<td>(0.008)</td>
<td>(0.001)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>F-statistic for zero discretionary accruals performance matched on ROA(t)</td>
<td>15.603</td>
<td>10.758</td>
<td>11.834</td>
<td>17.223</td>
<td>11.424</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.000)</td>
<td>(0.001)</td>
</tr>
</tbody>
</table>

This table repeats the tests reported in Table 3 for 122 AAER firm years that are alleged to have manipulated earnings through working capital accruals (i.e., accounts receivable, inventory or accounts payable). Tests of earnings management are conducted using the models below:

Healy: \( WC_{ACCD_{it}} = a + b PART_{it} + c PARTP1_{it} + d PARTP2_{it} + e_{it} \)

where

\( WC_{ACCD_{it}} = (\Delta CA_{it} - \Delta CL_{it} - \Delta Cash_{it} + \Delta STD_{it})/A_{it-1} \): Working capital accruals
\( \Delta CA \): Change in current assets (act)
\( \Delta CL \): Change in current liabilities (lct)
\( \Delta Cash \): Change in cash (che)

- 51 -
ΔSTD: Change in short-term debt (dlc)
A: Total assets (at)
PART\(_i\): Partitioning variable that is set to 1 for 122 AAER firm years that are alleged to have manipulated earnings through working capital accruals (i.e., accounts receivable, inventory or accounts payable)
PARTP\(_1\): Partitioning variable that is set to 1 for the first year following the final earnings management year and 0 otherwise
PARTP\(_2\): Partitioning variable that is set to 1 for the second year following the final earnings management year and 0 otherwise

* Characters in brackets are Compubst mnemonics and the same definitions apply to the following models.

** Jones: **
\[ WC_{ACC,i,t} = a + b\text{PART}_{i,t} + c\text{PARTP}_{1,i} + d\text{PARTP}_{2,i} + f_1\text{REV}_{i,t} + f_2\ PPE_{i,t} + e_{i,t} \]
where
\[ \Delta\text{REV}_{i,t} = (\text{Revenue}_{i,t} - \text{Revenue}_{i,t-1})/A_{i,t}; \] Change in revenue (sale)
\[ PPE_{i,t} = \text{PP&E}_{i,t}/A_{i,t}; \] Gross property, plant and equipment (ppeg)

** Modified Jones: **
\[ WC_{ACC,i,t} = a + b\text{PART}_{i,t} + c\text{PARTP}_{1,i} + d\text{PARTP}_{2,i} + f_1\text{REV}_{i,t} + f_2\ PPE_{i,t} + e_{i,t} \]
where
\[ \Delta\text{REV}_{i,t} = (\text{Revenue}_{i,t} - \text{Revenue}_{i,t-1}) - (\text{Net Accounts Receivable}_{i,t} - \text{Net Accounts Receivable}_{i,t-1})/A_{i,t-1}; \] Change in revenue (sale) – Change in net accounts receivable (rect)
\[ PPE_{i,t} = \text{PP&E}_{i,t}/A_{i,t}; \] Gross property, plant and equipment (ppeg)

** DD: **
\[ WC_{ACC,i,t} = a + b\text{PART}_{i,t} + c\text{PARTP}_{1,i} + d\text{PARTP}_{2,i} + f_1\text{CF}_{i,t-1} + f_2\ CF_{i,t} + f_3\ PPE_{i,t} + e_{i,t} \]
where
\[ \text{CF}_{i,t} = \text{Earnings before Extraordinary Items}_{i,t} (ib) - \text{WC}_{ACC,i,t} \]

** McNichols: **
\[ WC_{ACC,i,t} = a + b\text{PART}_{i,t} + c\text{PARTP}_{1,i} + d\text{PARTP}_{2,i} + f_1\text{REV}_{i,t} + f_2\ PPE_{i,t} + f_3\ CF_{i,t} + f_4\ CF_{i,t-1} + f_5\ PPE_{i,t} + e_{i,t} \]
where
\[ \Delta\text{REV}_{i,t} = (\text{Revenue}_{i,t} - \text{Revenue}_{i,t-1})/A_{i,t}; \] Change in revenue (sale)
\[ PPE_{i,t} = \text{PP&E}_{i,t}/A_{i,t}; \] Gross property, plant and equipment (ppeg)

For each performance matching model, discretionary accruals are estimated based on the respective model below;

** Healy: **
Discretionary accruals = WC\(_{ACC,i,t}\)

** Jones: **
Discretionary accruals are residuals from WC\(_{ACC,i,t}\) = a + f_1\text{REV}_{i,t} + f_2\ PPE_{i,t} + e_{i,t}
where
\[ \Delta\text{REV}_{i,t} = (\text{Revenue}_{i,t} - \text{Revenue}_{i,t-1})/A_{i,t}; \] Change in revenue (sale)
\[ PPE_{i,t} = \text{PP&E}_{i,t}/A_{i,t}; \] Gross property, plant and equipment (ppeg)

** Modified Jones: **
Discretionary accruals are residuals from WC\(_{ACC,i,t}\) = a + f_1\text{REV}_{i,t} + f_2\ PPE_{i,t} + e_{i,t}
where
\[ \Delta\text{REV}_{i,t} = (\text{Revenue}_{i,t} - \text{Revenue}_{i,t-1}) - (\text{Net Accounts Receivable}_{i,t} - \text{Net Accounts Receivable}_{i,t-1})/A_{i,t-1}; \] Change in revenue (sale) – Change in net accounts receivable (rect)
\[ PPE_{i,t} = \text{PP&E}_{i,t}/A_{i,t}; \] Gross property, plant and equipment (ppeg)

** DD: **
Discretionary accruals are residuals from WC\(_{ACC,i,t}\) = a + f_1\text{CF}_{i,t-1} + f_2\ CF_{i,t} + f_3\ PPE_{i,t} + e_{i,t}
where
\[ \text{CF}_{i,t} = \text{Earnings before Extraordinary Items}_{i,t} (ib) - \text{WC}_{ACC,i,t} \]

** McNichols: **
Discretionary accruals are residuals from WC\(_{ACC,i,t}\) = a + f_1\text{REV}_{i,t} + f_2\ PPE_{i,t} + f_3\ CF_{i,t} + f_4\ CF_{i,t-1} + f_5\ PPE_{i,t} + e_{i,t}
where
\[ \Delta\text{REV}_{i,t} = (\text{Revenue}_{i,t} - \text{Revenue}_{i,t-1})/A_{i,t}; \] Change in revenue (sale)
\[ PPE_{i,t} = \text{PP&E}_{i,t}/A_{i,t}; \] Gross property, plant and equipment (ppeg)
\[ \text{CF}_{i,t} = \text{Earnings before Extraordinary Items}_{i,t} (ib) - \text{WC}_{ACC,i,t} \]

Panel A reports the coefficients and t-statistics for each model based on a pooled regression using 161,119 observations between 1971 and 2003. PART is set to 1 for 122 AAER firm years that are alleged to have manipulated earnings through working capital accruals (i.e., accounts receivable, inventory or accounts payable) between 1971 and 2003 (zero otherwise), PARTP1 is set to 1 for the first year following those 122 AAER firm years (zero otherwise). PARTP2 is set to 1 for the second year following those 122 AAER firm years (zero otherwise). If earnings are managed for several consecutive years, PARTP1 is set to 1 for the first (second) year following the final year of earnings management. Panel B reports the mean performance matched discretionary accruals and associated t-statistics for 122 AAER firm years that are alleged to have manipulated earnings through working capital accruals (i.e., accounts receivable, inventory or accounts payable) between 1971 and 2003, based on PART only (i.e., b=0), PART and PARTP1 (i.e., b=0), PART, PARTP1 and PARTP2 (i.e., b=0), and performance matched discretionary accruals. For reversal models, a pooled cross-sectional regression is run using all 161,119 observations and χ²-tests are conducted based on heteroscedasticity-consistent standard errors. Then, χ²-statistics and associated p-values are reported. Note that b is the
coefficient on PART, c is the coefficient on PARTP1 and d is coefficient on PARTP2. For performance matching models, t-tests are conducted for performance matched discretionary accruals of 122 AAER firm years that are alleged to have manipulated earnings through working capital accruals (i.e., accounts receivable, inventory or accounts payable), and F-statistics and associated p-values are reported by squaring t-statistics to maintain comparability with $\chi^2$-statistics used for reversal models.
Table 5. Rejection rates when earnings management years are selected from the highest and lowest deciles of earnings performance

<table>
<thead>
<tr>
<th></th>
<th>Significantly Positive</th>
<th>Significantly Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Highest</td>
<td>Lowest</td>
</tr>
<tr>
<td><strong>Healy</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Test Level: One-tailed 5%</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b = 0</td>
<td>98.8%**</td>
<td>0.0%**</td>
</tr>
<tr>
<td>b – c = 0</td>
<td>27.8%**</td>
<td>0.0%**</td>
</tr>
<tr>
<td>b – c – d = 0</td>
<td>7.2%**</td>
<td>1.8%**</td>
</tr>
<tr>
<td>Performance matched on ROA(t-1)</td>
<td>7.8%**</td>
<td>4.0%</td>
</tr>
<tr>
<td>Performance matched on ROA(t)</td>
<td>6.9%*</td>
<td>5.5%</td>
</tr>
<tr>
<td><strong>Jones</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Test Level: One-tailed 5%</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b = 0</td>
<td>48.7%**</td>
<td>0.0%**</td>
</tr>
<tr>
<td>b – c = 0</td>
<td>4.6%</td>
<td>0.0%**</td>
</tr>
<tr>
<td>b – c – d = 0</td>
<td>1.2%**</td>
<td>1.4%**</td>
</tr>
<tr>
<td>Performance matched on ROA(t-1)</td>
<td>7.1%**</td>
<td>5.2%</td>
</tr>
<tr>
<td>Performance matched on ROA(t)</td>
<td>5.5%</td>
<td>5.4%</td>
</tr>
<tr>
<td><strong>Modified Jones</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Test Level: One-tailed 5%</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b = 0</td>
<td>69.9%**</td>
<td>0.0%**</td>
</tr>
<tr>
<td>b – c = 0</td>
<td>12.1%**</td>
<td>0.2%**</td>
</tr>
<tr>
<td>b – c – d = 0</td>
<td>4.1%</td>
<td>1.8%**</td>
</tr>
<tr>
<td>Performance matched on ROA(t-1)</td>
<td>7.3%**</td>
<td>5.0%</td>
</tr>
<tr>
<td>Performance matched on ROA(t)</td>
<td>5.8%</td>
<td>5.7%</td>
</tr>
<tr>
<td><strong>DD</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Test Level: One-tailed 5%</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b = 0</td>
<td>100.0%**</td>
<td>0.0%**</td>
</tr>
<tr>
<td>b – c = 0</td>
<td>62.0%**</td>
<td>0.0%**</td>
</tr>
<tr>
<td>b – c – d = 0</td>
<td>20.5%**</td>
<td>0.4%**</td>
</tr>
<tr>
<td>Performance matched on ROA(t-1)</td>
<td>9.2%**</td>
<td>3.3%*</td>
</tr>
<tr>
<td>Performance matched on ROA(t)</td>
<td>7.9%**</td>
<td>3.6%*</td>
</tr>
<tr>
<td><strong>McNichols</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Test Level: One-tailed 5%</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b = 0</td>
<td>89.3%**</td>
<td>0.0%**</td>
</tr>
<tr>
<td>b – c = 0</td>
<td>20.7%**</td>
<td>0.0%**</td>
</tr>
<tr>
<td>b – c – d = 0</td>
<td>6.8%*</td>
<td>0.4%**</td>
</tr>
<tr>
<td>Performance matched on ROA(t-1)</td>
<td>7.2%**</td>
<td>4.3%</td>
</tr>
<tr>
<td>Performance matched on ROA(t)</td>
<td>5.5%</td>
<td>4.2%</td>
</tr>
</tbody>
</table>

** *, * Significantly different from the specified test level at 1 and 5 percent level respectively using a two-tailed binomial test

This table reports rejection rates for the null hypothesis of no earnings management for observations in the highest and lowest deciles of earnings performance (i.e., earnings before extraordinary items (ib)) divided by beginning total assets (at)). Tests of earnings management are conducted using the following models:

**Healy**: WC\_ACC\_i,t = a + b\_PART\_i,t + c\_PARTP\_i,t + d\_PARTP2\_i,t + e\_i,t

where

WC\_ACC\_i,t = (ΔCA\_i,t – ΔCL\_i,t – ΔCash\_i,t + ΔSTD\_i,t)/A\_i,t: Working capital accruals

ΔCA: Change in current assets (act)

ΔCL: Change in current liabilities (lct)

ΔCash: Change in cash (che)

ΔSTD: Change in short-term debt (dlc)

A: Total assets (at)

PART\_i,t: Partitioning variable that is set to 1 for randomly selected 100 firm years and 0 otherwise

PARTP\_i,t: Partitioning variable that is set to 1 in the year following randomly selected 100 firm years and 0 otherwise

PARTP2\_i,t: Partitioning variable that is set to 1 in the second year following randomly selected 100 firm years and 0 otherwise

* Characters in brackets are Compustat mnemonics and the same definitions apply to the following models.

**Jones**: WC\_ACC\_i,t = a + b\_PART\_i,t + c\_PARTP\_i,t + d\_PARTP2\_i,t + f\_REV\_i,t + f\_PE\_i,t + e\_i,t

where
$ \Delta \text{REV}_{it} = \frac{(\text{Revenue}_{it} - \text{Revenue}_{it-1})}{A_{it}}$: Change in revenue (sale)

$ \text{PPE}_{it} = \text{PP&E}_{it}/A_{it}$: Gross property, plant and equipment (ppeg)

**Modified Jones:** $\text{WC\_ACC}_{it} = a + b \text{PART}_{it} + c \text{PARTP1}_{it} + d \text{PARTP2}_{it} + f_1 \Delta \text{REV}_{it} + f_2 \text{PPE}_{it} + \epsilon_{it}$

where

$ \Delta \text{REV}_{it} = (\text{Revenue}_{it} - \text{Revenue}_{it-1})/A_{it}$: Change in revenue (sale)

$ \text{PPE}_{it} = \text{PP&E}_{it}/A_{it}$: Gross property, plant and equipment (ppeg)

$ \text{DD}: \text{WC\_ACC}_{it} = a + b \text{PART}_{it} + c \text{PARTP1}_{it} + d \text{PARTP2}_{it} + f_1 \text{CF}_{it} + f_2 \text{CF}_{it-1} + f_3 \text{CF}_{it+1} + \epsilon_{it}$

where

$ \text{CF}_{it} = \text{Earnings before Extraordinary Items}_{it}$ (ib) – $\text{WC\_ACC}_{it}$

**McNichols:** $\text{WC\_ACC}_{it} = a + b \text{PART}_{it} + c \text{PARTP1}_{it} + d \text{PARTP2}_{it} + f_1 \Delta \text{REV}_{it} + f_2 \text{PPE}_{it} + f_3 \text{CF}_{it} + f_4 \text{CF}_{it-1} + f_5 \text{CF}_{it+1} + \epsilon_{it}$

where

$ \Delta \text{REV}_{it} = (\text{Revenue}_{it} - \text{Revenue}_{it-1})/A_{it}$: Change in revenue (sale)

$ \text{PPE}_{it} = \text{PP&E}_{it}/A_{it}$: Gross property, plant and equipment (ppeg)

$ \text{DD}: \text{Discretionary accruals are residuals from WC\_ACC}_{it} = a + f_1 \Delta \text{REV}_{it} + f_2 \text{PPE}_{it} + \epsilon_{it}$

where

$ \Delta \text{REV}_{it} = (\text{Revenue}_{it} - \text{Revenue}_{it-1})/A_{it}$: Change in revenue (sale)

$ \text{PPE}_{it} = \text{PP&E}_{it}/A_{it}$: Gross property, plant and equipment (ppeg)

**Healy:** Discretionary accruals = $\text{WC\_ACC}_{it}$

**Jones:** Discretionary accruals are residuals from $\text{WC\_ACC}_{it} = a + f_1 \Delta \text{REV}_{it} + f_2 \text{PPE}_{it} + \epsilon_{it}$

where

$ \Delta \text{REV}_{it} = (\text{Revenue}_{it} - \text{Revenue}_{it-1})/A_{it}$: Change in revenue (sale)

$ \text{PPE}_{it} = \text{PP&E}_{it}/A_{it}$: Gross property, plant and equipment (ppeg)

**Modified Jones:** Discretionary accruals are residuals from $\text{WC\_ACC}_{it} = a + b \text{PART}_{it} + c \text{PARTP1}_{it} + d \text{PARTP2}_{it} + f_1 \Delta \text{REV}_{it} + f_2 \text{PPE}_{it} + f_3 \text{CF}_{it} + f_4 \text{CF}_{it-1} + f_5 \text{CF}_{it+1} + \epsilon_{it}$

where

$ \Delta \text{REV}_{it} = (\text{Revenue}_{it} - \text{Revenue}_{it-1})/A_{it}$: Change in revenue (sale)

$ \text{PPE}_{it} = \text{PP&E}_{it}/A_{it}$: Gross property, plant and equipment (ppeg)

$ \text{DD}: \text{Discretionary accruals are residuals from WC\_ACC}_{it} = a + f_1 \text{CF}_{it} + f_2 \text{CF}_{it-1} + f_3 \text{CF}_{it+1} + \epsilon_{it}$

where

$ \text{CF}_{it} = \text{Earnings before Extraordinary Items}_{it}$ (ib) – $\text{WC\_ACC}_{it}$

**McNichols:** Discretionary accruals are residuals from $\text{WC\_ACC}_{it} = a + f_1 \Delta \text{REV}_{it} + f_2 \text{PPE}_{it} + f_3 \text{CF}_{it} + f_4 \text{CF}_{it-1} + f_5 \text{CF}_{it+1} + \epsilon_{it}$

where

$ \Delta \text{REV}_{it} = (\text{Revenue}_{it} - \text{Revenue}_{it-1})/A_{it}$: Change in revenue (sale)

$ \text{PPE}_{it} = \text{PP&E}_{it}/A_{it}$: Gross property, plant and equipment (ppeg)

$ \text{CF}_{it} = \text{Earnings before Extraordinary Items}_{it}$ (ib) – $\text{WC\_ACC}_{it}$

All 209,530 firm year observations between 1950 and 2009 are first grouped into decile portfolios based on earnings performance. Rejection rates for the null hypothesis of no earnings management are reported based on PART only (i.e., b=0), PART and PARTP1 (i.e., b=0, b=-d=0), PART, PARTP1 and PARTP2 (i.e., b=c=-d=0). To this end, 100 firm years are randomly selected from the highest earnings performance decile and designated as earnings management years, and PART is set to 1 for those 100 firm years (zero otherwise). Additionally, PARTP1 is set to 1 for the first year following those 100 earnings management years (zero otherwise) and PARTP2 is set to 1 for the second year following those 100 earnings management years (zero otherwise). Then, a pooled cross-sectional regression is run using all 209,530 observations and $\chi^2$-tests are conducted based on heteroscedasticity-consistent standard errors. This procedure is repeated 1000 times and the percent of the 1000 regressions is determined where the $\chi^2$-statistic is significant at the five percent level (one-tailed test). Then, a binomial test is performed to determine if the percentage is significantly different from the specified test level (i.e., 5%). Note that b is the coefficient on PART, c is the coefficient on PARTP1 and d is coefficient on PARTP2. The same procedure is performed for the lowest earnings performance decile. For performance matching models, performance matched discretionary accruals are computed for those 100 firm years selected for the reversal models and t-tests are conducted to examine whether the mean performance matched discretionary accruals of 100 randomly selected observations are significant at five percent level (one-tailed test). For performance matching on ROA(t), each of 100 randomly selected observations is matched with another observation from the same two-digit SIC code with the closest return on assets in the current year. Performance matched discretionary accruals for each of 100 randomly selected observations are then computed by subtracting the matched firm’s discretionary accruals. Performance matching on ROA(t-1) is done in a similar manner to performance matching on ROA(t), based on the closest prior year’s return on assets. This procedure is repeated 1000 times and the percent of the 1000 repetitions is determined where the t-statistic is significant at the five percent level (one-tailed test). Then, a binomial test is performed to determine if the percentage is significantly different from the specified test level (i.e., 5%). The same procedure is performed for the lowest earnings performance decile.
Table 6. Rejection rates when earnings management years are selected from the highest and lowest deciles of various economic characteristics

**Panel A. Significantly positive**

<table>
<thead>
<tr>
<th></th>
<th>Healy Test Level: One-tailed 5%</th>
<th>Jones Test Level: One-tailed 5%</th>
<th>Modified Jones Test Level: One-tailed 5%</th>
<th>DD Test Level: One-tailed 5%</th>
<th>McNichols Test Level: One-tailed 5%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sales Growth</td>
<td>Size</td>
<td>Cash Flows</td>
<td>Earnings Growth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Highest</td>
<td>Lowest</td>
<td>Highest</td>
<td>Lowest</td>
<td>Highest</td>
</tr>
<tr>
<td>b = 0</td>
<td>88.8%**</td>
<td>0.0%**</td>
<td>0.0%**</td>
<td>0.0%**</td>
<td>0.0%**</td>
</tr>
<tr>
<td>b – c = 0</td>
<td>50.4%**</td>
<td>0.1%**</td>
<td>6.4%**</td>
<td>1.2%**</td>
<td>0.0%**</td>
</tr>
<tr>
<td>b – c – d = 0</td>
<td>37.4%**</td>
<td>0.3%**</td>
<td>41.8%**</td>
<td>1.5%**</td>
<td>0.0%**</td>
</tr>
<tr>
<td>Performance matched on ROA(t-1)</td>
<td>76.2%**</td>
<td>0.1%**</td>
<td>0.2%**</td>
<td>1.6%**</td>
<td>0.0%**</td>
</tr>
<tr>
<td>Performance matched on ROA(t)</td>
<td>78.8%**</td>
<td>0.0%**</td>
<td>0.2%**</td>
<td>1.1%**</td>
<td>0.0%**</td>
</tr>
<tr>
<td>b = 0</td>
<td>5.0%</td>
<td>1.9%**</td>
<td>0.0%**</td>
<td>0.5%**</td>
<td>0.0%**</td>
</tr>
<tr>
<td>b – c = 0</td>
<td>3.7%</td>
<td>8.2%**</td>
<td>3.6%**</td>
<td>2.0%**</td>
<td>0.0%**</td>
</tr>
<tr>
<td>b – c – d = 0</td>
<td>3.4%*</td>
<td>11.2%**</td>
<td>13.4%**</td>
<td>2.4%**</td>
<td>0.0%**</td>
</tr>
<tr>
<td>Performance matched on ROA(t-1)</td>
<td>5.6%</td>
<td>13.0%**</td>
<td>0.4%**</td>
<td>3.4%**</td>
<td>0.0%**</td>
</tr>
<tr>
<td>Performance matched on ROA(t)</td>
<td>6.5%*</td>
<td>13.0%**</td>
<td>0.3%**</td>
<td>2.8%**</td>
<td>0.0%**</td>
</tr>
<tr>
<td>b = 0</td>
<td>14.2%**</td>
<td>1.1%**</td>
<td>0.5%**</td>
<td>0.5%**</td>
<td>0.0%**</td>
</tr>
<tr>
<td>b – c = 0</td>
<td>9.4%**</td>
<td>4.8%</td>
<td>6.6%**</td>
<td>1.9%**</td>
<td>0.0%**</td>
</tr>
<tr>
<td>b – c – d = 0</td>
<td>8.3%**</td>
<td>8.0%**</td>
<td>18.0%**</td>
<td>2.4%**</td>
<td>0.0%**</td>
</tr>
<tr>
<td>Performance matched on ROA(t-1)</td>
<td>18.4%**</td>
<td>4.1%</td>
<td>0.4%**</td>
<td>2.3%**</td>
<td>0.0%**</td>
</tr>
<tr>
<td>Performance matched on ROA(t)</td>
<td>23.3%**</td>
<td>4.2%</td>
<td>0.2%**</td>
<td>1.9%**</td>
<td>0.0%**</td>
</tr>
<tr>
<td>b = 0</td>
<td>89.5%**</td>
<td>0.0%**</td>
<td>0.0%**</td>
<td>0.0%**</td>
<td>0.0%**</td>
</tr>
<tr>
<td>b – c = 0</td>
<td>40.7%**</td>
<td>0.0%**</td>
<td>6.4%**</td>
<td>1.2%**</td>
<td>0.0%**</td>
</tr>
<tr>
<td>b – c – d = 0</td>
<td>31.1%**</td>
<td>0.2%**</td>
<td>55.6%**</td>
<td>1.3%**</td>
<td>0.0%**</td>
</tr>
<tr>
<td>Performance matched on ROA(t-1)</td>
<td>82.8%**</td>
<td>0.1%**</td>
<td>0.0%**</td>
<td>2.3%**</td>
<td>0.0%**</td>
</tr>
<tr>
<td>Performance matched on ROA(t)</td>
<td>84.2%**</td>
<td>0.0%**</td>
<td>0.0%**</td>
<td>2.7%**</td>
<td>0.0%**</td>
</tr>
</tbody>
</table>

**Significantly different from the specified test level at 1 and 5 percent level respectively using a two-tailed binomial test**
Table 6 (continued)
**Panel B. Significantly negative**

<table>
<thead>
<tr>
<th>Test Level: One-tailed 5%</th>
<th>Sales Growth</th>
<th>Earnings Growth</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Healy</strong></td>
<td>Highest</td>
<td>Lowest</td>
</tr>
<tr>
<td>b = 0</td>
<td>0.0%**</td>
<td>98.6%**</td>
</tr>
<tr>
<td>b – c = 0</td>
<td>0.0%**</td>
<td>46.8%**</td>
</tr>
<tr>
<td>b – c – d = 0</td>
<td>0.2%**</td>
<td>17.6%**</td>
</tr>
<tr>
<td>Performance matched on ROA(t-1)</td>
<td>0.0%**</td>
<td>47.1%**</td>
</tr>
<tr>
<td>Performance matched on ROA(t)</td>
<td>0.0%**</td>
<td>41.3%**</td>
</tr>
<tr>
<td><strong>Jones</strong></td>
<td>Highest</td>
<td>Lowest</td>
</tr>
<tr>
<td>b = 0</td>
<td>6.5%</td>
<td>12.9%**</td>
</tr>
<tr>
<td>b – c = 0</td>
<td>7.8%**</td>
<td>2.6%**</td>
</tr>
<tr>
<td>b – c – d = 0</td>
<td>6.2%</td>
<td>1.9%**</td>
</tr>
<tr>
<td>Performance matched on ROA(t-1)</td>
<td>5.9%</td>
<td>1.0%**</td>
</tr>
<tr>
<td>Performance matched on ROA(t)</td>
<td>3.6%*</td>
<td>1.1%**</td>
</tr>
<tr>
<td><strong>Modified Jones</strong></td>
<td>Highest</td>
<td>Lowest</td>
</tr>
<tr>
<td>b = 0</td>
<td>1.9%**</td>
<td>19.8%**</td>
</tr>
<tr>
<td>b – c = 0</td>
<td>2.1%**</td>
<td>5.2%**</td>
</tr>
<tr>
<td>b – c – d = 0</td>
<td>2.1%**</td>
<td>2.5%**</td>
</tr>
<tr>
<td>Performance matched on ROA(t-1)</td>
<td>1.5%**</td>
<td>4.9%**</td>
</tr>
<tr>
<td>Performance matched on ROA(t)</td>
<td>0.5%**</td>
<td>4.9%**</td>
</tr>
<tr>
<td><strong>DD</strong></td>
<td>Highest</td>
<td>Lowest</td>
</tr>
<tr>
<td>b = 0</td>
<td>0.0%**</td>
<td>99.7%**</td>
</tr>
<tr>
<td>b – c = 0</td>
<td>0.0%**</td>
<td>49.4%**</td>
</tr>
<tr>
<td>b – c – d = 0</td>
<td>0.2%**</td>
<td>19.7%**</td>
</tr>
<tr>
<td>Performance matched on ROA(t-1)</td>
<td>0.0%**</td>
<td>45.1%**</td>
</tr>
<tr>
<td>Performance matched on ROA(t)</td>
<td>0.0%**</td>
<td>38.0%**</td>
</tr>
<tr>
<td><strong>McNichols</strong></td>
<td>Highest</td>
<td>Lowest</td>
</tr>
<tr>
<td>b = 0</td>
<td>4.7%</td>
<td>19.3%**</td>
</tr>
<tr>
<td>b – c = 0</td>
<td>9.5%**</td>
<td>3.1%**</td>
</tr>
<tr>
<td>b – c – d = 0</td>
<td>8.5%**</td>
<td>2.4%**</td>
</tr>
<tr>
<td>Performance matched on ROA(t-1)</td>
<td>3.4%**</td>
<td>0.9%**</td>
</tr>
<tr>
<td>Performance matched on ROA(t)</td>
<td>2.7%**</td>
<td>0.5%**</td>
</tr>
</tbody>
</table>

**Significantly different from the specified test level at 1 and 5 percent level respectively using a two-tailed binomial test**

This table reports rejection rates for tests of the null hypothesis of no earnings management for observations in the highest and lowest deciles of various economic characteristics. Sales growth is the current year sales divided by the sales in the prior year. Size is the market value at fiscal year end. Cash flows are earnings before extraordinary (ib) items less working capital accruals (WC_ACC). Earnings growth is the median of analysts’ long-term earnings growth forecast from I/B/E/S for the last month of the fiscal year. Tests of earnings management are conducted using

- 57 -
the following models:

**Healy:** WC\_ACC\_it = a + b\_PART\_it + c\_PARTP1\_it + d\_PARTP2\_it + \epsilon\_it

where:

WC\_ACC\_it = (\Delta CA\_it - \Delta CL\_it - \Delta Cash\_it + \Delta STD\_it)/A\_it; Working capital accruals

\Delta CA: Change in current assets (act)

\Delta CL: Change in current liabilities (lt)

\Delta Cash: Change in cash (che)

\Delta STD: Change in short-term debt (dlc)

A: Total assets (at)

PART\_it: Partitioning variable that is set to 1 for randomly selected 100 firm years and 0 otherwise

PARTP1\_it: Partitioning variable that is set to 1 in the year following randomly selected 100 firm years and 0 otherwise

PARTP2\_it: Partitioning variable that is set to 1 in the second year following randomly selected 100 firm years and 0 otherwise

* Characters in brackets are Compustat mnemonics and the same definitions apply to the following models.

**Jones:** WC\_ACC\_it = a + b\_PART\_it + c\_PARTP1\_it + d\_PARTP2\_it + f\_\Delta REV\_it + f\_2 PPE\_it + \epsilon\_it

where:

\Delta REV\_it = (Revenue\_it − Revenue\_it−1)/A\_it; Change in revenue (sale)

PPE\_it = PP&EA\_it/A\_it; Gross property, plant and equipment (ppegp)

**Modified Jones:** WC\_ACC\_it = a + b\_PART\_it + c\_PARTP1\_it + d\_PARTP2\_it + f\_\Delta REV\_it + f\_2 PPE\_it + f\_3 CF\_it−1 + f\_4 CF\_it−1−1 + \epsilon\_it

where:

\Delta REV\_it = ((Revenue\_it − Revenue\_it−1) − (Net Accounts Receivable\_it − Net Accounts Receivable\_it−1))/A\_it; Change in revenue (sale) - Change in net accounts receivable (rect)

PPE\_it = PP&EA\_it/A\_it; Gross property, plant and equipment (ppegp)

**DD:** WC\_ACC\_it = a + b\_PART\_it + c\_PARTP1\_it + d\_PARTP2\_it + f\_\Delta REV\_it + f\_2 PPE\_it + f\_3 CF\_it−1 + f\_4 CF\_it−1−1 + f\_5 CF\_it−1−2 + \epsilon\_it

where:

CF\_it = Earnings before Extraordinary Items\_it (ib) − WC\_ACC\_it

**McNichols:** WC\_ACC\_it = a + b\_PART\_it + c\_PARTP1\_it + d\_PARTP2\_it + f\_\Delta REV\_it + f\_2 PPE\_it + f\_3 CF\_it−1 + f\_4 CF\_it−1−1 + f\_5 CF\_it−1−2 + \epsilon\_it

where:

\Delta REV\_it = (Revenue\_it − Revenue\_it−1)/A\_it; Change in revenue (sale)

PPE\_it = PP&EA\_it/A\_it; Gross property, plant and equipment (ppegp)

For each performance matching model, discretionary accruals are estimated based on the respective model below:

**Healy:** Discretionary accruals = WC\_ACC\_it

**Jones:** Discretionary accruals are residuals from WC\_ACC\_it = a + f\_\Delta REV\_it + f\_2 PPE\_it + \epsilon\_it

where:

\Delta REV\_it = (Revenue\_it − Revenue\_it−1)/A\_it; Change in revenue (sale)

PPE\_it = PP&EA\_it/A\_it; Gross property, plant and equipment (ppegp)

**Modified Jones:** Discretionary accruals are residuals from WC\_ACC\_it = a + f\_\Delta REV\_it + f\_2 PPE\_it + \epsilon\_it

where:

\Delta REV\_it = ((Revenue\_it − Revenue\_it−1) − (Net Accounts Receivable\_it − Net Accounts Receivable\_it−1))/A\_it; Change in revenue (sale) - Change in net accounts receivable (rect)

PPE\_it = PP&EA\_it/A\_it; Gross property, plant and equipment (ppegp)

**DD:** Discretionary accruals are residuals from WC\_ACC\_it = a + f\_\Delta REV\_it + f\_2 PPE\_it + f\_3 CF\_it−1 + f\_4 CF\_it−1−1 + \epsilon\_it

where:

CF\_it = Earnings before Extraordinary Items\_it (ib) − WC\_ACC\_it

**McNichols:** Discretionary accruals are residuals from WC\_ACC\_it = a + f\_\Delta REV\_it + f\_2 PPE\_it + f\_3 CF\_it−1 + f\_4 CF\_it−1−1 + f\_5 CF\_it−1−2 + \epsilon\_it

where:

\Delta REV\_it = (Revenue\_it − Revenue\_it−1)/A\_it; Change in revenue (sale)

PPE\_it = PP&EA\_it/A\_it; Gross property, plant and equipment (ppegp)

CF\_it = Earnings before Extraordinary Items\_it (ib) − WC\_ACC\_it

All 209,530 firm year observations (53,025 firm year observations for earnings growth) between 1950 and 2009 are first grouped into decile portfolios based on various economic characteristics. Rejection rates for the null hypothesis of no earnings management are reported based on PART only (i.e., b=0), PART and PARTP1 (i.e., b−c=0), PART, PARTP1 and PARTP2 (i.e., b−c−d=0). To this end, 100 firm years are randomly selected from the respective highest decile and designated as earnings management years, and PART is set to 1 for those 100 firm years (zero otherwise). Additionally, PARTP1 is set to 1 for the first year following those 100 earnings management years (zero otherwise) and PARTP2 is set to 1 for the second year following those 100 earnings management years (zero otherwise). Then, a pooled cross-sectional regression is run using all 209,530 observations (53,025 observations for earnings growth) and \chi^2-tests are conducted based on heteroscedasticity-consistent standard errors. This procedure is repeated 1000 times and the percent of the 1000 regressions is determined where the \chi^2-statistic is significant at the five
percent level (one-tailed test). Then, a binomial test is performed to determine if the percentage is significantly different from the specified test level (i.e., 5%). Note that $b$ is the coefficient on $\text{PART}$, $c$ is the coefficient on $\text{PARTP1}$ and $d$ is coefficient on $\text{PARTP2}$. The same procedure is performed for the respective lowest decile. For performance matching models, performance matched discretionary accruals are computed for those 100 firm years selected for the reversal models and $t$-tests are conducted to examine whether the mean performance matched discretionary accruals of 100 randomly selected observations are significant at five percent level (one-tailed test). For performance matching on ROA(t), each of 100 randomly selected observations is matched with another observation from the same two-digit SIC code with the closest return on assets in the current year. Performance matched discretionary accruals for each of 100 randomly selected observations are then computed by subtracting the matched firm’s discretionary accruals. Performance matching on ROA(t-1) is done in a similar manner to performance matching on ROA(t), based on the closest prior year’s return on assets. This procedure is repeated 1000 times and the percent of the 1000 repetitions is determined where the $t$-statistic is significant at the five percent level (one-tailed test). Then, a binomial test is performed to determine if the percentage is significantly different from the specified test level (i.e., 5%). The same procedure is performed for the respective lowest decile.
Figure 1. The power of tests for earnings management using seeded earnings management that is partially reversed in the subsequent period. All tests employ 100 seeded earnings management observations, the Healy model and a 5% one-tailed test.

A: Induced earnings management = 1% of assets

<table>
<thead>
<tr>
<th>Reversal % in Year t+1</th>
<th>Frequency of Rejecting No Earnings Management (1% of Assets)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Freq. b &gt; 0</td>
</tr>
<tr>
<td>0%</td>
<td>20.7%</td>
</tr>
<tr>
<td>10%</td>
<td>20.7%</td>
</tr>
<tr>
<td>20%</td>
<td>20.7%</td>
</tr>
<tr>
<td>30%</td>
<td>20.7%</td>
</tr>
<tr>
<td>40%</td>
<td>20.7%</td>
</tr>
<tr>
<td>50%</td>
<td>20.7%</td>
</tr>
<tr>
<td>60%</td>
<td>20.7%</td>
</tr>
<tr>
<td>70%</td>
<td>20.7%</td>
</tr>
<tr>
<td>80%</td>
<td>20.7%</td>
</tr>
<tr>
<td>90%</td>
<td>20.7%</td>
</tr>
<tr>
<td>100%</td>
<td>20.7%</td>
</tr>
</tbody>
</table>

B: Induced earnings management = 2% of assets

<table>
<thead>
<tr>
<th>Reversal % in Year t+1</th>
<th>Frequency of Rejecting No Earnings Management (2% of Assets)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Freq. b &gt; 0</td>
</tr>
<tr>
<td>0%</td>
<td>46.7%</td>
</tr>
<tr>
<td>10%</td>
<td>46.7%</td>
</tr>
<tr>
<td>20%</td>
<td>46.7%</td>
</tr>
<tr>
<td>30%</td>
<td>46.7%</td>
</tr>
<tr>
<td>40%</td>
<td>46.7%</td>
</tr>
<tr>
<td>50%</td>
<td>46.7%</td>
</tr>
<tr>
<td>60%</td>
<td>46.7%</td>
</tr>
<tr>
<td>70%</td>
<td>46.7%</td>
</tr>
<tr>
<td>80%</td>
<td>46.7%</td>
</tr>
<tr>
<td>90%</td>
<td>46.7%</td>
</tr>
<tr>
<td>100%</td>
<td>46.7%</td>
</tr>
</tbody>
</table>

This figure is based on the Healy model.
\[
WC_{ACC_{i,t}} = a + bPART_{i,t} + cPARTP1_{i,t} + \epsilon_{i,t}
\]

where

\[
WC_{ACC_{i,t}} = (\Delta CA_{i,t} - \Delta CL_{i,t} - \Delta Cash_{i,t} + \Delta STD_{i,t})/A_{i,t-1} + \text{Induced earnings}_{i,t}
\]

for randomly selected 100 firm years

or

\[
WC_{ACC_{i,t}} = (\Delta CA_{i,t} - \Delta CL_{i,t} - \Delta Cash_{i,t} + \Delta STD_{i,t})/A_{i,t-1}
\]

for the rest

\(\Delta CA\): Change in current assets (act)

\(\Delta CL\): Change in current liabilities (lct)

\(\Delta Cash\): Change in cash (che)

\(\Delta STD\): Change in short-term debt (dlc)

\(\text{Induced earnings}_{i,t}\): 1% (figure A) or 2% (figure B) of beginning total assets

\(\text{PART}_{i,t}\): Partitioning variable that is set to 1 for randomly selected 100 firm years and 0 otherwise

\(\text{PARTP1}_{i,t}\): Partitioning variable that is set to 1 in the year following randomly selected 100 firm years and 0 otherwise

The simulations employ 209,530 firm year observations between 1950 and 2009. 100 firm years are randomly selected as earnings management years and 1% (figure A) or 2% (figure B) of beginning total assets are added to existing working capital accruals for the 100 randomly selected firm years. Then, \(\text{PART}\) is set to 1 for those 100 firm years (zero otherwise) and \(\text{PARTP1}\) is set to 1 for the year following those 100 earnings management years (zero otherwise). To examine improvement in detecting earnings management when reversal is incorporated, the reversal of those upward managed earnings in the following year is set to range from 0% up to 100% in increments of 10% as shown on the x-axis. Then, a pooled regression is run using all 209,530 observations for each increment. This procedure is repeated 1,000 times for each increment. The frequency is reported where \(b\) is significant at the five percent level (one tailed test) among 1,000 pooled regressions (i.e., Freq \(b > 0\)) and the frequency is also reported where \(b - c = 0\) is significant at the 10 percent level and that the difference \((b - c)\) is positive among 1,000 pooled regressions (i.e., Freq, \(b-c > 0\)). For the performance matching model, 1% (figure A) or 2% (figure B) of beginning total assets are also added to existing working capital accruals of those 100 randomly selected firm years for the reversal model and each of 100 randomly selected observations is matched with another observation from the same two-digit SIC code with the closest return on assets in the current year (i.e., year \(t\)). Performance matched discretionary accruals for each of 100 randomly selected observations are then computed by subtracting the matched firm’s discretionary accruals and it is determined whether the mean performance matched discretionary accruals of those 100 observations is significantly greater than zero. This procedure is repeated 1000 times and the frequency is reported where the mean performance matched discretionary accruals is greater than zero (i.e., Performance matched discretionary accruals > 0).
Figure 2. The power of tests for earnings management using various sample sizes. All tests employ the Healy model, induced earnings management of 2% that completely reverses in the subsequent period and a 5% one-tailed test.

Figure 2 shows the change in frequency of rejecting the null hypothesis of no earnings management as the sample size increases. The x-axis represents the sample size of earnings management, while the y-axis represents the frequency of rejecting the null hypothesis (% of 1000 regressions). The graph includes three lines: red for Freq b > 0, blue for Freq b-c > 0, and black for Performance Matched Discretionary Accruals > 0.

The table below shows the frequency of rejecting the null hypothesis for different sample sizes:

<table>
<thead>
<tr>
<th>Sample Size of Earnings Management</th>
<th>Freq b &gt; 0</th>
<th>Freq b-c &gt; 0</th>
<th>Performance Matched Discretionary Accruals &gt; 0</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>96.7%</td>
<td>67.0%</td>
<td>35.6%</td>
</tr>
<tr>
<td>200</td>
<td>67.8%</td>
<td>59.8%</td>
<td>50.6%</td>
</tr>
<tr>
<td>300</td>
<td>79.8%</td>
<td>65.9%</td>
<td>61.2%</td>
</tr>
<tr>
<td>400</td>
<td>87.1%</td>
<td>93.8%</td>
<td>73.2%</td>
</tr>
<tr>
<td>500</td>
<td>93.6%</td>
<td>95.9%</td>
<td>75.2%</td>
</tr>
<tr>
<td>600</td>
<td>96.1%</td>
<td>100.0%</td>
<td>86.2%</td>
</tr>
<tr>
<td>700</td>
<td>97.8%</td>
<td>100.0%</td>
<td>69.7%</td>
</tr>
<tr>
<td>800</td>
<td>99.2%</td>
<td>100.0%</td>
<td>94.1%</td>
</tr>
<tr>
<td>900</td>
<td>99.4%</td>
<td>100.0%</td>
<td>93.7%</td>
</tr>
<tr>
<td>1000</td>
<td>99.6%</td>
<td>100.0%</td>
<td>99.4%</td>
</tr>
</tbody>
</table>

This figure is based on the Healy model:

\[ WC_{ACC,i,t} = a + bPART_{i,t} + cPARTP1_{i,t} + e_{i,t} \]

where

- \( WC_{ACC,i,t} \) = \( \Delta \text{CA}_{i,t} - \Delta \text{CL}_{i,t} - \Delta \text{Cash}_{i,t} + \Delta \text{STD}_{i,t}/A_{i,t-1} \) + Induced earnings, for randomly selected 100 firm years
- or \( \Delta \text{CA}_{i,t} - \Delta \text{CL}_{i,t} - \Delta \text{Cash}_{i,t} + \Delta \text{STD}_{i,t}/A_{i,t-1} \) for the rest
- \( \Delta \text{CA} \): Change in current assets (act)
- \( \Delta \text{CL} \): Change in current liabilities (lct)
- \( \Delta \text{Cash} \): Change in cash (che)
- \( \Delta \text{STD} \): Change in short-term debt (dlc)
- Induced earnings,\(_i,t\): 1% (figure A) or 2% (figure B) of beginning total assets
- PART\(_{i,t}\): Partitioning variable that is set to 1 for randomly selected 100 firm years and 0 otherwise
- PARTP1\(_{i,t}\): Partitioning variable that is set to 1 in the year following randomly selected 100 firm years and 0 otherwise

The simulations employ 209,530 firm year observations between 1950 and 2009. 100 firm years are randomly selected as earnings management years and 2% of beginning total assets are added to existing working capital accruals for the 100 randomly selected firm years and those then subtracted from the following year’s working capital accruals, assuming 100% reversal in the following year. PART is set to 1 for those 100 firm years (zero otherwise) and PARTP1 is set to 1 for the year following those 100 earnings management years (zero otherwise). Then, a pooled regression is run using all 209,530 observations. This procedure is repeated 1,000 times. The frequency is reported where b is significant at the five percent level (one tailed test) among 1,000 pooled regressions (i.e., Freq b > 0) and the frequency is also reported where b - c = 0 is significant at the 10 percent level and that the difference (b - c) is positive among 1,000 pooled regressions (i.e., Freq. b-c > 0). The same procedure is repeated for 200 ~ 1,000 randomly selected firm years in increments of 100 as shown on the x-axis. For the performance matching model, 2% of beginning total assets are also added to existing working capital accruals of those 100 randomly selected firm years for the reversal model and each of 100 randomly selected observations is matched with another observation from the same two-digit SIC code with the closest return on assets in the current year (i.e., year t). Performance matched discretionary accruals for each of 100 randomly selected observations are then computed by subtracting the matched firm’s discretionary accruals and it is determined whether the mean performance matched discretionary accruals of those 100 observations is significantly greater than zero. This procedure is repeated 1000 times and the frequency is reported where the mean performance matched discretionary accruals is greater than zero (i.e., Performance matched discretionary accruals > 0). The same procedure is repeated for 200 ~ 1,000 randomly selected firm years in increments of 100 as shown on the x-axis.
Figure 3. Time series of working capital accruals for the Accounting and Auditing Enforcement Release (AAER) sample compared to a sample where PART is randomly assigned. Year 0 includes all years that the SEC alleges firms engage in manipulation.

* Variable definitions (Compustat mnemonics in brackets)
WC\textsubscript{ACC}\textsubscript{i,t} = (\Delta CA\textsubscript{i,t} – \Delta CL\textsubscript{i,t} + \Delta Cash\textsubscript{i,t} + \Delta STD\textsubscript{i,t})/A\textsubscript{i,t} - 1: Working capital accruals

\Delta CA: Change in current assets (act)
\Delta CL: Change in current liabilities (lct)
\Delta Cash: Change in cash (che)
\Delta STD: Change in short-term debt (dlc)
A: Total assets (at)

This figure compares the time series trend of the mean working capital accruals (WC\textsubscript{ACC}) for the AAER sample to that for the sample where earnings management years are randomly selected. For the AAER sample, year 0 represents the 406 AAER firm years between 1971 and 2003 and the mean working capital accruals are computed for year -5 to year 5. If earnings are managed for several consecutive years, “year 1” is the first year following the final year of earnings management. The rest of relative years are defined in the same manner. For the sample where earnings management years randomly selected, 100 firm years are randomly selected from 209,530 firm year observations between 1950 and 2009 and designated as year 0. This procedure is repeated 1,000 times. Then, the mean working capital accruals are computed for year -5 to year 5.
Figure 4. Time series of mean working capital accruals for the highest and lowest earnings performance deciles.

* Variable definitions (Compustat mnemonics in brackets)

\[ WC_{ACC,i,t} = (\Delta CA_{i,t} - \Delta CL_{i,t} - \Delta Cash_{i,t} + \Delta STD_{i,t})/A_{i,t-1} \]

- \(\Delta CA\): Change in current assets (act)
- \(\Delta CL\): Change in current liabilities (lct)
- \(\Delta Cash\): Change in cash (che)
- \(\Delta STD\): Change in short-term debt (dlc)
- \(A\): Total assets (at)

This figure compares the time series trend of the mean working capital accruals (WC_{ACC}) of the highest earnings performance decile to that of the lowest earnings performance decile. Earnings performance is defined as earnings before extraordinary items (ib) divided beginning total assets (at). 209,530 observations between 1950 and 2009 are first grouped into decile portfolios based on earnings performance. 100 firm years are first randomly selected from the highest earnings performance decile and designated as year 0. This procedure is repeated 1,000 times. Then, the mean working capital accruals are computed for year -5 ~ year 5. The same procedure is performed for the lowest earnings decile. For the time series trend of working capital accruals based on the total observations, all 209,530 firm years are designated as year 0 and the mean working capital accruals are computed for year -5 ~ year 5.
Figure 5. Time series of mean working capital accruals for highest and lowest deciles of various economic characteristics.

Figure 5-1. Sales growth

Figure 5-2. Size (Market value)
Figure 5 (continued)

Figure 5-3. Cash flows

Time Series of Mean Working Capital Accruals for Highest and Lowest Cash Flows Decile

* Variable definitions (Compustat mnemonics in brackets)
WC_ACC\_i,t = (\Delta CA\_i,t – \Delta CL\_i,t – \Delta Cash\_i,t + \Delta STD\_i,t)/A\_i,t-1; Working capital accruals
\Delta CA: Change in current assets (act)
\Delta CL: Change in current liabilities (lct)
\Delta Cash: Change in cash (che)
\Delta STD: Change in short-term debt (dlc)
A: Total assets (at)
This figure compares the time series trend of the mean working capital accruals (WC_ACC) of the highest decile to that of the lowest decile for various economic characteristics. Sales growth is the current year sales divided by the sales in the prior year. Size is the market value at fiscal year end. Cash flows are earnings before extraordinary (ib) items less working capital accruals (WC_ACC). Earnings growth is the median of analysts’ long-term earnings growth forecast from I/B/E/S for the last month of the fiscal year. For this figure, 209,530 observations (53,025 firm year observations for earnings growth) between 1950 and 2009 are first grouped into decile portfolios based on each economic characteristic. 100 firm years are first randomly selected from the respective highest decile and designated as year 0. This procedure is repeated 1,000 times. Then, the mean working capital accruals are computed for year -5 – year 5 for each economic characteristic. The same procedure is performed for the respective lowest decile. For the time series trend of working capital accruals based on the total observations, all 209,530 firm years (53,025 firm years for earnings growth) are designated as year 0 and the mean working capital accruals are computed for year -5 – year 5.