Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
Journal of Accounting and Economics 39 (2005) 163–197
JOURNAL OF Accounting & Economics
www.elsevier.com/locate/econbase

Performance matched discretionary accrual measures ☆

S.P. Kothari[a,*], Andrew J. Leone[b], Charles E. Wasley[b]

[a]*Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA 02142, USA*
[b]*William E. Simon Graduate School of Business Administration, University of Rochester, Rochester, NY 14627, USA*

Received 18 April 2001; received in revised form 22 September 2004; accepted 17 November 2004
Available online 23 January 2005

Abstract

We examine the specification and power of tests based on performance-matched discretionary accruals, and make comparisons with tests using traditional discretionary accrual measures (e.g., Jones and modified-Jones models). Performance matching on return on assets controls for the effect of performance on measured discretionary accruals. The results suggest that performance-matched discretionary accrual measures enhance the reliability of inferences from earnings management research when the hypothesis being tested does not imply that earnings management will vary with performance, or where the control firms are not expected to have engaged in earnings management.

JEL classification: M41; C12; C15; M42

Keywords: Discretionary accruals; Earnings management; Performance matching; Discretionary-accruals models

☆We gratefully acknowledge the comments and suggestions of an anonymous referee, Thomas Lys (editor), Wayne Guay, Prem Jain, Ross Watts, Jerry Zimmerman and workshop participants at Arizona State, UC-Irvine, Case Western, Colorado, Erasmus, Georgetown, MIT, Pennsylvania State, Rochester and Tilburg. S.P. Kothari acknowledges financial support from Arthur Andersen and Andrew Leone and Charles Wasley acknowledge the financial support of the Bradley Policy Research Center at the Simon School and the John M. Olin Foundation.

*Corresponding author. Tel.: +1 617 253 0994; fax: +1 617 253 0603.
E-mail address: kothari@mith.edu (S.P. Kothari).

0165-4101/$ - see front matter
doi:10.1016/j.jacceco.2004.11.002

1. Introduction

Use of discretionary accruals in tests of earnings management and market efficiency is widespread (see, for example, Defond and Jiambalvo, 1994, Rees et al., 1996; Teoh et al., 1998a, b). Earnings management studies "examine whether managers act as if they believe users of financial reporting data can be misled into interpreting reported accounting earnings as equivalent to economic profitability" (Fields et al., 2001, p. 279). Naturally, earnings management research is of interest not only to academics, but also to practitioners and regulators.

Inferences drawn from tests of hypotheses related to incentives for earnings management hinge critically on the researcher's ability to accurately estimate discretionary accruals. That is, all tests are joint tests of the researcher's model of discretionary accruals and earnings management.[1] This has spurred interest in research on the modeling of discretionary accruals and the empirical specification of the models. However, accurate estimation of discretionary accruals does not appear to be accomplished using existing models. Fields et al. (2001, p. 289) point out that "The only convincing conclusion appears to be that relying on existing accruals models to solve the problem of multiple method choices may result in serious inference problems," where multiple method choices refers to earnings management using accruals.

Our objective in this paper is to test whether a performance-matched discretionary-accrual approach (a type of control sample approach) is both well specified and powerful at estimating discretionary accruals. Use of such an accrual measure, subject to important caveats about type of hypotheses being tested, may enhance the reliability of inferences from earnings management studies with respect to discretionary accruals. We discuss below the kinds of hypothesis tests where matching may be beneficial.

Previous research examines the specification and power of various discretionary-accrual models (see Dechow et al., 1995), but not that of performance-matched accrual models. Dechow et al. (1995, p. 193) conclude that "all models reject the null hypothesis of no earnings management at rates exceeding the specified test levels when applied to samples of firms with extreme financial performance." These results illustrate the importance of a careful consideration of the hypotheses being tested, because firms with extreme performance are also likely to engage in earnings management. Under that hypothesis, discretionary accrual models may, in fact, correctly detect such manipulation (see Guay et al., 1996). Alternatively, the discretionary accrual models might be misspecified when applied to samples of firms with extreme performance in part because performance and estimated discretionary accruals exhibit a mechanical relation (as discussed below). To the extent the concern is model misspecification, and because earnings management research typically examines non-random samples (e.g., samples that firms self-select into by, for example, changing auditors), earnings management studies must employ some

[1] In the context of testing market's efficiency with respect to earnings management, the tests are joint tests of discretionary accrual models and market efficiency.

means of mitigating the misspecification to reduce the likelihood of incorrect inferences. In this vein, use of a control sample to address specification issues is common in the literature. By relying on a control sample to calibrate earnings management, the earnings management identified by our approach must be interpreted as 'abnormal' earnings management. In other words, adjusting for performance, firms identified as having managed earnings are in fact managing earnings at a rate higher than the comparison sample.

We examine properties of discretionary accruals adjusted for a performance-matched firm's discretionary accrual, where matching is on the basis of a firm's return on assets and industry membership. Our motivation to use ROA as the matching variable as opposed to other candidates (e.g., size, earnings growth, earnings yield, market-to-book, etc.) is two-fold. First, the Dechow et al. (1998) model of accruals discussed in Section 2 suggests ROA controls for the effect of performance on measured discretionary accruals. Second, matching on ROA follows Barber and Lyon's (1996) approach to detecting abnormal operating performance (Barber and Lyon do not focus on accruals) using a matched-firm research design. They find that matching on an operating performance measure similar to the ROA tends to be better than matching on other variables.

Performance matching cannot and does not solve all the problems arising from bad discretionary accrual models or from a researcher's failure to recognize the accrual management incentives that are unique to the research question being addressed. Rather, our approach provides additional controls for what is considered 'normal' earnings management. In other words, firms classified as having abnormally high or low levels of earnings management are those that manage more than would be expected given their level of performance. Researchers should consider using either the fitted values of our model (normal level of earnings management) or the residuals from the model (abnormal level of earnings management), depending on the specific hypotheses being tested (see Section 2.3 for further elaboration). Notwithstanding this caveat, the importance of controlling for the effect of performance in tests of earnings management is not surprising and has been recognized in some prior studies (e.g., Teoh et al., 1998a, b). We contribute to this literature as the first study to thoroughly examine and document the specification and power of performance-based discretionary accrual measures across a wide variety of settings representative of those encountered in accounting research.

Conceptually, our motivation for controlling for performance stems from the simple model of earnings, cash flows, and accruals in Dechow et al. (1998). This model shows that working capital accruals increase in forecasted sales growth and earnings because of a firm's investment in working capital to support the growth in sales. Therefore, if performance exhibits momentum or mean reversion (i.e., performance deviates from a random walk), then expected accruals would be non-zero. Firms with high growth opportunities often exhibit persistent growth patterns (i.e., earnings momentum). Similarly, accounting conservatism can produce earnings persistence (i.e., momentum) in the presence of good news and mean reversion in the presence of bad news (Basu, 1997). There is also evidence of mean reversion conditional on extreme earnings performance (see Brooks and Buckmaster, 1976).

As a result, accruals of firms that have experienced unusual performance are expected to be systematically non-zero. A correlation between performance and accruals is problematic in tests of earnings management because commonly used discretionary accrual models (e.g., the Jones (1991) and modified-Jones models) are mis-specified when applied to samples experiencing extreme performance (see Dechow et al., 1995).[2]

While we control for the impact of performance on estimated discretionary accruals using a performance-matched firm's discretionary accrual, an alternative is to formally model accruals as a function of performance (see Fields et al. (2001) for a discussion of this issue). However, doing so requires imposing a specific functional form linking accruals to past performance in the cross-section. Because of the lack of a theory, we control for performance using a performance-matched firm's discretionary accrual. Using a performance-matched firm's discretionary accrual does not impose a particular functional form linking accruals to performance in a cross-section of firms. Instead, the assumption underlying performance matching is, at the portfolio level, the impact of performance on accruals is identical for the test and matched control samples. For comparative purposes we also conduct tests that control for performance on discretionary accruals using a linear regression (i.e., ROA is added to the Jones and modified-Jones models as an additional regressor). The comparison reveals that tests of discretionary accruals using a performance-matched approach are better specified than those using a linear regression-based approach. This result is due in part to the non-linear relation between accruals and performance.

While adjustment of discretionary accruals for those of performance-matched samples is common in the literature, researchers choose from a wide range of firm characteristics on which to match without systematic evidence to guide their choice. Lack of such guidance hinders inter-study comparability of results. For example, Defond and Subramanyam (1998) match on cash flows, Teoh et al., (1998a) match on industry and net income, while Defond and Jiambalvo (1994) match on year and industry. Alternatively, Perry and Williams (1994) match on industry and size. A slightly different approach is adopted in Holthausen and Larcker (1996) who define a ''control firm'' as the median performance of the subset of firms in the same industry and Kasznik (1999) who uses the median performance of firms matched on return on assets. We provide a systematic treatment of the specification and power of the test using performance-based discretionary accruals. This analysis should aid in the design of future earnings management and market efficiency studies.

Summary of results: The main result from our simulation analysis is that discretionary accruals estimated using the Jones or the modified-Jones model, *and* adjusted for a performance-matched firm's discretionary accrual, tend to be the best specified measures of discretionary accruals across a wide variety of simulated event conditions. We report results using performance matching on the basis of industry and return on assets for the current year, ROA_t, and the past year, ROA_{t-1}.

[2]Recent research attempts to develop accrual models as a function of performance (see Kang and Sivaramakrishnan, 1995; Healy, 1996; Dechow et al., 1998; Peasnell et al., 2000; Barth et al., 2001).

Matching based on ROA_t performs better than matching on ROA_{t-1}. We believe matching on ROA_t produces less misspecified tests because the performance-related error in estimating the discretionary accrual of a treatment firms affects the treatment firm's ROA_t, which is matched with a control firm's ROA_t. Thus, the impact of performance-related accrual on the properties of subsequent period's estimated discretionary accrual of the treatment firm is better controlled for when matching is on ROA_t than by matching on a lagged (i.e., stale) determinant, ROA_{t-1}. The ROA performance-matched accrual measures' superior performance compared to other measures of discretionary accruals parallels the result in the context of operating performance measures and long-horizon stock returns (see Barber and Lyon, 1996, 1997; Lyon et al., 1999; Ikenberry et al., 1995).

Performance-matched discretionary accruals exhibit only a modest degree of mis-specification when firms are randomly selected from an extreme quartile of stocks ranked on the basis of firm characteristics such as the book-to-market ratio, firm size, sales growth, and earnings yield (i.e., performance). However, in the same samples, comparative results based on traditional discretionary accrual measures exhibit a far greater degree of mis-specification.

A caveat related to our analysis is that firms in stratified-random samples might be engaging in earnings management for contracting, political or capital market reasons. Thus, the well-specified rejection rate of the performance-matched approach might be an indication of a tendency to under-reject the null hypothesis (see Guay et al., 1996). In this context, our result that performance-matched measures are well specified is applicable only insofar as a researcher desires to calibrate the degree of earnings management (i.e., discretionary accruals) by the treatment sample relative to a matched sample that has not experienced a contracting, political, or capital market-related treatment event (also see Section 2), but is otherwise identical to the treatment sample in all other economic respects. Obviously, the success of the matched-firm approach hinges on the researcher's ability to identify an appropriate control sample. This, in turn, depends on the specific earnings management hypothesis being tested.

Section 2 provides the motivation for using a performance-matched approach to measure discretionary accruals and Section 3 describes the simulation procedure. Section 4 summarizes the results on the specification of the test and Section 5 reports results for the power of the test. Section 6 reports the results of a wide range of sensitivity analyses and Section 7 summarizes and discusses recommendations for future research.

2. Motivation for performance matching

Economic intuition, extant models of accruals, earnings, and cash flows, and empirical evidence all suggest that accruals are correlated with a firm's contemporaneous and past performance (see, for example, Guay et al. 1996; Healy, 1996; Dechow et al., 1998, 1995; Barth et al., 2001). While the Jones and modified-Jones models attempt to control for contemporaneous performance, empirical assessments of these models suggest that estimated discretionary accruals are significantly influenced by a firm's contemporaneous and past performance (e.g., Dechow et al., 1995). In this section we describe the relation between firm performance and accruals. This framework provides the motivation for developing a control for firm performance when estimating discretionary accruals and when comparing discretionary accruals between samples of firms.

2.1. Properties of earnings, cash flows and accruals

To formalize a relation between firm performance and accruals, we begin with a simple version of the model of earnings, cash flows and accruals discussed in Dechow et al. (1998). Ignoring the depreciation accrual and assuming: (i) sales, S_t, follow a random walk, (ii) cash margin of sales is a constant percentage π, (iii) α fraction of sales are on credit and (iv) all expenses are cash, Dechow et al. (1998) show that:

$$\mathrm{CF}_t = \pi S_t - \alpha\varepsilon_t \tag{1}$$

$$A_t = \alpha\varepsilon_t, \quad \text{and} \tag{2}$$

$$X_t = \mathrm{CF}_t + \alpha\varepsilon_t = \pi S_t, \tag{3}$$

where CF_t is cash flow, A_t is accrual, $\varepsilon_t = S_t - S_{t-1}$ is change in sales (or sales shock if earnings follow a random walk), and X is accounting earnings. In this simple setting, expected accruals are zero because sales follow a random walk.

$$E_t(A_{t+1}) = E_t(\alpha\varepsilon_{t+1}) = 0, \tag{4}$$

and the forecast of future cash flows is current earnings. More specifically,

$$E_t(\mathrm{CF}_{t+1}) = E_t(\pi S_{t+1} - \alpha\varepsilon_{t+1}) = \pi S_t = X_t. \tag{5}$$

The above analysis suggests that as long as the assumptions about the parameters and about the random walk property for sales, and therefore earnings, are descriptive, expected accruals are zero.[3] However, as seen from (4), if forecasted sales changes are not zero (i.e., sales depart from a random walk) or when profit margins or other parameters affecting accruals change, then forecasted earnings changes as well as accruals are non-zero. The direction of forecasted sales and earnings changes depend on whether performance is expected to mean revert or to exhibit momentum. Extreme one-time increases or decreases in performance are likely to produce mean reversion, whereas growth stocks might exhibit momentum for a period of time. Mean reversion or momentum in sales and earnings performance is quite likely for firms exhibiting unusual past performance. This predictability in future performance generates predictable future accruals. Unless the discretionary accrual models adequately filter out this *performance-related* predictable component of accruals, there is a danger of spurious indication of discretionary accruals. Previous research

[3] This conclusion also holds for models that capture the complexity of accounts payables and fixed costs (see Dechow et al., 1998). However, the result cannot be demonstrated as cleanly as for the simple model we present.

(e.g., Dechow et al., 1995; Guay et al., 1996) suggests the likelihood of a spurious indication of discretionary accruals is extremely high in samples experiencing unusual past performance (i.e., non-random samples).[4]

2.2. Controlling for the effect of performance on accruals

Theoretically, the need to control for the effect of current or past year's return on assets on estimated discretionary accruals is guided by the modeling of earnings, cash flows and accruals summarized above. In particular, Eq. (4) for the prediction of accruals suggests that when sales changes are predictable, earnings changes will also be predictable and expected accruals will be non-zero.[5] In samples of firms that are not random with respect to prior firm performance, earnings changes are predictable and accruals are also expected to be non-zero.

One means of controlling for the influence of prior firm performance on estimated discretionary accruals is to expand the set of independent variables used in traditional regression models of discretionary accruals (e.g., the Jones model). In this spirit, we augment the Jones and modified-Jones models to include current or past year's return on assets. Our motivation to use return on assets as a performance measure is two-fold. First, by definition, earnings deflated by assets equals return on assets, which in turn measures performance. Second, prior research analyzing long-run abnormal stock return performance and abnormal operating performance finds matching on ROA results in better specified and more powerful tests compared to other matching variables (see, for example, Barber and Lyon, 1996, 1997; Lyon et al., 1999; Ikenberry et al., 1995).

An alternative to the regression-based approach to control for the effect of performance on estimated discretionary accruals is to adjust a firm's estimated discretionary accrual by that of a performance-matched firm. Such an approach would also mitigate the likelihood that the estimated discretionary accruals are systematically non-zero (i.e., lead to invalid inferences about accrual behavior). Specifically, the performance-matched discretionary accrual measure adjusts a firm's estimated discretionary accrual by subtracting the corresponding discretionary accrual of a firm matched on the basis of industry and current or prior year's return on assets.

The relative efficacy of the matched-firm approach versus including a performance variable in the discretionary accrual regression model is an empirical issue. The regression approach imposes stationarity of the relation through time or in the cross-section, and perhaps more importantly, imposes linearity on the relation between the magnitude of performance and accruals. For statistical as well as economic reasons, we expect the mapping of current performance into future performance, or the

[4]In the presence of mean reversion, momentum, and/or other departures from a random walk property of sales, the inclusion of sales change as an explanatory variable in a discretionary accrual regression model is not sufficient to forecast all of the firm's non-discretionary accruals related to sales.

[5]As the simple model suggests, an alternative to return on assets would be to match on past sales growth. However, matching on return on assets serves to incorporate other factors contributing to the firm's accrual generating process, which our simple model does not capture, but which are likely to affect the magnitude of nondiscretionary accruals.

mapping of performance into returns, to be non-linear (e.g., Brooks and Buckmaster, 1976; Beaver et al., 1979; Freeman and Tse, 1992; Basu, 1997; Watts, 2003). Previous research shows that extreme performance is mean reverting, whereas average performance is quite persistent, which implies a non-linear relation between current and future performance across the entire cross-section.

Economic reasons for the non-linearity are rooted in accounting conservatism and incentives for earnings management (see Watts and Zimmerman, 1986; Basu, 1997; Watts, 2003). Accounting conservatism dictates that losses, but not gains, be anticipated. For example, asset write-offs, goodwill impairment, and restructuring charges all entail reporting the capitalized amounts of losses. In contrast, gains from asset revaluations and capitalized amounts of expected benefits from research and development and/or patents are not included in earnings until realized in future periods. Therefore, reported earnings include capitalized amounts of losses, whereas predominantly the gains included in earnings are flow amounts. Capitalized amounts are far less persistent compared to gains, which imparts a non-linearity in the relation between current and future earnings. A similar non-linearity is predicted as a result of management's tendency to take a "big bath" in bad economic times.

Unless a discretionary accrual model, like the Jones or modified-Jones model, is improvised to address non-linearities, we do not expect the regression approach to be effective at controlling for non-zero estimated discretionary accruals in stratified-random samples. We do not entertain non-linear regression approaches to control for the effect of performance on accruals in part because theory to guide the non-linear modeling is currently unavailable. This means experimentation with a range of non-linear specifications might be warranted. Such an exercise is beyond the scope of our study and potentially suffers from over-fitting of the data.

In contrast to the regression approach, the matched-firm approach does not impose any particular functional form on the relation between performance and accruals. It simply assumes that, on average, the treatment and control firms have the same estimated non-event discretionary accruals. Ultimately, the success of the matched-firm approach hinges on the precision with which matching can be done and the homogeneity in the relation between performance and accruals for the matched and the sample firm. As a result, we examine both the linear regression and the matched-firm approach as a means to control for the effect of performance on estimated discretionary accruals.

2.3. Does controlling for performance over-correct for performance-related accruals?

An important question related to our approach is will the use of industry and performance-matched control firms remove, in part, discretionary accruals resulting from treatment firms' earnings management activities? This would make it more difficult to reject the null hypothesis when it is false (i.e., the power of test using performance-based discretionary accruals would be reduced). This concern of potentially "throwing the baby out with the bath water" arises because matched (control) firms in the industry might have similar incentives to manage earnings when compared to the treatment firms.

While, on the surface, such a concern seems reasonable, controlling for performance-related accruals is nevertheless warranted. In an earnings management study, researchers are typically interested in testing whether an event (e.g., a seasoned equity offer) influences reported earnings performance in the pre- and post-event years. If the treatment firms' earnings performance in the post-event period is indistinguishable from that of the control firms, then the conclusion would be that the firms experiencing the event do not manage earnings any more or less than the matched firms that do not experience the event. Of course, it is possible that both treatment and control firms manage earnings, but this is not what the researcher is interested in testing. More precisely, central to the researcher's study is the hypothesis that the *event itself* contributes to earnings management for reasons beyond other known or observable factors like performance. This point can be made more transparent by considering the three components of estimated discretionary accruals: (i) discretionary accruals related to the "treatment" event (e.g., a seasoned equity offer), which is zero for the control firm; (ii) discretionary accruals arising from other incentives (e.g., bonus contract, meeting analysts' forecasts), which influence both treatment and control firms; and (iii) an accrual correlated with performance. The success of the performance-matched approach is predicated on the assumption that estimated discretionary accruals arising from (ii) and (iii) are, on average, the same for the treatment and control firms. This, of course, is the essence of and rationale for the typical matched-firm research design (see, for example, Campbell and Stanley, 1963; Cook and Campbell, 1979). Therefore, when the estimated discretionary accruals of the treatment and control firms are differenced, only the discretionary accrual related to the event of interest remains. To the extent the *non-event* accrual items (ii) and (iii) are systematically different between the treatment and control firms, the performance-matched discretionary accrual approach would not be as effective in isolating the discretionary accrual of interest (i.e., item (i)). The key point here is that the power of test using performance-based discretionary accrual measures is not sacrificed so long as the researcher seeks to estimate the earnings management impact of the treatment event itself (i.e., item (i)).

To summarize, performance matching can and will remove earnings management that is motivated by (poor or superior) performance because both treatment and matched control firms by design experience similar performance. Thus, performance-matched discretionary accruals represent "abnormal" earnings management, not total earnings management. Since it's designed to capture the earnings management effect that is beyond that attributable to performance, the use of performance-matched discretionary accruals is appropriate in controlling for the well-known misspecification of the discretionary-accrual models associated with performance.

3. The simulation procedure

This section describes the simulation procedure used to assess the specification and power of the test using alternative measures of discretionary accruals. We discuss sample selection (Section 3.1), discretionary accrual measures (Section 3.2), performance matching (Section 3.3), and the test statistics (Section 3.4). Section 3.5 presents descriptive statistics and Section 3.6 reports serial correlation properties for all discretionary accrual measures. The descriptive statistics provide preliminary evidence of potential biases inherent to traditional measures of discretionary accruals. Such biases contribute to test statistic misspecification in actual empirical studies.

3.1. Sample selection

We begin with all firm-year observations from the COMPUSTAT Industrial Annual, and Research files from 1962 through 1999. Consistent with prior discretionary accrual research, we exclude firm-year observations that do not have sufficient data to compute total accruals (described in Section 3.2) or the variables needed to estimate the Jones model. We also exclude all firm-year observations where there are fewer than ten observations in any two-digit SIC code in any given year. This is designed to exclude observations for which the regression-model-based discretionary accrual estimates are likely to be imprecise. Collectively, these filters yield a sample of roughly 210,000 observations. Since we match firms on the basis of performance (described below) and analyze stratified sub-samples based on performance (e.g., book-to-market, market capitalization, earnings/price ratio, sales growth and operating cash flow), the sample size is reduced to about 123,000 after excluding observations that cannot be performance matched or that do not have data to calculate the variables used to form the sub-samples.[6]

We report simulation results for 250 samples of 100 firms each. We draw samples without replacement from the full sample or from stratified subsets. The subsets are the lowest and highest quartiles of firms ranked on book-to-market, past sales growth, earnings-to-price, size (market value of equity, referred to as large and small firms) and operating cash flow. To construct the subsets, each year we rank all firm-year observations on the basis of each partitioning characteristic (e.g., book-to-market or size, measured at the beginning of the year). Each year we only retain the upper and lower quartiles of the sample. For each partitioning variable, we then pool observations across all years to form two sub-samples, one based on pooling all data from the annual upper quartiles and another based on pooling all data from the annual lower quartiles.

[6]An issue that arises is how different are the firm-years excluded from our analysis as a result of the performance matching-requirement (roughly 80,000) from the firm-years included in our analysis (roughly 123,000). While the included and excluded firms have significantly different (based on *t*-tests and two sample Wilcoxon tests) E–P ratios, book-to-market ratios, market values of equity, total accruals and operating cash flow to total asset ratios, economically the mean and median differences are quite small. For example, excluded firms have mean (median) E–P ratios of -0.05 (0.06) compared to –0.06 (0.05) for included firms. Corresponding values for excluded (included) firms book-to-market ratios are mean $= 0.81$ and median $= 0.64$ (mean $= 0.86$ and median $= 0.67$), total accruals are mean $= -0.01$ and median $= -0.03$ (mean $= -0.03$ and median $= -0.03$) and market values of equity are mean = \$454.5M and median = \$51.9M (mean = \$570.8M and median = \$50.5M).

3.2. Discretionary accrual measures

Among the various discretionary accrual models, Dechow et al. (1995) report that the Jones and the modified-Jones models (i.e., the modification by Dechow et al.) perform the best. The main difference between the two models is that the modified-Jones model attributes the entire change in receivables to earnings management (see details below). We begin our analysis with the Jones and modified-Jones models. We estimate the performance-matched Jones model discretionary accrual as the difference between the Jones model discretionary accrual and the corresponding discretionary accrual for a performance-matched firm. We similarly estimate the performance-matched modified-Jones model discretionary accrual. To compare the effectiveness of performance matching, versus a regression-based approach, we estimate an additional discretionary accrual measure where we include return on assets (ROA) in the models. For both the regression-based approach and the performance-matched firm approach we present results based on current or last year's ROA as a means to control for firm performance.

To estimate the discretionary accrual models, we define total accruals (TA) as the change in non-cash current assets minus the change in current liabilities excluding the current portion of long-term debt, minus depreciation and amortization, scaled by lagged total assets. With reference to COMPUSTAT, total accruals $= (\Delta\text{Data4} - \Delta Data1 - \Delta\text{Data5} + \Delta\text{Data34} - \text{Data14})/\text{lagged Data6}$. The Jones model discretionary accrual is estimated cross-sectionally each year using all firm-year observations in the same two-digit SIC code.

$$\text{TA}_{it} = \beta_0 + \beta_1(1/\text{ASSETS}_{it-1}) + \beta_2\Delta\text{SALES}_{it} + \beta_3\text{PPE}_{it} + \varepsilon_{it}, \tag{6}$$

where ΔSALES_{it} is change in sales scaled by lagged total assets, ASSETS_{it-1}, and PPE_{it} is net property, plant and equipment scaled by ASSETS_{it-1}. Use of assets as the deflator is intended to mitigate heteroskedasticity in residuals. White (1980) statistics for the annual, cross-sectional, industry models show that deflation reduces, but does not eliminate heteroskedasticity.

While prior research typically does not include a constant in the above model, we include a constant in the estimation for several reasons. First, it provides an additional control for heteroskedasticity not alleviated by using assets as the deflator. Second, it mitigates problems stemming from an omitted size (scale) variable (see Brown et al., 1999). Finally, we find that discretionary accrual measures based on models without a constant term are less symmetric, making power of the test comparisons less clear-cut. Thus, model estimations including a constant term allow us to better address the power of the test issues that are central to our analysis. Where appropriate, we comment on the differences between results based on models including versus excluding a constant.

We use residuals from the annual cross-sectional industry regression model in (6) as the Jones model discretionary accruals. To obtain modified-Jones model discretionary accruals we follow prior studies that estimate the modified-Jones model cross-sectionally and subtract the change in accounts receivable (ΔAR_{it}) from

$\Delta SALES_{it}$ prior to estimating model (6). See DeFond and Park (1997), Subramanyam (1996) and Guidry et al. (1999) as examples.

Our approach to estimate the modified Jones model (i.e., cross-sectionally) differs from that used by Dechow et al. (1995) in a *time-series* setting. They assume that sales are not managed in the estimation period, but that the entire change in accounts receivable in the event year represents earnings management. Therefore, Dechow et al. use the parameters from the Jones model estimated in the pre-event period for each firm in their sample, and apply those to a modified sales change variable defined as $(\Delta SALES_{it} - \Delta AR_{it})$ to estimate discretionary accruals in the event period. This approach is likely to generate a large estimated discretionary accrual whenever a firm experiences extreme growth in the test period compared to the estimation period.[7] To mitigate this problem and because we do not have a "pre-event" period where we can assume that changes in accounts receivable are unmanaged, we estimate the model as if all changes in accounts receivable arise from earnings management. That is, we cross-sectionally estimate the modified-Jones model using sales changes net of the change in accounts receivables [i.e., we use $\Delta SALES_{it} - \Delta AR_{it}$].

As noted above, we also estimate a model that is similar to the Jones and modified-Jones models, except that it is augmented to include ROA_{it} or ROA_{it-1}. This approach is designed to provide a comparison of the effectiveness of performance matching versus including a performance measure in the accruals regression. The model is

$$\begin{aligned}\mathrm{TA}_{it} = \delta_0 + \delta_1(1/\mathrm{ASSETS}_{it-1}) + \delta_2\Delta\mathrm{SALES}_{it} \\ + \delta_3\mathrm{PPE}_{it} + \delta_4\mathrm{ROA}_{it(\text{or } it-1)} + \upsilon_{it}.\end{aligned} \tag{7}$$

3.3. Performance matching

We match each firm-year observation with another from the same two-digit SIC code and year with the closest return on assets in the current year, ROA_{it} (net income divided by total assets).[8] Performance matching is also done on the basis of two-digit SIC code, year and ROA_{it-1}. We discuss the trade-off between these two alternatives when we present descriptive statistics for estimated discretionary accruals (see

[7]As an example of a treatment sample experiencing high growth, consider Teoh et al. (1998b, p. 68) description of their IPO firms: "The mean and median sales growth scaled by assets, an explanatory variable in the Jones (1991) model for accruals, are 54% and 28%. Loughran and Ritter (1995) also report high sales growth for new issuers." Although Teoh et al. (1998a, b) tabulate results using the modified-Jones model, they report that their results are robust to using the Jones model.

[8]In calculating ROA, we use net income rather than net income plus net-of-tax interest expense (the traditional measure used to calculate ROA) to avoid potential problems associated with estimating a tax rate. However, using net income imparts error in our matching procedure if leverage varies substantially within an industry. While we do not believe the error to be severe in the simulations we perform in the study, researchers should consider the trade off between potential errors in estimating the appropriate tax rate with the potential benefits of more precise matching as it relates to their particular setting, when deciding between net income and net income plus net-of-tax interest expense.

Table 1). We define the Jones-model performance-matched discretionary accrual for firm *i* in year *t* as the Jones-model discretionary accrual in year t minus the matched firm's Jones-model discretionary accrual for year *t*. Performance-matched modified-Jones model discretionary accrual is defined analogously.

3.4. Test statistics

For each of the 250 randomly selected samples (per event condition), we assess the significance of the mean discretionary accrual using a *t*-test. The *t*-test is defined as the equal-weighted sample mean discretionary accrual divided by an estimate of its standard error and assumes cross-sectional independence in the estimated discretionary accruals of the sample firms. This assumption seems justified given that we construct samples by selecting firms without regard to time period or industry membership (i.e., our samples are not clustered by industry and/or calendar time). The test statistic is

$$\overline{DA}/(s(DA)/\sqrt{N}) \sim t_{N-1}, \tag{8}$$

where

$$\overline{DA} = \frac{1}{N}\sum_{i=1}^{N} DA_{it}, \tag{9}$$

and

$$s(DA) = \sqrt{\sum_{i=1}^{N}(DA_{it} - \overline{DA})^2/N-1}, \tag{10}$$

where DA_{it} is the discretionary accrual of firm i in year t (based on one of the alternative discretionary accrual models described above), $\overline{DA}$ is the mean discretionary accrual for the sample, $s(DA)$ is the estimated standard deviation of $\overline{DA}$ and N is sample size (i.e., 100).

3.5. Descriptive statistics for discretionary accrual measures under the null hypothesis

Table 1 reports descriptive statistics for total accruals and discretionary accruals based on the Jones and Modified Jones models with and without performance matching. Panel A contains results for the full sample while Panel B contains results for various stratified-random samples (all values in the table are reported as a percent of total assets). From Panel A, the ratio of total accruals to beginning total assets is −3.03%. The negative value is due largely to depreciation. The inter-quartile range of −8.4% to 1.87% of total assets, coupled with a standard deviation of 11.62% of total assets indicates that the distribution of total accruals to total assets is leptokurtic relative to a standard normal distribution. Across the discretionary accrual measures in Panel A, average values are positive (negative) in three (eight) of the 11 total cases. Since these are regression residuals, they are expected to average to

Table 1

Descriptive statistics for various discretionary accrual measures

Panel A reports the mean, standard deviation, lower quartile, median and upper quartile values for the entire sample. Panel B reports means and medians for samples formed on the basis of book-to-market ratio, sales growth, earnings-to-price (EP) ratio, firm size (market value of equity) and operating cash flow. The samples in Panel B are from the lower and upper quartiles of the firms ranked on each partitioning variable at the end of the year t. The performance-matched discretionary accrual measures are constructed by matching each treatment firm with a control firm based on return on assets in period t or $t-1$. Firm-year accrual observations are from the COMPUSTAT Industrial Annual and Research files from 1963 through 1999. We exclude observations if they do not have sufficient data to construct the accrual measures or if the absolute value of total accruals scaled by total assets exceeds one. We eliminate observations where there are fewer than ten observations in a two-digit industry code for a given year and where a performance-matched firm cannot be obtained. The underlying accrual models (Jones and modified-Jones) include a constant term. All discretionary accrual measures are reported as a percent of total assets and all variables are winsorized at the 1st and 99th percentiles. The final sample size is 122,798

Panel A. Descriptive Statistics for Discretionary Accrual Measures:[a]

Description	Mean	Standard Deviation	Lower Quartile	Median	Upper Quartile
Total accruals	−3.03	11.62	−8.40	−3.46	1.87
Jones model	−0.19	9.98	−4.62	0.03	4.39
Modified-Jones model	−0.29	10.33	−4.86	−0.08	4.37
Jones model with ROA_{t-1}	−0.03	9.62	−4.38	0.07	4.37
Jones model with ROA_t	0.00	10.65	−4.79	−0.04	4.51
Modified Jones model with ROA_{t-1}	−0.04	9.94	−4.55	0.00	4.42
Modified Jones model with ROA_t	−0.03	10.98	−5.01	−0.15	4.52
Performance-matched Jones model ROA_{t-1}	0.08	14.38	−6.88	0.04	7.07
Performance-matched Jones model ROA_t	−0.02	15.50	−7.29	0.00	7.28
Performance-matched modified-Jones model ROA_{t-1}	0.09	14.83	−7.03	0.04	7.26
Performance-matched modified-Jones model ROA_t	−0.02	15.93	−7.45	0.00	7.43

Panel B. Means (Medians) of Discretionary Accrual Measures for Stratified-Random Sub-Samples:[a]

	Book/Market		Sales Growth		E/P Ratio		Size		Operating Cash Flow	
Description	High	Low	High	Low	High	Low	Large	Small	High	Low
Total accruals	−3.54	−3.95	1.31	−7.68	−1.33	−8.63	−3.18	−5.15	−0.29	−7.55
	(−3.63)	(−3.9)	(−0.23)	(−6.5)	(−2.33)	(−7.83)	(−3.77)	(−4.7)	(−1.34)	(−7.34)

Jones model	−0.31	−1.16	**0.40**	−1.15	0.28	−3.23	0.20	−1.59	0.30	−2.47
	(−0.09)	(−0.44)	(0.29)	(−0.33)	(0.21)	(-2.2)	(0.22)	(−1.01)	(0.25)	(−1.82)
Modified-Jones model	−0.61	−1.14	1.38	−2.21	0.31	−3.85	0.25	−2.06	0.69	−3.02
	(−0.26)	(−0.26)	(1.14)	(−1.05)	(0.23)	(−2.58)	(0.16)	(−1.14)	(0.51)	(−1.99)
Jones model with ROA_{t-1}	**0.08**	−0.79	0.63	−0.75	0.22	−2.24	−0.08	−0.82	−0.28	−1.52
	(0.19)	(−0.35)	(0.4)	**(−0.15)**	(0.14)	(−1.41)	**(0.02)**	(−0.36)	(−0.36)	(−1.03)
Jones model with ROA_t	−0.12	−0.49	0.83	−0.66	**−0.10**	−1.45	−0.23	−1.02	−0.69	**−0.38**
	(0.01)	(−0.41)	(0.38)	(−0.21)	(−0.08)	(−0.97)	(−0.11)	(−0.75)	(−0.74)	**(−0.39)**
Modified-Jones model with ROA_{t-1}	−0.14	−0.63	1.74	−1.69	0.28	−2.67	**0.00**	−1.12	**0.12**	−1.89
	(0.02)	(−0.38)	(1.16)	(−0.89)	(0.12)	(−1.82)	**(−0.02)**	(−0.62)	**(−0.21)**	(−1.44)
Modified-Jones model with ROA_t	−0.35	−0.31	1.90	−1.65	−0.15	−1.73	−0.20	−1.35	−0.48	−0.57
	(−0.15)	(−0.43)	(1.06)	(−0.98)	(−0.17)	(−1.24)	(−0.19)	(−1.03)	(−0.79)	(−0.65)
Performance-matched Jones model ROA_{t-1}	0.45	−0.69	0.72	−0.45	0.14	−1.58	−0.28	−0.36	−0.80	−0.77
	(0.27)	(−0.47)	(0.44)	(−0.07)	(0.03)	(−1.25)	(−0.11)	(−0.19)	(−0.48)	(−0.74)
Performance-matched Jones model ROA_t	−0.16	−0.18	0.53	**0.06**	−0.30	**−0.17**	−0.87	**−0.14**	−1.20	0.99
	(0.0)	(−0.11)	**(0.21)**	(0.2)	**(−0.02)**	**(−0.15)**	(−0.32)	**(0.0)**	(−0.7)	(0.7)
Performance-matched modified-Jones model ROA_{t-1}	0.30	−0.57	1.81	−1.35	0.20	−1.92	−0.22	−0.60	−0.50	−1.04
	(0.21)	(−0.36)	(1.31)	(−0.81)	(0.06)	(−1.53)	(−0.07)	(−0.37)	(−0.29)	(−0.99)
Performance-matched modified-Jones model ROA_t	−0.27	**−0.07**	1.51	−0.75	−0.40	−0.21	−0.91	−0.30	−1.28	1.04
	(0.0)	**(−0.05)**	(1.04)	(−0.36)	(−0.06)	(−0.17)	(−0.32)	(−0.12)	(−0.74)	(0.67)

[a]Total Accruals (TA_{it}) is defined as the change in non-cash current assets minus the change in current liabilities excluding the current portion of long-term debt minus depreciation and amortization [with reference to COMPUSTAT data items, $TA = (\Delta Data4 - \Delta Data1 - \Delta Data5 + \Delta Data34 - Data14)/\text{lagged } Data6$]. Discretionary accruals from the Jones model are estimated for each industry and year as follows: $TA_{i,t} = \alpha_0 + \alpha_1/ASSETS_{i,t-1} + \alpha_2 \Delta SALES_{i,t} + \alpha_3 PPE_{i,t} + \varepsilon_{it}$, where $\Delta SALES_{i,t}$ is change in sales scaled by lagged total assets and $PPE_{i,t}$ is net property, plant and equipment scaled by lagged assets. Discretionary accruals from the modified-Jones model are estimated for each industry and year as for the Jones model except that the change in accounts receivable is subtracted from the change in sales. Discretionary accruals from the Jones Model (Modified-Jones model) with ROA are similar to the Jones Model (Modified-Jones model) except for the inclusion of current or lagged year's ROA as an additional explanatory variable. For performance matched discretionary accruals, we match firms on ROA in period t or $t-1$. To obtain a performance-matched Jones model discretionary accrual for firm i we subtract the Jones model discretionary accrual of the firm with the closest ROA that is in the same industry as firm i. A similar approach is used for the modified Jones model.

zero. However, some deviation from zero arises because we winsorize extreme observations by setting the values in the bottom and top one percent to the values of the 1st and 99th percentiles (consistent with prior research).[9]

Results in panel A show that performance matching on current ROA yields discretionary accrual estimates that have both mean ($= -0.02\%$) and median ($= 0\%$) close to zero for both Jones- and modified-Jones models. We also find that adding ROA_t or ROA_{t-1} to the Jones and modified-Jones model yields erratic performance improvement. Specifically, while the average discretionary accruals are close to zero for both models, the medians differ between the Jones or the modified-Jones model, and whether ROA_t or ROA_{t-1} is included in the regression model.

Performance matching increases the standard deviation of the Jones model discretionary accruals from about 10% of total assets to about 14–16% of total assets for the performance-matched Jones model discretionary accrual. The 40–50% increase in variability is approximately the increase one would expect if the estimated discretionary accrual of the sample firm were uncorrelated with that of the matched firm. Assuming independence, the variance of the difference between two random variables with identical variances is twice the variance of the individual random variables. Therefore, the standard deviation would be the square root of two or 1.41 times the standard deviation of the individual random variable.

Consistent with claims in previous research, descriptive statistics in panel B document the inability of discretionary accrual models to generate mean-zero estimates when applied to stratified-random samples. Bold numbers in Panel B correspond to the mean and median value closest to zero in each column of the table. The bias (non-zero values) in the discretionary accrual measures in Panel B is of concern because the greater the bias the more likely it is that the null hypothesis of zero discretionary accruals will be spuriously rejected.

The results reveal that performance matching based on ROA_t and using the Jones model produces the lowest mean and median values (in absolute magnitude). This approach produces the lowest mean value in three of the ten cases and lowest median in five of the ten cases. The next best performing accrual measure is the Modified Jones Model that includes ROA_{t-1} as an additional regressor (lowest mean and median value two times each). In summary, the performance matching approach based on ROA_t and using the Jones model produces means and medians in performance-related sub-samples that are closest to zero more often than the other measures.

A final observation on the results in Panel B is that the mean and median performance-matched discretionary accruals for the operating cash flow sub-sample are substantially different depending on whether matching is on ROA_t or ROA_{t-1}. Matching on ROA_t for the operating cash flow samples mechanically influences the performance-matched discretionary accrual. Holding ROA constant, high operating cash flow stocks must necessarily have low accruals compared to the matched ROA firm. Thus, we expect a negative average for the current year's performance-matched

[9] Whether winsorization imparts any bias that leads to erroneous inferences depends on the research context (see Kothari et al., 2005).

discretionary accrual for high operating cash flow stocks and a positive value for the low operating cash flow stocks. This is precisely what is observed in panel B. This mechanical relation is not obtained when matching is on $\mathrm{ROA}_{t-1.}$

3.6. Serial correlations

Under the *null hypothesis* of no earnings management, the typical discretionary accrual or earnings management study implicitly assumes estimated discretionary accruals to have zero mean and exhibit no serial correlation. For example, a study of earnings management around IPOs would hypothesize that the accruals managed around the IPO reverse in subsequent years. The null hypothesis is that discretionary accruals in the IPO and subsequent years are zero and that the (serial) correlation between the IPO-year discretionary accruals with the subsequent years' discretionary accruals is zero. That is, under the null hypothesis, a zero coefficient in a regression of subsequent years' discretionary accruals on the IPO-year discretionary accruals is predicted. Thus, from a statistical perspective, discretionary-accrual estimates (i.e., error terms from the models) that are serially uncorrelated satisfy one of the distributional properties of the test statistic under the null hypothesis.

In non-random samples, total accruals themselves are likely to be correlated, which can lead to serially correlated estimates of discretionary accruals. The serial correlation in total accruals arises due to economic/operating reasons (e.g., actions by management such as expanding receivables or inventories in periods of growth). A major objective of the discretionary accrual models like the Jones model is to filter out non-discretionary accruals from total accruals to obtain estimates of discretionary accruals that have a zero mean and are serially uncorrelated as expected under the null hypothesis of no earnings management. We expect well-specified discretionary accrual models to be successful in filtering out non-discretionary accruals that are serially correlated.

We report estimates of the serial correlation in various discretionary accrual measures. Serial correlations are slope coefficients from the following cross-sectional regression model estimated annually from $t = 1962$ to 1999:

$$X_{it} = \alpha + \beta X_{it-1} + \varepsilon_{it}, \tag{11}$$

where X_{it} is the current value of the variable of interest (e.g., return on assets, total accruals, Jones- or modified-Jones model discretionary accrual). The serial correlation estimate from the cross-sectional regression in (11) assumes it is identical across the firms in the cross-section (see Fama and French, 2000). While this is unlikely to be true, the regression estimate is unbiased and thus it is an estimate of the cross-sectional average serial correlation. We attempt to mitigate the variation in serial correlation across firms by estimating the model for sub-samples that are, a priori, likely to be homogeneous. A distinct advantage of using (11) compared to a firm-specific time-series regression is that sample attrition and survival bias stemming from requiring a long time series of data for each firm are avoided.

Table 2 reports the average of the annual cross-sectional serial correlation estimates for each variable and sub-sample. Significance tests for a zero mean are based on t-tests where the Fama–MacBeth (1973) standard error used to calculate the t-statistic incorporates the Newey–West (1987) autocorrelation correction for five lags.

The serial correlation estimates reveal that return on assets is positively autocorrelated for the entire sample as well as for all sub-samples. The serial correlation is lower for earnings yield and operating cash flow portfolios, consistent with mean reversion in extreme earnings (e.g., Brooks and Buckmaster, 1976) and extreme cash flows. Serial correlation in total accruals is positive for the all-firm sample, and positive, but much lower (or zero) for high book-to-market, and low sales growth, low earnings yield, low operating cash flow to asset firms. These findings indicate that unusual or extreme past performance imparts a transitory component to accruals.

The results in Table 2 show that discretionary accrual models tend to reduce serial correlation, and that performance matching dampens serial correlation the most. For example, serial correlation in the modified-Jones model discretionary accruals is -0.072 for the low sales growth stocks, which is reduced to -0.041 or -0.047 by performance matching on ROA_{t-1} or ROA_t, respectively. Corresponding numbers for small stocks are -0.091 and -0.040 or -0.046, respectively. In summary, performance matching seems to be better at generating discretionary accrual estimates with properties under the null hypothesis of no earnings management. The extent to which this improvement affects test specification and power is addressed next.

4. Specification of the test: Type I error rates for various discretionary accrual measures

This section reports results on the specification of the test under the null hypothesis of zero discretionary accruals. We report the percentage of times out of 250 simulated samples the null hypotheses of non-negative (Table 3, panel A) and non-positive (Table 3, panel B) discretionary accruals are rejected at the 5% level of significance (upper or lower one-tailed test). These rejection rates measure each metric's Type I error rate. The 95% confidence interval for the rejection rate of 5% ranges from 2% to 8%. If the actual rejection rate falls below (above) 2% (8%), the test is misspecified as it rejects too infrequently (frequently), and is biased in favor of (against) the null hypothesis (results using a 1% significance level lead to similar inferences).

4.1. Rejection rates under the alternative hypothesis of negative discretionary accruals

Panel A of Table 3 reports rejection rates for one-tailed tests of the alternative hypothesis of negative discretionary accruals. To facilitate interpretation of the results, rejection rates that are significantly less than the nominal significance level of the test (i.e., tests that are mis-specified because they are too conservative) appear in bold italic type while rejection rates that are significantly greater than the nominal

Table 2

Serial correlation in ROA, total accruals and various discretionary accrual measures for the entire sample and select subsamples.

The table reports the mean value of the slope coefficient of the following annual regression: $X_{it} = \alpha + \beta X_{it-1} + \varepsilon_{it}$, where $X_{it}(X_{it-1})$ is the value (lagged value) of the particular variable of interest (i.e., ROA, total accruals, Jones model discretionary accruals, Modified-Jones model discretionary accruals, performance-matched Jones model accruals or performance-matched Modified-Jones Model accruals). Results are reported for the full sample (All Firms) and subsamples based on book-to-market, sales growth, earnings-to-price ratio, firm size and operating cash flow. The sub-samples are firm-year observations from the lower and upper quartiles of the firms ranked on each partitioning variable at the end of the year t. The performance-matched discretionary accrual measures are constructed by matching each treatment firm to a control firm based on return on assets in period t or $t-1$. Firm-year accrual observations are constructed from the COMPUSTAT Industrial Annual and Research files from 1963 through 1999. We exclude observations if they do not have sufficient data to construct the accrual measures described below or if the absolute value of total accruals scaled by total assets is greater than one. We eliminate observations where there are fewer than ten observations in a two-digit industry code for a given year and where a performance-matched firm cannot be obtained. The underlying accrual models (Jones and Modified Jones) include a constant term. All variables are winsorized at the 1st and 99th percentiles. The final sample size is 122,798

Variable[a]	All Firms	Book/Market		Sales Growth		E/P Ratio		Size		Oper. Cash Flows	
		High	Low	High	Low	High	Low	Large	Small	High	Low
ROA	0.738**	0.549**	0.779**	0.687**	0.664**	0.411**	0.428**	0.763**	0.661**	0.361**	0.402**
Total accruals	0.189**	0.056**	0.256**	0.273**	0.031	0.114**	0.074	0.356**	0.098**	0.292**	0.058
Jones model accruals	**0.001**	−0.053**	0.029	−0.065	−0.077**	−0.057**	**0.019**	0.131**	−0.102*	0.051**	−0.005
Modified-Jones model accruals	0.015*	−0.043**	0.052**	−0.075	−0.072**	−0.045**	0.020	0.137**	−0.091*	0.063**	−**0.004**
Performance-matched Jones ROA_{t-1}	−0.025**	−0.037**	−0.033	−**0.002**	−0.046	−0.057**	0.066	**0.023***	−0.044**	−0.006	0.059
Performance-matched Jones ROA_{t}	−0.006	−0.047*	−**0.001**	0.003	−0.049**	−0.036**	−0.069**	0.080*	−0.048**	0.028	−0.051**
Performance-matched modified-Jones ROA_{t-1}	−0.023**	−**0.036****	−0.033	−0.012	−**0.041**	−0.057**	0.072	0.025*	−**0.040****	−**0.004**	0.066
Performance-matched modified-Jones ROA_{t}	−0.002	−0.046*	0.012	0.007	−0.047**	−**0.030***	−0.063**	0.080*	−0.046**	0.031*	−0.048**

**, *denotes that t-statitics are significant at 0.01 and 0.05, respectively. t-tests are adjusted for autocorrelation using the Newey-West (1987) correction with 5 lags.

[a]Return on Assets (ROA) is net income (COMPUSTAT data item 18) scaled by lagged total assets. Total accruals is defined as the change in non-cash current assets minus the change in current liabilities excluding the current portion of long-term debt minus depreciation and amortization [with reference to COMPUSTAT data items, TA = (ΔData4 − ΔData1 − ΔData5 + ΔData34 − Data14)/lagged Data6]. Discretionary accruals from the Jones model are estimated for each industry and year as follows: $\text{TA}_{i,t} = \alpha_0 + \alpha_1/\text{ASSETS}_{i,t-1} + \alpha_2\Delta\text{SALES}_{i,t} + \alpha_3\text{PPE}_{i,t} + \varepsilon_{it}$, where TA_{it} (Total Accruals) is as defined above, $\Delta\text{SALES}_{i,t}$ is change in sales scaled by lagged total assets ($\text{ASSETS}_{i,t-1}$), and $\text{PPE}_{i,t}$ is net property, plant and equipment scaled by $\text{ASSETS}_{i,t-1}$. Discretionary accruals from the modified-Jones model are estimated for each industry and year as for the Jones model except that the change in accounts receivable is subtracted from the change in sales. Discretionary accruals from the Jones Model (modified-Jones model) with ROA are similar to the Jones model (modified-Jones model) except for the inclusion of current or lagged year's ROA as an additional explanatory variable. For performance matched discretionary accruals, we match firms on ROA in period t or $t-1$. To obtain a performance-matched Jones model discretionary accrual for firm i we subtract the Jones model discretionary accrual of the firm with the closest ROA that is in the same industry as firm i. A similar approach is used for the modified Jones model.

Table 3
A Comparison of the Type I error rates of alternative discretionary accrual measures for the full sample and upper and lower quartiles of sub-samples formed on the basis of book-to-market ratio, sales growth, earnings-to-price (EP) ratio, firm size and operating cash flow measured at the end of year t

The table reports the percentage of 250 samples of 100 firms each where the null hypothesis of zero discretionary accrual is rejected at the 5% level (upper and lower one-tailed tests). The significance of the mean discretionary accrual in each sample is based on a cross-sectional t-test. Performance-matched discretionary accrual measures are constructed by matching each treatment firm with a control firm based on return on assets in period t or $t-1$. Firm-year accrual observations are constructed from the COMPUSTAT Industrial Annual and Research files from 1963 through 1999. We exclude observations if they do not have sufficient data to construct the accrual measures if the absolute value of total accruals scaled by total assets is greater than one. We eliminate observations where there are fewer than ten observations in a two-digit industry code for a given year and where a performance-matched firm cannot be obtained. The underlying accrual models (Jones and Modified Jones) include a constant term. All variables are winsorized at the 1st and 99th percentiles. The final sample size is 122,798

	All Firms	Book-to-Market		Sales Growth		EP Ratio		Size		Operating Cash Flow	
		High	Low	High	Low	High	Low	Large	Small	High	Low
Panel A. H_A: Accruals < 0[a] (Figures in bold (***bold italic***) signify rejection rates that significantly exceed (fall below) the 5% nominal significance level of the test and indicate that such tests are biased against (in favor of) the null hypothesis)											
Rejection rates for the Jones model:											
Cross sectional within-industry	4.0	**8.4**	**12.8**	*1.2*	**18.8**	***1.2***	**68.0**	2.0	**25.6**	2.4	**34.4**
ROA_{t-1} included as a regressor	6.4	4.4	**14.8**	***1.6***	**14.0**	4.4	**42.0**	5.6	**12.4**	**9.2**	**25.2**
Performance matched on ROA_{t-1}	**8.4**	2.4	**11.2**	***0.4***	**8.4**	4.8	**19.2**	**9.6**	6.0	**12.8**	**12.8**
ROA_t included as a regressor	4.4	4.4	**12.4**	2.0	**12.0**	**9.2**	**20.0**	6.8	23.2	**16.0**	6.4
Performance matched on ROA_t	4.4	6.4	8.0	2.8	5.2	**8.4**	3.2	**14.4**	**10.4**	**17.6**	***1.2***
Rejection rates for the modified-Jones model:											
Cross sectional within-industry	4.4	**14.0**	**10.4**	***0.0***	**46.4**	***0.8***	**74.8**	2.0	**32.0**	***0.4***	**40.8**
ROA_{t-1} included as a regressor	7.6	8.0	**12.4**	***0.0***	**36.8**	3.6	**50.8**	4.8	**17.6**	3.2	**30.0**
Performance matched on ROA_{t-1}	7.6	2.8	**8.8**	***0.0***	**19.2**	5.2	**24.8**	**8.8**	**9.6**	7.6	**13.2**
ROA_t included as a regressor	4.8	**8.4**	**8.4**	***0.0***	**38.4**	**10.4**	**24.0**	7.6	**27.6**	**12.8**	**9.2**
Performance matched on ROA_t	4.4	6.4	**8.4**	***1.2***	**14.0**	**9.2**	4.4	**14.4**	**10.4**	**18.0**	***1.2***

Panel B. H_A: Accruals > 0[a] (Figures in bold (***bold italic***) signify rejection rates that are significantly exceed (fall below) the 5% nominal significance level of the test and indicate that such tests are biased against (in favor of) the null hypothesis)

Rejection rates for the Jones model:											
Cross sectional within-industry	6.0	2.4	2.4	**11.2**	***1.2***	**14.0**	***0.0***	7.2	***0.8***	**9.6**	***0.4***
ROA_{t-1} included as a regressor	6.4	6.8	***1.6***	8.0	***1.6***	**9.6**	***0.4***	5.6	***1.2***	4.4	***0.4***
Performance matched on ROA_{t-1}	4.0	**13.6**	2.0	6.0	2.4	6.4	***1.6***	2.4	3.2	***0.8***	**1.6**
ROA_t included as a regressor	5.6	2.0	3.6	**15.2**	***1.6***	4.0	***0.0***	3.6	***1.6***	**0.8**	4.4
Performance matched on ROA_t	5.6	2.8	5.6	**8.8**	4.4	3.2	4.4	***0.4***	7.2	**0.8**	**12.8**
Rejection rates for the modified-Jones model:											
Cross sectional within-industry	5.2	***1.2***	3.2	**32.4**	***0.0***	**15.6**	***0.0***	**9.2**	***0.4***	**21.2**	***0.0***
ROA_{t-1} included as a regressor	4.8	2.8	2.0	**28.0**	***0.4***	**8.4**	***0.0***	**8.8**	***0.4***	8.0	***0.0***
Performance matched on ROA_{t-1}	5.2	**9.6**	2.0	**16.0**	***0.0***	8.0	***1.2***	2.8	***1.6***	3.2	***0.8***
ROA_t included as a regressor	4.8	***1.6***	4.8	**37.6**	***0.4***	3.6	***0.0***	3.6	***0.4***	2.4	4.0
Performance matched on ROA_t	4.8	2.4	6.8	**20.8**	2.4	3.6	3.6	***0.4***	6.0	***0.8***	**13.6**

[a]Discretionary accruals from the Jones model are estimated for each industry and year as follows: $TA_{i,t} = \alpha_0 + \alpha_1/ASSETS_{i,t-1} + \alpha_2\Delta SALES_{i,t} + \alpha_3 PPE_{i,t} + \varepsilon_{it}$, where TA_{it} (Total Accruals) is defined as the change in non-cash current assets minus the change in current liabilities excluding the current portion of long-term debt minus depreciation and amortization [with reference to COMPUSTAT data items, $TA = (\Delta Data4 - \Delta Data1 - \Delta Data5 + \Delta Data34 - Data14)/\text{lagged } Data6$], $\Delta SALES_{i,t}$ is change in sales scaled by lagged total assets ($ASSETS_{i,t-1}$), and $PPE_{i,t}$ is net property, plant and equipment scaled by $ASSETS_{i,t-1}$. Discretionary accruals from the Jones Model (Modified-Jones Model) with ROA are similar to the Jones Model (Modified-Jones Model) except for the inclusion of current or lagged year's ROA as an additional explanatory variable. For performance matched discretionary accruals, we match firms on ROA in period t or $t-1$. To obtain a performance-matched Jones model discretionary accrual for firm i we subtract the Jones model discretionary accrual of the firm with the closest ROA that is in the same industry as firm i. A similar approach is used for the modified Jones model.

significance level of the test (i.e., tests that are mis-specified because they reject the null hypothesis too often) appear in bold type.

First, we observe that all discretionary accrual measures exhibit some degree of misspecification. No single measure is well-specified under the null hypothesis in each and every one of the 11 sample partitions (columns). Second, while under-rejection of the null hypothesis occurs, misspecification of the test is due primarily to rejecting the null too often. Finally, as we discuss more fully below, the best specified test across the sample partitions contained in panel A is the performance-matched Jones model discretionary accrual where matching is done on ROA_t.

A closer look at the rejection rates in panel A reveals that within a given sample partition and when compared to the other accrual measures within that partition, the Jones and Modified Jones models exhibit the highest Type I error rates. For example, in the low E/P portfolio the rejection rate for the Jones (modified-Jones) model is 68.0% (74.8%). The corresponding rejection rate for the low operating cash flow to total assets partition is 34.4% (40.8%). High rejection rates using the Jones and modified-Jones models are not surprising as Dechow et al. (1995) report similar evidence for samples selected from extreme deciles of stocks ranked according to earnings and cash flow performance. By extending their results we find that even if firms are sampled from less extreme populations (i.e., quartiles in our study) and based on a variety of economic characteristics, the Jones and modified-Jones models excessively reject the null hypothesis of no discretionary accruals.

Misspecification problems are attenuated, but not eliminated when ROA_t or ROA_{t-1} is included in the Jones- and modified-Jones regressions, as the rejection rates remain in excess of 8%. For example, with ROA_t (ROA_{t-1}) added to the Jones model, rejection rates are still excessive as six (six) of the 11 rejection rates exceed 8%. For the modified Jones model, the corresponding rejection rates still exceed 8% in eight (five) of the 11 cases when ROA_t (ROA_{t-1}), respectively are added to the regression model. To provide some specific comparisons, with ROA_t (ROA_{t-1}) included in the Jones model, the rejection rate in samples of low sales growth is 12.0% (14.0%) compared to 18.8% for the Jones model itself. Corresponding numbers for low EP ratio stocks are 20.0% (42.0%) versus 68.0%; for small firms they are 23.2% (12.4%) versus 25.6% and finally for low operating cash flow firms they are 6.4% (25.2%) versus 34.4%.[10]

4.2. Performance matching

Rejection rates for performance-matched discretionary accrual measures in panel A of Table 3 reveal a lesser degree of misspecification compared to other models. For example, performance-matched Jones model discretionary accruals based on with

[10]Un-tabulated results (available upon request) show discretionary accruals calculated as the Jones or modified-Jones model discretionary accrual minus the industry mean or median Jones or modified-Jones model discretionary accrual do not cure the excessive rejection rates of the models. Previous research (e.g., DeFond and Park, 1997) uses industry-adjustment as a means of mitigating the likelihood of spurious rejection. Our results suggest that such attempts are unlikely to be successful.

ROA_t indicate negative discretionary accruals close to 5% of the time except when sampling is restricted to high E/P stocks, large or small market capitalization stocks or high operating cash flow stocks. The high rejection rate of 17.6% for the high operating cash flow firms is neither surprising nor unexpected. In fact, it is obtained mechanically because earnings are the sum of operating cash flows and accruals. Since we match on earnings performance (i.e., ROA) in the current year, treatment firms selected from the high operating cash flow quartile, by construction, will have lower accruals than matched firms that do not always belong to the high operating cash flow quartile. This mechanical relation is absent when matching is on ROA_{t-1}, and as a result the rejection rate in that case is lower, 12.8%, although still higher than the upper bound of the 95% confidence interval.[11]

While performance matching does not cure all excessive Type I error rates, performance matching based on ROA_t and using the Jones model performs the best across the sample partitions. Moreover, even when this measure is mis-specified, its excessive rejection rates tend to be comparable and sometimes lower than those of the other accrual measures. Nonetheless, while the results for performance-matched Jones model using ROA_t indicate that it tends to be the most reliable overall, it will not solve misspecification problems in samples of very large or small market capitalization stocks.

4.3. Comparing the Jones with the modified-Jones model

The results in Table 3, panel A show that differences between the rejection rates of the Jones and modified-Jones models are generally small except in the case of low sales growth samples. The rejection frequency for the low sales growth quartile firms based on the modified-Jones model is 46.4% compared to 18.8% for the Jones model.[12] A potential explanation for this large difference is that the modified-Jones model assumes that all credit sales represent accrual manipulation. As we note earlier, the credit-sales related assumption causes the modified-Jones model discretionary accrual to be positively correlated with sales growth. Therefore, for samples from low sales growth quartile firms, the performance-matched

[11]Thus, a limitation of matching on current ROA for extreme operating cash flow firms is it induces test misspecification. To illustrate, assume the treatment (T) firm belongs to the high operating cash flow quartile, but both T and the control firm (C) have identical reported ROA because they are matched on ROA. Define the estimated discretionary accruals of the treatment and control firms as DA_T and DA_c. Assume $DA_C = 0$, for the control firm. If the alternative hypothesis is the treatment firm's discretionary accruals are negative, the performance-matching procedure will likely be biased in favor of the alternative hypothesis. This occurs because $ROA_T = ROA_C$ due to matching, and the treatment firm belongs to the high operating cash flow quartile, but the control firm may not. Therefore, it is expected that the treatment firm's accruals are lower than the control firm's accruals. This makes it likely that $DA_T < DA_C$, i.e., the null is over-rejected. Therefore, matching on current ROA for extreme cash flow firms will cause test misspecification, as seen from Table 3.

[12]Assuming independence, a difference of about three percentage points between the rejection rates using two models is statistically significant.

modified-Jones-model discretionary accrual is likely to be systematically negative, as seen from the excessive rejection rate. Thus, unless a researcher is confident that credit sales represent accrual manipulation, the modified-Jones-model is expected to spuriously conclude earnings management. The results do indicate that while performance matching on the basis of ROA_t does not eliminate the bias in the modified-Jones model in low sales growth samples, it does for the Jones model.

4.4. Rejection rates for the alternative hypothesis of positive discretionary accruals

Simulation results of testing for positive discretionary accruals appear in panel B of Table 3. While misspecification in panel A was due primarily to excessive rejections of the null hypothesis, misspecification in panel B is primarily (but not entirely) a result of too infrequently rejecting the null hypothesis. For example, in samples of low earnings yield, low sales growth, small market capitalization and low cash flow firms, virtually all models except performance matching on ROA_t (based on either the Jones or modified-Jones model) conclude positive discretionary accruals too infrequently (less than 2.0% of the time).

To provide some specific comparisons, consider the low sales growth partition. Here a performance-based measure using ROA_t and based on the Jones (modified-Jones) model, rejects the null 4.4% (2.4%) of the time (i.e., is well-specified) while three out of four (all four) of the other accrual measures based on the Jones (modified-Jones) model reject at a rate of 1.6% (0.4%) or less. Corresponding numbers in low E/P samples show a rejection rate of 4.4% (3.6%) for the performance-based measure using ROA_t and the Jones (modified-Jones) model while all four (all four) of the other accrual measures reject at a rate of 1.6% (1.2%) or less based on the Jones (modified-Jones) model, respectively. This feature of the results also shows up in small capitalization stock samples which have a rejection rate of 7.2% (6.0%) for the performance-based measure using ROA_t and the Jones (modified-Jones) model while three out of four (all four) of the other accrual measures reject at a rate of 1.6% (1.6%) or less.

Finally, the inclusion of ROA as an additional regressor in the Jones and modified-Jones accrual models does little to improve their specification. Performance matching using ROA_t (based on either the Jones or modified-Jones models) is the best approach except in samples of high sales growth or low cash flow as a percent of total assets. In these latter two settings the ROA performance-based measures are misspecified.

4.5. Summary

As expected on the basis of previous research, the Jones and modified-Jones models are severely misspecified in stratified random samples. Over-rejection of the null hypothesis is apparent primarily in tests of negative discretionary accruals, whereas under-rejection frequently occurs when testing for positive discretionary accruals. Overall, all of the discretionary accrual measures examined exhibit some degree of misspecification; no single measure is well specified under the null hypothesis in each and every sample partition. However, under a wide variety of sampling conditions, the evidence seems to support the use of a performance-matched measure based ROA_t and the Jones model (and to a lesser extent the same measure based on the modified-Jones model). While there are instances where even this measure is misspecified, the results suggest that this performance-matched discretionary accrual measure are likely to be a viable alternative to existing discretionary accrual models for use in the research on earnings management.

5. The power of the test based on performance-matched discretionary accrual measures

Table 4 summarizes the results of comparing the power of the Jones and modified-Jones models with and without performance matching based on ROA_t and ROA_{t-1}.We report rejection frequencies for random and stratified-random samples of 100 firms with plus/minus 1%, 2%, 4%, or 10% accrual added to each firm's estimated discretionary accrual. The percentage accrual refers to accrual as a percentage of the firm's total assets. For each sample the indicated seed level is added to total accruals *before* estimating the respective discretionary accrual model. Panel A (B) reports results for the Jones (modified-Jones) model where the seeded abnormal accrual is negative and panel C (D) reports corresponding results where the seeded abnormal accrual is positive. In our tests we model earnings management that is 50% revenue-based. In particular, we assume that half of the abnormal accrual arises from credit sales and also add half of the seed to the change in sales and change in accounts receivable before estimating the discretionary accrual models.[13]

Even though the results in Table 3 show that the Jones and modified-Jones discretionary accrual measures suffer more from misspecification than performance-matched accrual measures, Table 4 reports results on the power of the test for all of these alternative discretionary accrual measures. The rationale is as follows. While a researcher could discard all discretionary accrual measures except the one subject to the least misspecification, such an approach implicitly assumes the cost of a Type I error is high while that of a Type II error is low. However, if the cost of a Type II error is high and that of a Type I error low, a researcher would make a different trade-off between power and specification. In such a setting, a researcher would prefer a test with a higher probability of rejecting the null hypothesis, even though the probability of a false rejection (Type I error) is greater than that using another model (e.g., another discretionary accrual measure). Since we do not know the relative costs of Type I and II errors, we report results for the power of the test for all discretionary accruals measures.[14] The results provide future researchers with the

[13] We also conduct the analysis assuming 0% and 100% revenue-based earnings management. The results are qualitatively similar to those reported in Table 4 and are available from the authors.

[14] Amemia (1994, p. 185) notes that "Classical statisticians usually fail to do this, because a consideration of the costs tends to bring in a subjective element."

Table 4
A comparison of the power of the test for the full sample and upper and lower quartiles of sub-samples formed on the basis of book-to-market ratio, sales growth, earnings-to-price (EP) ratio, firm size and operating cash flow

For each sample the indicated seed level is added to total accruals before estimating the respective discretionary accrual model. We assume that half of the abnormal accrual arises from credit sales and also add half of the seed to the change in sales and change in accounts receivable before estimating the discretionary accrual models. The table reports the percentage of 250 samples of 100 firms each where the null hypothesis of zero discretionary accruals is rejected at the 5% level (upper and lower one-tailed tests). The significance of the mean discretionary accrual of each sample is based on a cross-sectional t-test. The performance-matched discretionary accrual measures are constructed by precision matching each treatment firm with a control firm based on return on assets in period $t-1$. Firm-year accrual observations are constructed from the COMPUSTAT Industrial Annual and Research files from 1963 through 1999. We exclude observations if they do not have sufficient data to construct the accrual measures or if the absolute value of total accruals scaled by total assets is greater than one. We eliminate observations where there are fewer than ten observations in a two-digit industry code for a given year and where a performance-matched firm cannot be obtained. The underlying accrual models (Jones and Modified Jones) include a constant term. All variables are winsorized at the 1st and 99th percentiles. The final sample size is 122,798

Seeded abnormal accrual (%)	All Firms	Book-to-Market		Sales Growth		EP Ratio		Size		Operating Cash Flow	
		High	Low	High	Low	High	Low	Large	Small	High	Low
Panel A: Jones Model H_A: Accruals<0 (Figures in bold (*bold italic*) signify rejection rates that are significantly exceed (fall below) the 5% nominal significance level of the test and indicate that such tests are biased against (in favor of) the null hypothesis).[a]											
Rejection rates for the Jones model											
0	4.0	**8.4**	**12.8**	***1.2***	**18.8**	***1.2***	**68.0**	2.0	**25.6**	2.4	**34.4**
−1	20.0	35.6	34.4	6.8	46.8	16.8	81.6	22.0	44.4	13.6	53.6
−2	37.6	70.0	45.6	16.0	76.8	43.2	91.6	69.6	66.0	26.8	72.0
−4	88.4	98.8	85.6	53.6	96.8	90.8	98.4	99.6	98.0	76.4	93.6
−10	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Rejection rates for the performance-matched jones with matching on ROA_t											
0	4.4	6.4	8.0	2.8	5.2	**8.4**	3.2	**14.4**	**10.4**	**17.6**	***1.2***
−1	13.2	19.2	10.0	5.2	12.0	26.0	10.4	42.0	14.4	27.6	6.8
−2	25.6	39.6	15.6	11.6	20.0	40.4	22.4	64.0	28.0	45.2	8.4
−4	58.4	76.0	41.6	32.4	46.0	84.8	39.6	89.6	53.6	81.2	20.4
−10	99.6	100.0	96.0	91.2	97.6	100.0	94.4	100.0	97.2	100.0	80.4
Rejection rates for the performance-matched Jones Model with Matching on ROA_{t-1}											
0	**8.4**	2.4	**11.2**	***0.4***	**8.4**	4.8	**19.2**	**9.6**	6.0	**12.8**	**12.8**
−1	13.6	12.0	22.0	6.0	23.2	17.2	29.2	26.0	14.8	28.0	18.8
−2	29.6	31.2	36.4	13.2	37.2	36.8	51.2	61.6	30.4	45.2	29.6
−4	72.0	74.0	65.2	41.6	76.8	85.6	79.2	96.4	65.2	84.8	66.0
−10	100.0	100.0	98.4	98.0	99.6	100.0	100.0	100.0	100.0	100.0	99.6

Panel B: Modified Jones Model H_A: Accruals < 0 (Figures in bold (***bold italic***) signify rejection rates that are significantly exceed (fall below) the 5% nominal significance level of the test and indicate that such tests are biased against (in favor of) the null hypothesis).[a]

Rejection rates for the modified Jones model

0	4.4	**14.0**	**10.4**	***0.0***	**46.4**	**0.8**	**74.8**	2.0	**32.0**	***0.4***	**40.8**
−1	21.6	52.4	29.6	2.0	80.4	16.0	88.8	19.6	56.4	5.2	63.6
−2	37.2	78.8	40.4	6.8	94.0	41.6	95.2	66.0	74.4	17.6	78.4
−4	87.2	99.6	81.2	21.6	100.0	89.6	99.6	99.2	98.0	61.6	95.2
−10	100.0	100.0	100.0	98.8	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Rejection rates for the performance-matched modified Jones model with matching on ROA_t

0	4.4	6.4	**8.4**	***1.2***	**14.0**	**9.2**	4.4	**14.4**	**10.4**	**18.0**	***1.2***
−1	11.6	19.2	8.4	*1.2*	19.6	28.4	9.2	41.6	16.4	25.2	6.4
−2	24.0	40.4	12.8	4.8	34.8	40.8	22.0	61.6	28.0	44.8	7.2
−4	54.8	76.0	36.0	15.2	63.2	82.0	42.4	88.4	56.0	78.4	19.2
−10	99.6	100.0	95.2	81.6	99.6	100.0	93.6	100.0	96.0	100.0	80.0

Rejection rates for the performance-matched modified Jones model with matching on ROA_{t-1}

0	7.6	2.8	**8.8**	***0.0***	**19.2**	5.2	**24.8**	**8.8**	**9.6**	7.6	**13.2**
−1	13.2	15.2	18.8	2.8	42.4	18.4	33.2	24.4	16.8	25.2	20.4
−2	28.8	38.4	35.2	4.0	57.2	32.4	54.4	59.2	36.0	36.8	33.2
−4	72.0	78.8	60.4	24.4	90.4	83.2	83.6	95.6	71.2	80.0	68.4
−10	100.0	100.0	98.0	92.8	100.0	100.0	100.0	100.0	100.0	100.0	99.6

Panel C: Jones Model H_A: Accruals > 0 (Figures in bold (***bold italic***) signify rejection rates that are significantly exceed (fall below) the 5% nominal significance level of the test and indicate that such tests are biased against (in favor of) the null hypothesis).[a]

Rejection rates for the Jones model

0	6.0	2.4	2.4	**11.2**	***1.2***	**14.0**	***0.0***	7.2	***0.8***	**9.6**	***0.4***
1	20.4	18.4	4.0	23.2	3.2	39.2	***0.0***	38.8	2.4	33.2	***0.8***
2	40.4	56.4	10.4	42.8	17.2	83.2	***0.4***	73.2	8.0	56.4	4.0
4	88.8	98.4	40.4	78.4	65.6	99.6	12.4	97.6	45.2	93.6	25.6
10	100.0	100.0	98.8	100.0	100.0	100.0	96.8	100.0	99.6	100.0	100.0

Rejection rates for the performance-matched Jones model with matching on ROA_t

0	5.6	2.8	5.6	**8.8**	4.4	3.2	4.4	***0.4***	7.2	***0.8***	**12.8**
1	14.0	13.6	6.8	15.2	13.6	12.8	8.0	5.2	7.2	2.0	24.8
2	29.6	27.2	14.8	24.8	31.6	28.4	14.4	18.8	17.2	6.0	28.0
4	65.2	71.2	36.0	51.2	58.8	70.0	34.8	65.6	40.8	35.2	52.8
10	99.6	100.0	95.6	97.6	99.2	100.0	95.2	99.6	98.0	96.8	97.6

Rejection rates for the Performance-Matched Jones Model with Matching on ROA_{t-1}

0	4.0	**13.6**	2.0	6.0	2.4	6.4	***1.6***	2.4	3.2	***0.8***	**1.6**
1	14.0	24.8	8.0	19.2	10.0	21.2	2.0	15.6	8.4	7.2	8.0
2	32.8	50.4	12.0	30.8	23.6	50.8	4.4	36.8	19.6	16.0	8.8
4	75.2	91.6	28.0	57.2	54.4	91.6	22.0	84.8	54.0	53.2	40.4
10	100.0	100.0	96.0	99.6	100.0	100.0	95.2	100.0	99.2	99.6	98.4

Table 4 (*continued*)

Seeded abnormal accrual (%)	All Firms	Book-to-Market		Sales Growth		EP Ratio		Size		Operating Cash Flow	
		High	Low	High	Low	High	Low	Large	Small	High	Low
Panel D: Modified Jones Model H_A: Accruals > 0 (Figures in bold (*bold italic*) signify rejection rates that are significantly exceed (fall below) the 5% nominal significance level of the test and indicate that such tests are biased against (in favor of) the null hypothesis).[a]											
Rejection rates for the modified Jones model											
0	5.2	1.2	3.2	**32.4**	***0.0***	**15.6**	***0.0***	**9.2**	***0.4***	**21.2**	***0.0***
1	21.2	10.8	4.0	55.2	***0.4***	44.0	***0.0***	42.4	0.8	47.6	***0.0***
2	38.0	43.6	13.2	69.2	4.4	82.8	***0.4***	77.6	4.4	72.0	2.4
4	88.4	96.4	43.6	93.6	34.0	99.6	8.8	98.4	37.6	97.2	22.0
10	100.0	100.0	98.8	100.0	99.6	100.0	94.4	100.0	100.0	100.0	98.8
Rejection rates for the Performance-matched modified Jones model with matching on ROA_t											
0	4.8	2.4	6.8	**20.8**	2.4	3.6	3.6	***0.4***	6.0	***0.8***	**13.6**
1	12.8	10.8	9.2	28.0	4.0	12.4	8.0	3.2	8.0	2.4	24.4
2	26.8	23.6	14.0	39.6	14.0	22.8	13.2	17.6	16.0	5.6	29.2
4	60.0	64.8	36.4	66.4	43.2	58.0	34.8	62.4	35.2	30.8	54.0
10	99.6	100.0	94.8	98.8	97.6	100.0	94.4	99.2	96.0	94.0	98.0
Rejection rates for the Performance-Matched Modified Jones Model with Matching on ROA_{t-1}											
0	5.2	**9.6**	2.0	**16.0**	***0.0***	8.0	***1.2***	2.8	***1.6***	3.2	***0.8***
1	14.8	22.4	8.4	33.6	2.8	22.4	***1.2***	17.2	6.0	10.4	6.4
2	32.4	46.4	12.0	51.6	10.8	52.4	3.6	37.2	17.6	24.0	6.8
4	74.8	88.4	32.8	75.6	40.4	92.0	18.0	86.0	50.0	63.6	36.8
10	100.0	100.0	96.4	99.6	99.2	100.0	95.6	100.0	99.2	100.0	98.4

[a]Discretionary accruals from the Jones model are estimated for each industry and year as follows: $TA_{i,t} = \alpha_0 + \alpha_1/ASSETS_{i,t-1} + \alpha_2 \Delta SALES_{i,t} + \alpha_3 PPE_{i,t} + \varepsilon_{it}$, where TA_{it} is defined as the change in non-cash current assets minus the change in current liabilities excluding the current portion of long-term debt minus depreciation and amortization [with reference to COMPUSTAT data items, $TA = (\Delta Data4 - \Delta Data1 - \Delta Data5 + \Delta Data34 - Data14)/$ lagged Data6]. $\Delta SALES_{i,t}$ is change in sales scaled by lagged total assets ($ASSETS_{i,t-1}$), and $PPE_{i,t}$ is net property, plant and equipment scaled by $ASSETS_{i,t-1}$. Discretionary accruals from the Jones Model (Modified-Jones Model) with ROA are similar to the Jones Model (Modified-Jones Model) except for the inclusion of current or lagged year's ROA as an additional explanatory variable. For performance matched discretionary accruals, we match firms on ROA in period t or $t-1$. To obtain a performance-matched Jones model discretionary accrual for firm i we subtract the Jones model discretionary accrual of the firm with the closest ROA that is in the same industry as firm i. A similar approach is used for the modified Jones model.

information necessary to choose among competing models on the basis of both specification and power.[15]

Focusing first on panel A (Jones model), across sample partitions the Jones model, on average, detects a −1%, −2% and −4% discretionary accrual, as a percentage of assets, about 34.1%, 55.9% and 89.1% of the time. Alternatively, a performance-matched measure based on ROA_t and the Jones model, on average, detects a −1%, −2% and −4% discretionary accrual about 17.0%, 29.2% and 56.7% of the time, while average rejection rates across sample partitions for the performance-matched measure based on ROA_{t-1} are 19.2%, 36.6% and 73.3%. As suggested above, the relatively high rejection rates for the Jones model are largely attributable to misspecification. For example, the rejection rate for low EP Ratio firms is 68% with no seeding of discretionary accruals. Therefore, it is not surprising that the rejection rate is 81.6% for a discretionary accrual seeding of only −1% of assets. Obviously, researchers should not necessarily interpret these findings as the Jones model being highly powerful in detecting earnings management.

The modified-Jones Model (panel B) detects on average a −1%, −2% and −4% discretionary accrual, as a percentage of assets, about 39.6%, 57.3% and 84.8% of the time across sample partitions. Alternatively, a performance-matched measure based on ROA_t and the modified-Jones model, on average, detects a −1%, −2% and −4% discretionary accrual about 17.0%, 29.2% and 55.6% of the time. Corresponding average rejection rates across sample partitions for the performance-matched measure based on ROA_{t-1} are 21.0%, 37.8% and 73.5%. Across all sample partitions and accrual measures, a −10% discretionary accrual, as a percentage of assets, is almost certain to be detected all of the time.

The Jones model (panel C) detects on average a 1%, 2% and discretionary accrual, as a percentage of assets, about 16.7%, 35.7% and 67.8% of the time. Alternatively, a performance-matched measure based on ROA_t and the Jones model will, on average, detect a 1%, 2% and 4% discretionary accrual about 11.2%, 21.9% and 52.9% of the time, while average rejection rates across sample partitions for the performance-matched measure based on ROA_{t-1} are 12.6%, 26.0% and 59.3%.

Finally, the modified-Jones Model (panel D) detects on average a 1%, 2% and 4% discretionary accrual, as a percentage of assets, about 20.6%, 37.1% and 65.4% of the time. Alternatively, a performance-matched measure based on ROA_t and the modified-Jones model, on average, detects a 1%, 2% and 4% discretionary accrual about 11.2%, 20.2% and 49.6% of the time. Corresponding average rejection rates across sample partitions for the performance-matched measure based on ROA_{t-1} are 13.2%, 26.8% and 59.9%. As was the case in panels A and B for a −10% discretionary accrual, across sample partitions and accrual measures in panel C and D, a 10% discretionary accrual is also almost certain to be detected all of the time.

The power results reported in Table 4 do not produce a single measure that is uniformly the most powerful across all the sample partitions considered. In a similar vein, the results do not suggest that any single measure is inferior to all others across the sample partitions examined. Subject to the point we made above about the

[15]We thank the referee for recommending this interpretation of the power results.

trade-off between Type I and Type II errors, collectively, we interpret the results from Table 3 (Type I errors) and Table 4 (power of the test) as evidence in support of using a performance-matched measure based ROA_t and the Jones model (and to a lesser extent the same measure based on the modified-Jones model).

In summary, from an economic perspective, the results of the power tests suggest that plausible levels of positive earnings management, say, for example, discretionary accruals of 2% (4%) of total assets will empirically be detected about 30% (60%) of the time. Corresponding values for negative earnings management of the same magnitude are 40% (70%). Finally, the higher rejection rates of the Jones and modified-Jones model are due largely to misspecification and not greater power.

6. Sensitivity analyses

We perform a variety of additional tests. These tests confirm the main conclusion that performance matching improves test statistic specification. We briefly summarize the motivation for and main findings from those tests.

6.1. Sensitivity to including a constant in the Jones and modified-Jones regression models

The Type I error rate and power of the test are clearly a function of the underlying properties of the discretionary accrual measures reported in panel B of Table 1. The asymmetry (relative to zero) of the discretionary accrual measures reported there directly affects the Type I error rates and power we report. As noted above, our estimations of the parameters of the Jones and modified-Jones models include a constant term in the regressions. Untabulated results indicate that failure to include a constant magnifies misspecification of the Jones and modified-Jones models. For example, when no intercept is included in the cross sectional within-industry Jones model, rejection rates increase by more than 20% over those reported in Panel A of Table 3 for the high book-to-market, low book-to-market, low sales growth, low E/P ratio and low operating cash flows sub-samples where over-rejection already occurs. Based on this evidence, inclusion of a constant term when estimating traditional accrual models appears warranted.

6.2. Matching on ROA versus including ROA as a regressor in discretionary accrual models

Given a concern for controlling for performance on measured discretionary accruals, the choices available include adding a performance measure (e.g., ROA) to the accrual model as an additional regressor and matching treatment and control firms on the basis of the performance variable. An advantage of the former is that it imposes one less data requirement as data availability for a control firm is not required. A disadvantage is that it imposes a specific functional form (linearity) on the relation between performance and accruals. The results reported above indicate, that performance matching performs at least as well and, typically better than adding ROA as an additional regressor to the accrual regression models. In this section we provide some preliminary evidence to explain why this is the case.

We conjecture that matching performs better than the regression approach, at least in part, because of a non-linear relation between performance (measured by ROA) and accruals. Butler et al. (2004) find that firms experiencing extremely poor performance engage in liquidity-related transactions (e.g., factoring receivables and delaying payment of accounts payables) and record asset impairment charges to reflect economic declines in asset values. Asset write-downs have a direct and negative impact on income and are a reflection of accounting conservatism. In contrast, firms experiencing significant growth record large accruals (e.g., increases in inventory and accounts receivable), but these accruals do not have a one-to-one income consequence.[16]

Our analysis of the decile portfolios ranked on the basis of total accruals (not reported) bears out a non-linear relation between accruals and ROA. Firms falling in the lowest decile of total accruals have accruals that are proportionately much lower than their ROA compared to firms in the second accrual decile. At the other extreme, firms in the highest accrual decile have accruals that greatly exceed ROA because they are experiencing unusual growth (note that both total accruals and ROA have total assets in the denominator). For example (un-tabled results show), median sales growth in the highest accrual decile is 49% (24% in the 9th decile) compared to just 2% for firms in the 1st decile (6% for firms in the 2nd decile). This means most of the negative accruals for firms in the first decile are not due to declining sales, but due to reasons like asset write-down that affects both accruals and ROA equally. In contrast, for firms in the high accrual deciles, sales growth is the primary cause of extreme accruals and these growth-related accruals (e.g., increases in inventory, accounts receivable, etc.) have a proportionately smaller impact on ROA compared to the case of extreme negative accruals. Overall, we expect and find that ROA is more (less) closely associated with accruals when accruals are extremely negative (positive). This suggests a non-linear relation between accruals and ROA. Therefore a linear regression approach to estimating discretionary accruals is unlikely to perform as well as a matched-firm approach.

6.3. Sample size

Many earnings management studies estimate average discretionary accruals using very large samples (e.g., over 1,000 in tests of earnings management in the initial public offerings studies like Teoh et al., 1998b). While we do not analyze samples of 1,000 (or more) firms, we do provide evidence on specification and power in larger samples by examining samples of 200 and 300 firms. In general, across the stratified-random samples we use, misspecification is exacerbated in larger samples under all discretionary accrual measures. The highest rejection rates under the null hypothesis

[16]Butler et al. (2004) provide evidence that these asset-impairment charges are not the consequence of earnings management.

tend to be for the Jones- and modified-Jones models. Not surprisingly, the results indicate that the power of the test increases rapidly with sample size (e.g., regardless of the seeded accrual, in 300 firm samples rejection rates approach 100% for performance-matched samples). Overall, the results are consistent with those reported in Section 5 in that performance-matched accrual measures based on ROA_t offer specification and power advantages over the other measures although such advantages do not occur in each and every sample partition (as was the case in Section 5).[17]

6.4. Multi-year horizon

Many earnings management studies examine accrual behavior over multi-year horizons (e.g., Teoh et al., 1998a, b). To aid in drawing inferences from such research, we examine the specification and power of the discretionary accrual measures over 3- and 5-year horizons. We calculate a discretionary accrual each year of the 3- or 5-year period by re-estimating the Jones- and modified-Jones models every year. We then aggregate the annual discretionary accruals and test whether the cumulative average discretionary accrual significantly differs from zero. We find that the Jones- and modified-Jones models remain misspecified, but the rejection rates are less extreme than when the event period is one year. The performance-matched discretionary accrual models continue to offer specification and power advantages over the Jones and modified-Jones models across the sample partitions examined.

6.5. Relaxing the within-industry restriction on Jones model estimation

To relax data availability conditions, we implement the Jones model cross-sectionally using all non-financial firms, instead of just firms within-the same industry. Results are consistent with those reported earlier in Section 5. These results indicate that while matching is important, within-industry estimation is less important.

6.6. Alternative discretionary accrual measures

We also examine properties of other discretionary accrual measures including total accruals minus the industry median total accruals, total accruals minus a matched firm's total accruals and Jones (modified-Jones) model discretionary accrual minus the industry median discretionary accrual using the Jones (modified-Jones) model. None of these alternative discretionary accrual measures out-performs

[17]The increase in power with sample size suggests the cross-sectional independence assumption underlying the t-tests is a reasonable approximation. If the data were highly correlated cross-sectionally, then incremental reduction in the cross-sectional variance of discretionary accruals as a function of sample size would be small and we would not expect to observe an appreciable increase in power with increased sample size.

the performance-matched measures found in Section 5 to offer the most consistent performance across the sample partitions examined.

7. Summary and implications for future research

Researchers frequently use measures of discretionary accruals in tests for earnings management and market efficiency. In such studies, the Jones and modified-Jones models are the most popular choices for estimating discretionary accruals even though previous research shows that both the Jones and modified-Jones models are severely misspecified when applied to stratified-random samples of firms (e.g., Dechow et al., 1995; Guay et al., 1996).

We present detailed simulation evidence on the properties of alternative measures of discretionary accruals based on random and stratified-random samples. We also examine the properties of discretionary accrual models over multi-year horizons and their sensitivity to sample size. Under most circumstances, performance-matched discretionary accruals are well specified and powerful. While we observe some misspecification using the performance-matched discretionary accrual models in some settings, on balance, such measures are the most reliable from sample-to-sample in terms of Type I error rates. We do not interpret our findings as evidence that a performance-based discretionary accrual measure is the best measure in every conceivable setting. Rather we conclude that our findings suggest that a performance-matched discretionary accrual measure is useful in mitigating type I errors in cases where the researchers' partitioning variable of interest is correlated with performance. Our results also indicate that, researchers should include a constant term when estimating the Jones and modified-Jones models because doing so serves to further mitigate model misspecification. However, by design, our approach may increase the type II error rate. Specifically, our approach implies that our earnings measurement metric should be interpreted as "normal" earnings management. Controlling for performance, our approach classifies firms involved in average earnings management as not having managed earnings. As a result, depending on the specific setting, researches must decide whether the fitted values (after controlling for performance) or the residuals should be used as the variables of interest.

In addition to the above caveat, our study has three additional limitations. First, we ignore the consequences of the error embedded in estimated total accruals (and therefore in discretionary accruals) as a result of using the balance-sheet approach to estimating total accruals. According to Collins and Hribar (2002), the error in the balance-sheet approach of estimating accruals is correlated with firms' economic characteristics. Therefore, the error not only reduces the discretionary accrual models' power to detect earnings management, but also has the potential to generate incorrect inferences about earnings management. An interesting extension of our study would be to measure total accruals using the cash-flow statement approach advocated in Collins and Hribar (2002). Second, although we simulate several event conditions (e.g., multi-year performance, sample size, and various stratified-random samples), our results may not generalize to other research settings. In addition, we have made certain research design choices like cross-sectional within-industry estimation of the Jones and modified-Jones models, and re-estimation of the models each year as we examine a multi-year horizon that may not be appropriate in all accounting research settings. Finally, while we find that tests using performance-matched measures do not over-reject the null of zero discretionary accruals even in stratified-random samples, we cannot be sure that this necessarily indicates that the tests are always well specified. Accounting theory suggests stratified-random sample firms are more likely to engage in earnings management. Therefore, powerful tests should reject the null hypotheses in stratified-random samples. Nonetheless, performance-matched accrual measures provide evidence on whether the extent of discretionary accruals in stratified-random samples exceeds that in matched samples with similar performance characteristics, except for the treatment event that is the focus of the researcher's hypothesis.

References

Amemia, T., 1994. Introduction to Statistics and Econometrics. Harvard University Press, Cambridge, MA.

Barber, B., Lyon, J., 1996. Detecting abnormal operating performance: The empirical power and specification of test statistics. Journal of Financial Economics 41, 359–399.

Barber, B., Lyon, J., 1997. Detecting long-run abnormal stock returns: the empirical power and specification of test statistics. Journal of Financial Economics 43, 341–372.

Barth, M., Cram, D., Nelson, K., 2001. Accruals and the prediction of future cash flows. The Accounting Review 76, 27–58.

Basu, S., 1997. The conservatism principle and the asymmetric timeliness of earnings. Journal of Accounting & Economics 24, 3–37.

Beaver, W., Clarke, R., Wright, F., 1979. The association between unsystematic security returns and the magnitude of earnings forecast errors. Journal of Accounting Research 17, 316–340.

Brooks, L., Buckmaster, D., 1976. Further evidence on the time series properties of accounting income. Journal of Finance 31, 1359–1373.

Brown, S., Lo, K., Lys, T., 1999. Use of R^2 in accounting research: Measuring changes in value relevance over the last four decades. Journal of Accounting & Economics 28, 83–115.

Butler, M., Leone, A., Willenborg, M., 2004. An empirical analysis of auditor reporting and its association with abnormal accruals. Journal of Accounting & Economics 37, 139–165.

Collins, D., Hribar, P., 2002. Errors in estimating accruals: implications for empirical research. Journal of Accounting Research 40, 105–135.

Campbell, D., Stanley, J., 1963. Experimental and Quasi-Experimental Designs for Research. Houghton-Mifflin Company.

Cook, T., Campbell, D., 1979. Quasi-Experimentation Design and Analysis Issues for Field Settings. Houghton-Mifflin Company.

Dechow, P., Sloan, R., Sweeney, A., 1995. Detecting earnings management. The Accounting Review 70, 193–225.

Dechow, P., Kothari, S., Watts, R., 1998. The relation between earnings and cash flows. Journal of Accounting & Economics 25, 133–168.

DeFond, M., Jiambalvo, J., 1994. Debt covenant violation and manipulation of accruals: Accounting choice in troubled companies. Journal of Accounting & Economics 17, 145–176.

DeFond, M., Park, C., 1997. Smoothing income in anticipation of future earnings. Journal of Accounting & Economics 23, 115–139.

DeFond, M., Subramanyam, K., 1998. Auditor changes and discretionary accruals. Journal of Accounting & Economics 25, 35–67.
Fama, E., French, K., 2000. Forecasting profitability and earnings. Journal of Business 73, 161–176.
Fama, E., MacBeth, J., 1973. Risk, return, and equilibrium: Empirical tests. Journal of Political Economy 81, 607–636.
Fields, T., Lys, T., Vincent, L., 2001. Empirical research on accounting choice. Journal of Accounting & Economics 31, 255–307.
Freeman, R., Tse, S., 1992. A nonlinear model of security price responses to unexpected earnings. Journal of Accounting Research 30, 185–209.
Guay, W., Kothari, S., Watts, R., 1996. A market-based evaluation of discretionary accrual models. Journal of Accounting Research 34, 83–105.
Guidry, F., Leone, A., Rock, S., 1999. Earnings-based bonus plans and earnings management by business unit managers. Journal of Accounting and Economics 26, 113–142.
Healy, P., 1996. Discussion of a market-based evaluation of discretionary accrual models. Journal of Accounting research 34, 107–115.
Holthausen, R., Larcker, D., 1996. The financial performance of reverse leveraged buyouts. Journal of Financial Economics 42, 293–332.
Ikenberry, D., Lakonishok, J., Vermaelen, T., 1995. Market underreaction to open market share repurchases. Journal of Financial Economics 39, 181–208.
Jones, J., 1991. Earnings management during import relief investigations. Journal of Accounting Research 29, 193–228.
Kang, S., Sivaramakrishnan, K., 1995. Issues in testing earnings management and an instrumental variable approach. Journal of Accounting Research 33, 353–367.
Kasznik, R., 1999. On the association between voluntary disclosure and earnings management. Journal of Accounting Research 37, 57–81.
Kothari, S., Sabino, J., Zach, T., 2005. Implications of survival and data restrictions on tests of market efficiency. Journal of Accounting & Economics forthcoming.
Loughran, T., Ritter, J., 1995. The new issues puzzle. Journal of Finance 50, 23–51.
Lyon, J., Barber, B., Tsui, C., 1999. Improved methods for tests of long-run abnormal stock returns. Journal of Finance 54, 165–201.
Newey, W., West, K., 1987. A simple, positive semi-definite heteroscedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.
Peasnell, K., Pope, P., Young, S., 2000. Detecting earnings management using cross-sectional discretionary accrual models. Accounting & Business Research 30, 313–326.
Perry, S., Williams, T., 1994. Earnings management preceding management buyout offers. Journal of Accounting & Economics 18, 157–179.
Rees, L., Gill, S., Gore, R., 1996. An investigation of asset write-downs and concurrent abnormal accruals. Journal of Accounting Research 34, 157–169.
Subramanyam, K.R., 1996. The pricing of discretionary accruals. Journal of Accounting and Economics 22, 249–282.
Teoh, S., Welch, I., Wong, T., 1998a. Earnings management and the long-run underperformance of seasoned equity offerings. Journal of Financial Economics 50, 63–100.
Teoh, S., Welch, I., Wong, T., 1998b. Earnings management and the long-run underperformance of initial public offerings. Journal of Finance 53, 1935–1974.
Watts, R., 2003. Conservatism in accounting Part I, Explanations and implications. Accounting Horizons 17, 207–221.
Watts, R., Zimmerman, J., 1986. Positive Accounting Theory. Prentice Hall, Englewood Cliffs, NJ.
White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 48, 817–838.