A Simple Approach to Better Distinguish Real Earnings Manipulation from Strategy Changes*

THEODORE E. CHRISTENSEN, *University of Georgia*†

ADRIENNA HUFFMAN, *The Brattle Group*

MELISSA F. LEWIS-WESTERN, *Brigham Young University*

KRISTEN VALENTINE, *University of Georgia*

ABSTRACT

Researchers typically infer real earnings management when a firm's operating and investing activities differ from industry norms. A significant problem with classifying deviations from industry averages as myopic earnings management is that companies can change their operating and investing decisions for strategic business reasons rather than to mislead stakeholders. Using principal components analysis, we systematically evaluate existing measures and develop a comprehensive real activities measure to better capture earnings manipulation. Our measure reflects (i) deviations from industry averages across multiple activities and (ii) other signals of manipulation. This approach is promising because, although there are many sources of abnormal activities, manipulation is more likely the cause when managers engage in multiple income-increasing abnormal activities that coincide with other signals that indicate an elevated risk of manipulation. This simple approach results in a metric that associates negatively with future operating performance and earnings persistence, yields high-power tests, and captures manipulation reasonably well across most life-cycle stages. Importantly, this approach performs better than the standard real earnings management metrics across all dimensions. Specifically, it generates the expected reduction in future earnings and reduced earnings persistence in 82% of the tests compared to 36% and 46% in common alternatives. Also, because this innovation does not require a long time-series or rely on future period realizations for classification, it can be useful in more research settings than other recent innovations in the literature.

Keywords: abnormal accruals, real earnings management, earnings management, financial reporting quality, beneficial earnings management

---

* Accepted by Paul Hribar. We are grateful to Paul Hribar (editor) and anonymous reviewers who helped shape this manuscript. We also thank Nerissa Brown (FARS discussant), Frank Heflin, Won Kim, Greg Martin, Samuel Melessa, Sugata Roychowdhury (AAA discussant), Tim Seidel, Rachel Scott, Steve Stubben, and participants at UC Berkeley, UC Davis, Tulane University, the Securities and Exchange Commission, and the University of North Carolina at Charlotte workshops, and the 2015 BYU Accounting Research Symposium, 2016 American Accounting Association (AAA) annual meeting, and 2017 AAA FARS mid-year meeting for their helpful comments. We thank Karen Hennes, Andy Leone, and Brian Miller for use of use of their restatement data. We appreciate research assistance provided by Rachel Scott. We are grateful for financial support from the Terry College of Business and the Marriott School of Management. A previous version of this paper was circulated under the title "Earnings Management Proxies: Prudent Business Decisions or Earnings Manipulation?" We provide SAS code to create the REMComp measure at the following website: https://github.com/kristengvalentine/CHLV2022_A_Simple_Approach_to_Better_Distinguish_Real_Earnings_Manipulation_from_Strategy_Changes

† Corresponding author.

Une approche simple pour mieux distinguer la manipulation des résultats réels et les changements de stratégie

RÉSUMÉ

En règle générale, les chercheurs infèrent que la gestion des résultats est réelle lorsque les activités opérationnelles et d'investissement d'une entreprise diffèrent des normes industrielles. Un problème important lié au fait de considérer les écarts aux moyennes industrielles comme une gestion myope des résultats est que les entreprises peuvent modifier leurs décisions opérationnelles et d'investissement pour des raisons commerciales stratégiques plutôt que pour induire les parties prenantes en erreur. Les auteurs évaluent de manière systématique les mesures existantes et développent une mesure complète des activités réelles pour mieux saisir la manipulation des résultats à l'aide de l'analyse en composantes principales. Leur mesure reflète (i) les écarts aux moyennes industrielles pour de multiples activités et (ii) d'autres signaux de manipulation. Cette approche est prometteuse, car bien qu'il existe de nombreuses sources d'activités anormales, la manipulation est plus probablement en cause lorsque les dirigeants s'engagent dans de multiples activités anormales augmentant les revenus et coïncidant avec d'autres signaux indiquant un risque élevé de manipulation. Cette approche simple permet d'obtenir une mesure qui s'associe négativement à la performance opérationnelle future et à la persistance des résultats, de mettre au point des tests à forte puissance et de bien saisir la manipulation dans la plupart des étapes du cycle de vie. Il est important de noter que cette approche est plus performante que les mesures standards de gestion réelle des résultats dans toutes les dimensions. Plus précisément, elle génère la diminution attendue des résultats futurs et la diminution de la persistance des résultats dans 82 % des tests, contre 36 % et 46 % pour les alternatives courantes. En outre, comme cette innovation ne nécessite pas de longues séries temporelles et ne repose pas sur des réalisations de périodes futures pour la classification, elle peut être utile dans un plus grand nombre de contextes de recherche en comparaison avec d'autres innovations récentes dans la littérature.

Mots-clés : régularisations anormales, gestion réelle des résultats, gestion des résultats, qualité de l'information financière, gestion bénéfique des résultats

1. Introduction

We explore a common form of managerial myopia, the purposeful use of manager discretion in structuring transactions to mislead stakeholders and refer to these actions as "real earnings *manipulation*" (REM; Healy and Wahlen 1999; Schipper 1989). Many studies examine "earnings management" broadly and "earnings manipulation" more specifically in a variety of contexts (Dechow et al. 2010; Roychowdhury et al. 2019). Although earnings manipulation is clearly of interest to academics, practitioners, and regulators, extant research also widely acknowledges that accurately measuring earnings manipulation can be extremely challenging (Badertscher et al. 2012; Dechow et al. 1995; Srivastava 2019). Moreover, although substantial improvements in the measurement of accrual-based earnings manipulation have emerged in recent years (Christensen et al. 2022), only a handful of studies have devoted attention to advancing REM measures. Furthermore, recent studies that develop improvements to existing REM measures leave room for additional innovations (Cohen et al. 2019; Srivastava 2019; Vorst 2016). Since refinements are ongoing and the relative benefits of the new measures are uncertain, researchers frequently use multiple standard REM proxies in a variety of contexts.[1] We systematically evaluate existing measures and propose a new comprehensive REM measure that better captures the construct of manipulation.

Because companies can change their operating and investing decisions for strategic reasons and it is difficult to distinguish between strategy adjustments and earnings manipulation, recent

1. Appendix 1 summarizes the number of papers published in the top five accounting journals over the last five years that examine real earnings management. Seventy-six percent of the 54 published papers use the standard real earnings management proxies to measure REM and these papers are subsequently cited by over 4,312 papers.

research recommends several modifications to the standard REM metrics (Cohen et al. 2019; Srivastava 2019; Vorst 2016). Srivastava (2019) views the problem to be severe and concludes that competitive strategy likely explains much of what prior research has classified as REM. To overcome this problem, Cohen et al. (2019), Srivastava (2019), and Vorst (2016) each propose alternative ways to distinguish earnings manipulation from strategy adjustments. However, there are benefits and costs to each approach.[2] We propose a simple innovation that (i) is easy to implement in a broad sample of firms, (ii) does not require a long time-series of data, and (iii) does not employ future period realizations in its development. We validate our proposed REM measure across various relevant dimensions to determine whether our approach results in a summary REM metric that performs as we would expect if it actually captures REM.

In addition, we compare our new measure to (i) the three standard REM metrics and (ii) summary measures constructed from the three standard metrics. These tests achieve two objectives. First, to the best of our knowledge, prior research has not systematically examined whether the standard REM metrics capture characteristics expected of REM (e.g., lower future performance when the costs of REM are realized). Second, these tests compare the performance of our proposed REM metric to the performance of the standard metrics to determine whether our proposed innovation provides meaningful improvement relative to the standard metrics.

We advance REM measurement by jointly considering the insights that (i) managers can simultaneously employ multiple strategies to manage real activities (Roychowdhury 2006) and (ii) REM is more likely to occur in particular settings. To the extent that managers concurrently use multiple tools to manage earnings, the combined consideration of these activities will better distinguish REM from strategy changes (Dickinson 2011; Porter 1980; Srivastava 2019). For example, observing declines in advertising expenditures concurrent with abnormal increases in production could suggest that the observed changes stem from an overarching plan to manipulate earnings rather than a change in expectations about future demand for the company's products. In addition, we improve our REM measure by incorporating information on settings in which managers face elevated incentives to employ REM.

We use principal components analysis (PCA) to develop a summary REM measure that reflects the concurrent use of multiple REM activities and three REM signals. This approach is promising because, although there are many possible motivations for abnormal activity levels aside from manipulation, it is less likely that anything other than manipulation can explain managers' use of multiple real activities that coincides with other signals of manipulation.[3] PCA generates a summary variable using weighted combinations of the individual metrics and it allows for different weightings across variables. Our PCA analysis results in a summary REM metric that reflects situations where managers cut discretionary expenses, overproduce inventory and offer steep discounts to prices or lenient credit terms. This metric is also higher when (i) there is less flexibility to use accrual-based earnings manipulation (as measured by balance sheet bloat) and (ii) when litigation risk is low.[4]

To validate our proposed PCA approach for detecting REM, we employ multiple tests in different settings and compare its performance to that of standard REM metrics. Specifically, the standard REM activities we examine are (i) abnormal cuts to discretionary expenses, (ii) unusually low cash flow from operations, and (iii) excess production (Cohen and Zarowin 2010; Gunny 2010; Roychowdhury 2006). Following prior research, we examine these three standard metrics individually and also the most commonly used combinations. A review of

2. For example, the approaches proposed by Vorst (2016) and by Srivastava (2019) each use future realizations in first-stage models or for classification of first-stage residuals, which may not be appropriate in addressing some types of research questions.

3. Although we emphasize managers' use of multiple REM activities to achieve income-increasing objectives throughout, our PCA approach is flexible enough to allow for consideration of managers' use of REM tactics to decrease income.

4. Huang et al. (2020) provide evidence that litigation risk reduces REM by constraining managers' ability to issue optimistic and misleading disclosures. That is, with higher litigation risk, opportunistic disclosure is less likely to occur and the need for REM to corroborate opportunistic disclosure declines.

recent research indicates that 76% of the studies examining real earnings management over the past five years use these methods.[5]

A primary characteristic of REM is lower future earnings as a result of current-period suboptimal operating and investing decisions (Roychowdhury 2006; Cohen and Zarowin 2010; Kothari et al. 2016; Vorst 2016). Thus, we examine the association between REM proxies and future operating performance because we expect REM to generate performance costs. The performance measures that we employ are future earnings and future cash flows in both one- and three-year horizons.[6] We investigate the relation between the REM proxies and future operating performance using a firm fixed effects research design, which tests whether the level of REM proxies is negatively associated with future operating performance relative to the firm's average performance in typical years. Thus, our analyses use the firm as its own control to help alleviate concerns regarding correlated omitted factors (Gow et al. 2016; Larcker and Rusticus 2010). We also include year fixed effects to control for economy-wide factors.

When we consider the standard metrics individually, we find that excess production and lower-than-expected cash flows are associated with *lower* future operating performance. In contrast, we do not find evidence that abnormal cuts to discretionary expenses are negatively associated with future performance. These results suggest that abnormal cuts to discretionary expenses—the most commonly used individual REM metric—are more susceptible to misclassification than other proxies. This result corroborates recent research suggesting that abnormal cuts to discretionary expenses do not reflect REM (Chen et al. 2018; Srivastava 2019). When examining two summary metrics based on simple sums of the standard metrics and calculating the metrics using the most common combinations of the standard metrics, we find that the first simple sum measure associates positively with future earnings. In contrast, the second simple sum measure negatively associates with future operating performance. Thus, the most commonly used approach to combining the standard metrics (i.e., using the two simple sum measures) produces REM variables that do not consistently associate negatively with future performance. In contrast, the PCA component—which reflects all three of the standard metrics and additional REM signals—negatively associates with future one- and three-year operating performance (both earnings and cash flows), consistent with REM capturing corporate myopia and short-termism. These results persist even when we supplement the model with the two simple sum measures, suggesting that the PCA approach is incremental to summations of the standard metrics.

To validate these results, we next consider restatements. The relation between REM and restatements should be positive or insignificant, but not negative. A negative relation would indicate that the abnormal activity reflects strategy rather than manipulation. The relation between the first simple sum measure and restatements is either significantly negative or not significantly different from zero. Both the second simple sum measure and our proposed PCA measure perform as expected for REM since they exhibit positive associations or no association with restatements. Overall, the future performance and restatement results suggest that our proposed PCA approach provides a summary REM metric that associates negatively with future performance, positively or insignificantly with restatements and performs better than the standard metrics.

Next, we examine earnings persistence. The future operating performance tests rely on prior researchers' assumption that earnings manipulation will generate observable future costs (Bowen et al. 2008; Demerjian et al. 2020; Kothari et al. 2016; Vorst 2016). However, the association between earnings manipulation and future operating performance may not be sufficiently strong to lead to observable decreases in future operating performance. Alternatively, additional earnings manipulation may mask these subsequent declines. To overcome this concern, we rely on a related stream of research that posits that although earnings manipulation might not result in observable decreases in future operating performance, it should be negatively associated with

5. Appendix 1 provides additional details.

6. We expect that three years is a sufficient horizon for the consequences of REM to be realized (Kothari et al. 2016).

earnings persistence (Sloan 1996; Richardson et al. 2005; Casey et al. 2017). We follow prior research (Casey et al. 2017; Richardson et al. 2005; Sloan 1996,) and estimate annual Fama and MacBeth earnings persistence regressions (Fama and MacBeth 1973). The results suggest that the first simple sum REM measure does not influence the persistence of cash flows, but our PCA REM measure is negatively associated with cash flow persistence.[7]

We next conduct a simulation analysis in which we seed income-increasing errors of 1% to 10% of assets for each activity we study (production, cash flows, and discretionary expenses), and then recalculate the abnormal metrics, the summary metrics, and the PCA metric. The results suggest that the component approach is more powerful in detecting REM than the standard metrics for all seeded error levels. For example, for seeded errors of 1%, the rejection rate for our PCA measure is 11.20% versus 7.6% for the two simple sum measures. At a 10% error level, our PCA measure is associated with a rejection rate of 84.80% versus 60.40% for the two simple sum measures.

Finally, we consider whether our approach is incremental to the improvement proposed by Vorst (2016). We replicate and corroborate his results and we control for his manipulation variables in our setting. We find that the PCA measure is still significantly negatively associated with future performance. These results suggest that the PCA approach identifies earnings manipulation that is incremental to the manipulation that Vorst (2016) identifies. Finally, we perform numerous additional tests and find that (i) the PCA approach works broadly across firm life-cycles (and better than the standard REM metrics), (ii) the PCA measure increases with earnings manipulation incentives, and (iii) our results are not sensitive to alternative definitions of performance, the control variables used, or the fixed effects employed.

We contribute to the earnings management literature in several ways. First, we systematically evaluate common REM measures and the extent to which they possess characteristics expected of a valid REM measure. Second, we develop and validate a simple REM measure that can be used in a variety of settings. Our PCA-generated REM metric possesses several characteristics expected of REM and it performs better than standard REM metrics across a variety of tests. Specifically, we find that our PCA variable exhibits a negative association with future performance and earnings persistence in 82% of the tests, while the first (second) simple sum measure negatively associates with future performance in 36% (46%) of tests.

Finally, our approach adds to the recent literature addressing the problem of measuring REM (Vorst 2016; Cohen et al. 2019; Srivastava 2019) in at least three ways. First, both Vorst's (2016) and Srivastava's (2019) approaches use future realizations or stock price data as either independent variables in first-stage identification models or for classification of first-stage residuals, whereas ours does not. The use of future information and stock prices limits the usefulness of these approaches in some settings (e.g., tests of market efficiency). Second, Srivastava (2019) provides improvements that stem from the observation that competitive strategies (and the resulting influence on operating and investing decisions) vary across firms within an industry, thereby diminishing the correspondence between the residual in first-stage manipulation models and earnings manipulation. Thus, he improves first-stage models of operating and investing decisions. Rather than improving the inherently complicated models of firm operating and investing decisions, we use the standard models to identify abnormal activities, but only classify those activities as manipulation if consideration of their combined use and other signals suggest that it is earnings manipulation. This alternative approach is promising since it results in an REM metric that is negatively associated with future operating performance and earnings persistence, yields

7. An underlying assumption in our future performance and persistence tests is that, on average, strategy adjustments would lead to improved performance or be neutral with respect to performance. In contrast, we expect earnings manipulation to generate future performance costs. We acknowledge, however, that our future performance tests cannot distinguish between poor strategy choices—that lead to lower future performance—and earnings manipulation. For this reason, we conduct simulation tests that do not share this weakness.

high-power tests, and performs reasonably well across most firm life-cycle stages. Third, we provide evidence that our PCA metric identifies earnings manipulation that is incremental to Vorst's (2016) approach. Ultimately, we introduce a well-specified REM measure, but future researchers can improve our approach by incorporating context-specific variables into the PCA that influence the likelihood that abnormal operating decisions reflect earnings manipulation.

2. Background and hypothesis development

Definition of earnings manipulation

The term "earnings management" has been widely used in the accounting, finance, and management literatures. The term, however, is vague and has been used to describe both opportunistic and informative discretionary reporting and operating activities. Schipper (1989) and Healy and Wahlen (1999) define manipulation as managers' use of judgments in financial reporting and in structuring transactions either to mislead some stakeholders about the underlying economic performance of the firm, or to influence contractual outcomes. Dechow and Skinner (2000) offer a similar definition and note that even within-GAAP discretion can be classified as manipulation. We focus on myopic or value-destroying earnings management, which we label "earnings manipulation."[8]

We contrast earnings manipulation with other types of earnings management like income smoothing or signaling, which attempts to (i) communicate private information or (ii) benefit the firm via the increased usefulness of earnings in contracting or more precise estimates of firm value. Real activities manipulation refers to managers' efforts to inflate earnings through changes in operating and investing decisions that can be defined as "departures from normal operational practices, motivated by managers' desire to mislead at least some stakeholders into believing certain financial reporting goals have been met in the normal course of operations" (Roychowdhury 2006, 337).

REM: Identification and measurement

Financial reporting is fundamental in our economy since it is the primary means by which we measure performance and allocate resources (Healy and Wahlen 1999). If real earnings management were easily unwound, then it will not influence users' investment decisions. Prior research, however, indicates that the quality of financial reporting, including earnings manipulation, influences investment, and credit decisions (see Dechow et al. 2010; Roychowdhury et al. 2019 for reviews). Thus, identifying and adjusting for earnings manipulation is important because it influences critical economic decisions. As a result, an evolving body of research assesses the validity of various innovations to REM measurement that build on the standard REM models proposed by Roychowdhury (2006) and Gunny (2010). These innovations use performance-matched samples, improvements in first-stage models of normal operating and investing decisions, and the classification of discretionary expense declines into those stemming from changing firm strategy or manipulation.

Performance matching

Cohen et al. (2019) use simulation analyses to examine how well-specified REM models are and whether performance matching improves model specification. They consider the three commonly used REM proxies (abnormal production, cash flow, and discretionary expenses) plus two related measures proposed by Gunny (2010): (i) abnormal selling and administrative costs and (ii) abnormal gains from the sale of fixed assets. On balance, they posit that performance matching may help, but conclude that performance-matched REM measures are not well specified

8. We focus on real earnings manipulation while Christensen et al. (2022) provide a broader framework focusing on accruals manipulation.

in many settings and that traditional and performance-matched metrics yield low-power tests. Their results suggest that there is ample room for improvements in REM measurement.

Improving first-stage models

Similar to our concern that commonly used REM proxies often do not accurately capture earnings manipulation, Srivastava (2019) posits that competitive strategy is a correlated omitted variable from first-stage REM models and that failure to account for this omitted variable leads to incorrect inferences. He finds that cohort age and competitive strategy are systematically associated with first-stage residuals, suggesting that firm characteristics heavily influence REM measures. This problem, however, is most severe for abnormal discretionary expenses. To combat this issue, he supplements the first-stage models with investment opportunity variables including market capitalization and the market-to-book ratio, future revenues, and past expenditures. These steps reduce the magnitude of the residuals and lead to residuals that are less correlated with competitive strategy.

Srivastava (2019) notes that despite these improvements, his approach is not a panacea for measuring REM because the approach overcorrects for errors in commonly occurring situations, such as when firms in the same industry also manage earnings or when earnings manipulation persists for more than one year. Prior research suggests that these situations are not infrequent. For example, managers often engage in consecutive years of income-increasing earnings manipulation (Badertscher 2011; Efendi et al. 2007; Ettredge et al. 2010) and industry peers often utilize similar earnings management strategies (Gleason et al. 2008; Kedia et al. 2015). In addition, the suggested supplemental first-stage variables rely on future information (future sales) and market-based information (market capitalization and the market-to-book ratio), both of which are problematic for tests of market understanding or efficiency. Finally, to the extent that competitive strategy, particularly as it relates to investment in R&D, is a distinct firm characteristic rather than an industry-cohort characteristic (Hirshleifer et al. 2018),[9] modeling strategy by industry-cohort is unlikely to be fully effective in removing the "strategy" effect from first-stage residuals. In short, developing high-quality models of the economic drivers of production and investment is a challenging task.

Suspect cuts to discretionary expenses

Srivastava (2019) finds that measurement errors associated with commonly used REM metrics are most severe for discretionary expenses. Recent research, however, has proposed improvements intended to better identify cuts to discretionary expenses that reflect earnings manipulation. Vorst (2016) posits that reversing cuts to discretionary expenses are more likely to reflect earnings manipulation than changing firm economics, which are more likely to be accompanied by sustained cuts to discretionary expenses. He identifies cuts to discretionary expenses that reverse in the following year and examines how reversing and non-reversing cuts associate with future operating performance. His evidence suggests that reversing cuts are associated with lower future operating performance whereas non-reversing cuts are not.

Bereskin et al. (2018) explore manipulation-motivated cuts to R&D. To do so, they classify abnormal cuts to R&D as manipulation in years when cutting R&D allows the firm to report earnings that beat the prior year's earnings. They then examine the effect of suspect R&D cuts on innovative outputs. After controlling for firms' investment opportunities, their results suggest that manipulation-motivated cuts to R&D are costly to the firm and result in a decrease in the number of future patents, future patent citations, and innovative efficiency.

Both Vorst (2016) and Bereskin et al. (2018) provide refinements in measuring REM through the cutting of discretionary expenses and they find that their approach is an improvement over

9. Hirshleifer et al. (2018, 2556) provide evidence that the ability to use innovative investment to create novel technology is a "competitive advantage" that "other firms have difficulty replicating."

standard REM measures. Although these advances are likely useful in numerous settings, they may not be appropriate for other settings. For example, Vorst's (2016) approach uses future realizations for classification, which may limit its usefulness in testing some research questions. In addition, Bereskin et al. (2018) assume that just beating last year's earnings is the primary motivation for earnings manipulation. Thus, their approach is less suited to situations where other incentives have a primary influence on operating decisions.

Correlation with accounting and auditing enforcement actions or restatements

In the United States, public firms' financial reports are subject to SEC review and firms failing to comply with current securities regulations may be the focus of an accounting and auditing enforcement action (AAER) or may be required to restate previously reported financials. Prior research argues that Type II errors are lower among AAER or restatement firms, although the use of these samples assumes that the SEC correctly identifies misstatement firms (Dechow et al. 2011). Prior research typically has not examined the relation between REM and AAERs or restatements, perhaps because REM results from suboptimal operating decisions rather than misapplication of GAAP. Some prior evidence suggests, however, that REM is associated with restatements because it is used more extensively as the firm's ability to engage in accrual-based earnings management declines (Badertscher 2011; Ettredge et al. 2010) or that managers use accruals and REM together to achieve reporting objectives (Beatty et al. 1995). Thus, it is not clear ex ante whether REM will associate positively with restatements, but it should not associate negatively with them. A negative association could suggest that the metric reflects changing strategy rather than manipulation.

Takeaways

Prior research provides ample evidence that (i) the process generating firms' operating and investing activities is not well understood and (ii) standard REM proxies do not uniformly capture the underlying construct of earnings manipulation. Due to the fundamental importance of earnings quality in our economy, however, and because new methodological approaches do not offer a perfect solution, researchers still frequently use many of the standard real earnings management proxies in different contexts. Appendix 1 tabulates the number of studies in the top-five accounting journals that use these proxies in a recent five-year period (2015–2020). There are 54 papers published in the top accounting journals during this period that study REM metrics, and 41 (76%) use abnormal cuts to discretionary expenses, abnormal production, or abnormal cash flow to measure REM following Roychowdhury (2006). In addition, the most common REM metric employed is abnormal cuts to discretionary expenses, the metric least likely to reflect manipulation. This evidence on the frequent use of these proxies despite known weaknesses emphasizes the importance of our examination and highlights the need for improvement.

Hypothesis 1

Prior research finds that managers sometimes make suboptimal, myopic operating and investing decisions to boost current earnings at the expense of long-term performance (Bhojraj et al. 2009; Cohen and Zarowin 2010; Graham et al. 2005; Roychowdhury 2006). Because these decisions artificially increase current earnings, these departures do not generally contribute to firm value and can have negative long-term consequences (Bhojraj et al. 2009; Cohen et al. 2008; Cohen and Zarowin 2010; Graham et al. 2005; Roychowdhury 2006). For example, Roychowdhury (2006, 338) concludes that "real activities manipulation can reduce firm value because actions taken in the current period to increase earnings can have a negative effect on cash flows in future periods" and that relative to accrual earnings manipulation, real activities manipulation imposes "greater long-term costs on the company." Similarly, Gunny (2010, 857) concludes that abnormal activities that generate positive future performance are not manipulation, but rather reflect a strategy to "enhance the firm's credibility and reputation with stakeholders."

As an example, if managers achieve an artificial boost to income by overproducing inventory to reduce the per-unit cost of goods sold, then future earnings will be lower as the firm bears the cost of this suboptimal decision. In the case of overproduction, the subsequent costs include the additional inventory holding costs associated with excess inventory and the price discounts that will apply if the stale inventory must be sold at a reduced price (Roychowdhury 2006). Similarly, if managers decrease their investment in R&D to meet short-term earnings goals, this action will reduce future earnings when the firm subsequently resumes the forgone R&D, or when the firm incurs the eventual costs of delayed investment (e.g., a competitor goes to market first due to the development delay or the firm misses testing schedules for product releases, which delays production and sales). Finally, offering steep price discounts or excessively lenient credit terms to boost sales will likely reduce future earnings. For example, Roychowdhury (2006, 338) notes that "aggressive price discounts to increase sales volume and meet some short-term earnings targets can lead customers to expect these discounts in future periods, as well. This argument can imply lower margins on future sales." With respect to lenient credit, Roychowdhury (2006, 340) predicts that this activity will lead to "lower cash inflow over the life of the sales." In summary, REM should result in lower future operating performance when the costs of REM are realized. Thus, we test the following hypothesis (stated in the alternative form):

Hypothesis 1 (H1). *Income-increasing REM is negatively associated with firms' future operating performance.*

A critical assumption underlying this hypothesis is that the predicted REM costs are observable. A substantial literature predicts that REM will generate costs (Cohen and Zarowin 2010; Roychowdhury 2006; Kothari et al. 2016). Thus, this idea is not controversial. The concern is whether future costs will reduce earnings enough to be observable to researchers. For example, the future costs may be too small to affect earnings, or additional REM may offset the decline in earnings from prior period's REM. Consequently, we perform additional tests that relax the assumption that REM will generate costs of sufficient magnitude to reduce future earnings. We subsequently examine the influence of REM on earnings persistence, perform simulation analyses, and consider how the metrics perform across firm life-cycles.

Hypothesis 2

If researchers were able to observe REM, then it is likely that they would find evidence of a negative relation between REM and future performance. However, since researchers cannot directly observe REM, they must instead infer its existence from observable behavior like abnormal changes in discretionary expenses, production, and cash flows. Yet, managers' incentives to inflate earnings are not the only influence on these operating and investing decisions. For example, prior research provides ample examples of situations where a firm's strategy may influence operating and investing decisions (Curtis et al. 2020; Dickinson 2011; Porter 1980; Srivastava 2019). Consequently, deviations in operating and investing decisions from industry peers may not reflect earnings manipulation. Instead, these deviations may be part of the firm's distinctive competitive strategy (Hirshleifer et al. 2018; Srivastava 2019). If so, researchers may reach incorrect conclusions if they assume that these actions reveal earnings manipulation when they actually reflect strategic value-enhancing decisions (Bowen et al. 2008; Srivastava 2019).

Our identification strategy relies on the idea that although there are countless economic factors that create viable alternative explanations for abnormal activities aside from earnings manipulation, there are fewer alternative economic explanations that can explain the combined income-increasing use of these activities. For example, it is difficult to determine whether cuts to R&D reflect manipulation or a value-enhancing response to the firm's current opportunities. One simple way to reduce misclassification is to consider other potential REM activity levels concurrently. For example, if R&D cuts are accompanied by decreased production, then it is less likely that REM is the cause of the changes and more likely that the changes reflect adjustments to strategy since the cuts to R&D are not accompanied by income-increasing activities in the other REM metrics.

Prior research also highlights situations where REM is more likely to be employed and thus situations where abnormal activities are more likely to reflect earnings manipulation. Huang et al. (2020) rely on an exogenous shock—an unanticipated court ruling that reduced litigation risk for firms headquartered in the Ninth Circuit—to examine the influence of litigation risk on REM. Their evidence is consistent with litigation risk deterring REM by constraining managers' ability to issue misleading disclosures. That is, their evidence suggests that managers use REM to corroborate misleading disclosures and that when litigation risk increases, misleading disclosure declines and so does the need to use REM to corroborate it. Based on their evidence, we expect that abnormal operating activities are more likely to reflect REM than strategy when litigation risk is low. A substantial literature also finds that managers' accrual and real earnings management decisions are related (Badertscher 2011; Fields et al. 2001; Zang 2012) and that the marginal cost of accrual-based earnings management increases with past earnings manipulation (Barton and Simko 2002; Laux 2014). As a result, we expect that managers are more likely to use REM when they have less flexibility to use accrual-based manipulation due to aggressive reporting choices in the recent past.

In summary, considering the combined use of multiple REM tools and situations where abnormal operating activities are more likely to reflect REM should improve our ability to identify REM. Clearly, this approach will not perfectly classify abnormal activities, but it may be useful if it (i) improves classification of abnormal activities (ii) relies on different limiting assumptions than other recent improvements, and (iii) provides a simple approach that can be broadly applied. These ideas lead to our second hypothesis (stated in the alternative form):

> Hypothesis 2 (H2). *A variable that reflects the combined use of the individual REM metrics and situations where abnormal operating activities are more likely to reflect earnings manipulation is more consistently negatively associated with future operating performance than the standard REM metrics.*

Out of necessity, researchers desiring to measure earnings manipulation make numerous assumptions. A critical assumption underlying our approach is that managers implement a REM plan most often using multiple tactics (e.g., overproduction and cuts to discretionary expenses). To the extent that managers use one tactic or the other, but refrain from concurrent use, our approach will be less beneficial. Prior research suggests that managers utilize multiple REM tactics to increase income. For example, Roychowdhury (2006, 338) posits that when managers undertake REM, they "engage in a range of activities" and researchers often predict that incentives will influence multiple REM activities concurrently (Badertscher 2011; Chan et al. 2015; Cheng et al. 2016; Zang 2012). Also, consistent with our H1 tests, we perform additional analyses that relax the assumption that REM will generate costs of sufficient magnitude to reduce future earnings. A second critical assumption is that abnormal activities that reflect strategy will either associate positively with future performance or not influence future performance. To the extent that managers implement poor strategy choices that generate future performance costs, we would not be able to distinguish poor strategy from earnings manipulation in our future performance tests. We note, however, that the simulation tests are less influenced by this concern. Moreover, we do not expect that firms implement value-decreasing strategies on average.

3. Data, variable definitions, and descriptive statistics

Sample selection

To test our hypotheses, we construct a broad sample of firms with sufficient data to calculate the earnings manipulation, future operating performance, and control variables using data from

TABLE 1
Sample selection

Sample selection criteria	Number of observations remaining in sample
Firm-years on Compustat, 1989–2015	259,138
Exclude observations with accounting changes, merger or acquisition activity, or discontinued operations	246,173
Exclude observations in regulated industries (2-digit SIC 60–63; 40–49)	203,187
Exclude observations without data required to calculate variables	55,804
Require 10 observations per industry-year for calculating REM metrics	54,017
Exclude singleton observations	53,180

Compustat and CRSP. The sample period begins in 1989 since we require statement of cash flow data and ends in 2015 to accommodate the measurement of future operating performance. We exclude observations with accounting changes, merger or acquisition activity, or discontinued operations, following McNichols (2002), and we exclude regulated industries (e.g., financial institutions and utilities) since earnings manipulation models are not well suited for these firms (Dechow et al. 2011).[10] We require firms to have non-missing data for variables used in our analyses. We also require firms to have 10 observations within a Fama and French 48 industry-year to allow sufficient observations for estimating normal activity levels. Finally, we exclude observations that enter the data set only once to avoid any bias that comes when there are singleton observations and fixed effects are nested within the groups used to cluster standard errors (Correia 2015). These requirements yield a final sample of 53,180 firm-year observations from 1989 to 2015. Table 1 outlines these sample selection procedures.

Variable definitions

Earnings manipulation

We use three empirical proxies to measure the standard REM activities: (i) cutting discretionary expenses (to increase income), (ii) extending credit to more risky customers (to increase sales), and (iii) overproducing inventory (to reduce per-unit costs) (Cohen and Zarowin 2010; Gunny 2010; Roychowdhury 2006). Our empirical models and estimation approach are based on Roychowdhury (2006). We estimate the REM models by Fama and French (1997) industry and year. We measure expected discretionary expenses with the following model:

$$\frac{Exp_t}{AT} = \alpha_0 + \beta_0 \frac{1}{AT} + \beta_1 \frac{Sales_{t-1}}{AT} + \varepsilon_t. \tag{1}$$

Exp is discretionary expenses calculated as the sum of R&D expenditures and SG&A in year t. *Sales* is the company's sales and *AT* measures the company's average total assets for year t. The residual from equation (1) measures abnormal discretionary expenses, where abnormally low discretionary expenses are consistent with income-increasing REM. Therefore, we multiply the error term by negative one so that it is increasing in the extent of REM (*AbnLowExp*).

10. Specifically, we exclude firm years where ACCHG_FN, AQA_FN or DO_FN are not blank. We identify regulated industries as SIC codes between 4000 and 4900 and between 6000 and 6399. We exclude discontinued operations per McNichols (2002), but acknowledge that discontinued operations could be used to manage earnings as evidenced in Barua et al. (2010). Barua et al. (2010), however, find that earnings management via discontinued operations has declined since the passage of SFAS 144, effective for fiscal years beginning after December 15, 2001, suggesting that this form of earnings management is less prevalent in recent periods.

We measure expected cash flows (given the level of sales) with the following model:

$$\frac{CFO_t}{AT} = \alpha_0 + \beta_0 \frac{1}{AT} + \beta_1 \frac{Sales_t}{AT} + \beta_2 \frac{\Delta Sales_t}{AT} + \varepsilon_t. \tag{2}$$

CFO is operating cash flows. The residual from equation (2) reflects abnormal cash flows, where lower-than-expected cash flows are the result of selling more product to customers on credit than in cash. Thus, we multiply the error term by negative one so that it is increasing in REM (*AbnLowCFO*).

We estimate normal production with the following model:

$$\frac{PROD_t}{AT} = \alpha_0 + \beta_0 \frac{1}{AT} + \beta_1 \frac{Sales_t}{AT} + \beta_2 \frac{\Delta Sales_t}{AT} + \beta_3 \frac{\Delta Sales_{t-1}}{AT} + \varepsilon_q. \tag{3}$$

We define production costs (*PROD*) as costs of goods sold plus the change in inventory. The residual from this model serves as our measure of REM from over production (*AbnHighProd*). We also calculate two summary metrics from the three standard individual metrics. *REM SUM1* is the sum of *AbnLowExp* and *AbnHighProd* and *REM SUM2* is the sum of *AbnLowExp* and *AbnLowCFO*.

We use the three individual earnings manipulation metrics, litigation risk, and balance-sheet bloat as inputs into a PCA. We measure litigation risk using two litigation-related variables. The first is an indicator for firms domiciled in the Ninth District prior to the famous 1999 *Silicon Graphics Inc.* US Circuit Court ruling (*LitigousCircuit*). The second is *LitigiousIndustry* an indicator for firms operating in highly litigious industries. Our measure for balance sheet bloat—*PastAM*—is an indicator variable set to one for firm-years with net operating assets above their industry median for that year, and zero otherwise.

We use PCA rather than a factor analysis because the assumptions underlying PCA are best suited for our setting.[11] PCA assumes that the resulting components are weighted combinations of the input variables. The benefit of this approach is that it allows for differing weights on each metric and creates a summary variable that is larger when multiple metrics are high. In contrast, factor analysis assumes that an underlying factor (e.g., earnings manipulation incentives) is a primary antecedent of each of the standard metrics and that the resulting factor reflects the common variation underlying each metric (rather than total variation). We view this approach to be less helpful in this setting because a primary concern identified by prior research is that the standard metrics are influenced by the firm's competitive strategy. Consequently, common variation might be explained by earnings manipulation incentives or competitive strategy. For these reasons, we view PCA as the approach most likely to create an improved REM measure.[12]

We retain the first PCA component, which has an eigenvalue of 1.609 and an intuitive factor pattern (*REMComp*). We retain one component rather than multiple components because using

11. See Allee et al. (2022) for an overview of the differences between PCA and factor analysis.
12. Like OLS, PCA relies on several assumptions about data inputs. Ideally, input variables would be continuous, normally distributed variables. As a result, there is considerable debate about the use of ordinal input variables with some viewing the approach as appropriate and others disagreeing (see https://www.theanalysisfactor.com/principal-component-analysis-for-ordinal-scale-items/ for a useful discussion). This issue is less problematic for PCA than factor analysis as conceptually factor analysis assumes that a *continuous* latent factor is the underlying cause of the observed weightings. With PCA there is no conceptual problem; rather the problem only stems from the correlation matrix, which is the default option in most statistical packages (Allee et al. 2022) and relies on the assumptions that variables are normally distributed. We use the standard PCA approach that uses the correlation matrix as it is the easiest to implement. To ensure this issue does not change inferences from our analyses, we replicate Table 4 after estimating our PCA with polychoric correlations (results not tabulated); the sign and significance of all of the results are the same as those reported in Table 4.

TABLE 2
Descriptive statistics

Panel A: Sample descriptive statistics

Variable	N	Mean	Median	SD
Future ROA_{t+1}	53,180	−0.030	0.035	0.339
Future $ROA3_{t+1,t+3}$	53,180	−0.034	0.030	0.313
Future CFO_{t+1}	53,180	0.044	0.079	0.211
Future $CFO3_{t+1,t+3}$	53,180	0.042	0.077	0.198
Restate	35,808	0.112	0.000	0.316
AbnLowExp	53,180	0.000	0.024	0.232
AbnLowCFO	53,180	0.000	−0.006	0.131
AbnHighProd	53,180	0.000	0.002	0.222
REMComp	53,180	0.000	0.112	1.000
REM SUM1	53,180	0.000	0.032	0.399
REM SUM2	53,180	0.000	0.018	0.232
Market Cap (untransformed)	53,180	2,099.95	182.98	7,294.96
Total Assets (untransformed)	53,180	1,643.86	173.32	5,235.00
Accruals Comp	53,180	0.000	0.026	1.000
ROA	53,180	−0.021	0.037	0.218
AbnormalReturn	53,180	0.074	−0.062	0.758
SalesGrowth	53,180	0.054	−0.003	0.440
MB	53,180	2.947	1.964	4.420
Size	53,180	5.262	5.155	2.121
NumAnalysts	53,180	0.967	0.693	1.078
BigNAuditor	53,180	0.816	1.000	0.387
RDMissing	53,180	0.324	0.000	0.468
Sales	53,180	1.248	1.112	0.823
Sales Change	53,180	0.086	0.068	0.289
CFO	53,180	0.047	0.079	0.177
Production Costs (*PROD*)	53,180	0.855	0.696	0.699
Discretionary Expenses (*EXP*)	53,180	0.384	0.319	0.308

Panel B: Rotated factor pattern for the REM components

Variable	*REMComp*
AbnLowExp	0.828
AbnHighProd	0.874
AbnLowCFO	0.147
LitigiousIndustry	−0.181
LitigiousCircuit	−0.168
PastAM	0.276
Eigenvalue	1.609
Proportion of variance explained	0.268

Notes: Although we rank variables as described in Appendix 2, we tabulate untransformed versions in this table for ease of interpretation. The sample consists of 53,180 annual observations from 1989 to 2015. We provide variable definitions in Appendix 2. In panel B, we use the three individual earnings manipulation metrics and other signals as inputs into a PCA.

one variable to measure the construct of REM is likely more useful than using multiple variables particularly when REM is used as a dependent variable. Also, using multiple variables could lead to “overfitting” where a high amount of total variation is explained, making it more likely that the multiple PCA variables also include firm strategy rather than REM.

Panel B of Table 2 indicates that all three of the standard metrics load positively on the component. This evidence suggests that the component is higher in the presence of the combined use of multiple REM tools. In addition, both litigation variables exhibit negative loadings consistent with the component decreasing when litigation risk increases. Also, as expected, prior use of accrual-based earnings management (as evidenced by balance sheet bloat) increases the component score. For each firm-year, we retain the PCA component score, which we use in subsequent analysis as our measure of REM. When used in regression analyses, we decile rank by year (from zero to one) the earnings manipulation variables.[13]

Future operating performance

We measure future operating performance using future earnings and cash flows. We define ROA as net income scaled by average total assets. Ex ante, it is unclear when the effect of suboptimal reporting decisions will occur. Prior research examines earnings management consequences beginning the first year following the suspect reporting period and up to three years later (Dechow and Dichev 2002; Bowen et al. 2008; Dechow et al. 2012; Demerjian et al. 2020). Therefore, we measure *FutureROA* as ROA in year $t + 1$ and *FutureROA3* as average ROA over years $t + 1$ to $t + 3$. We define CFO as operating cash flows scaled by average total assets. Similarly, we measure *FutureCFO* as CFO in year $t + 1$ and *FutureCFO3* as average CFO over years $t + 1$ to $t + 3$. We decile rank the future operating performance variables by year, where ranks range from zero to one, when used in regression analyses.

Restatements

Restate is an indicator variable coded one in year t where the firm subsequently restates year t financial information. We identify restatements from Hennes et al. (2008) for 1997–2006 and from the Audit Analytics database for 2007 onward. Appendix 2 provides details on restatement measurement.

Main control variables

Fields et al. (2001) and Zang (2012) note that managers use both accrual and real activities management together, and that failure to include both types of earnings manipulation in models can lead to inaccurate conclusions. Thus, we include a control for accrual-based earnings manipulation in our regressions.[14] Because there are numerous approaches for measuring accrual-based manipulation, we calculate commonly used approaches and use them as inputs into a PCA, which indicates one accrual-based earnings manipulation component (*Accruals Comp*), which we use as our control for accrual-based earnings manipulation. Appendix 2 provides additional estimation details.

Prior research examines the association between current period earnings characteristics, including earnings management, and future operating performance (Bowen et al. 2008; Core et al. 1999; Demerjian et al. 2013; Demerjian et al. 2020). The control variables used in our future performance tests follow from these studies. Specifically, we control for the firm's current period performance (*ROA*, *AbnormalReturn*), growth (*SalesGrowth*, *MB*), and size (*Size*) (see Bowen et al. 2008; Core et al. 1999; Demerjian et al. 2013; Demerjian et al. 2020). In addition, we include a control for the number of analysts following the firm (*NumAnalysts*), as firms with greater analyst following experience greater investor recognition, which increases the likelihood that earnings manipulation is detected (Beneish 1997; Dechow et al. 2011). Following Becker

13. We winsorize the input variables to the REM models before estimating equations (1)–(3) and before applying sample selection criteria. We estimate REM models after applying sample selection criteria. There is no post-estimation winsorization of the REM measures as we decile rank the measures before using them in regressions.

14. Our results are not sensitive to this decision and remain when we exclude the accrual component and all controls variables (results not tabulated). See the "Calibration tests" section for a discussion of robustness tests.

et al. (1998), we also control for whether the firm is audited by a Big N audit firm (*BigNAuditor*), which may influence firms' propensities to engage in earnings manipulation. Finally, following Koh and Reeb (2015), we replace missing values of R&D expenditures with zero and include an indicator for firms with missing R&D (*RDMissing*).

Additional variables used in Vorst (2016) replication

Vorst (2016) measures future operating performance using industry-adjusted measures. In our replication, we also use industry-adjusted measures.[15] *AbnROA* is earnings before extraordinary items (IB), scaled by lagged assets and adjusted for the industry-year median, where industries are defined using 2-digit SIC codes following Vorst (2016). Similarly, *AbnCFO* is CFO scaled by lagged assets and adjusted for the industry-year median. Vorst's (2016) earnings manipulation variables are reversing cuts in R&D and reversing cuts in SG&A. *Reversing R&D cut* is an indicator variable set equal to one when Abnormal R&D in year t is in the bottom quintile and there is a reversal in year $t + 1$, zero otherwise. Similarly, *Reversing SG&A cut* is an indicator variable set equal to one when Abnormal SG&A in year t is in the bottom quintile and there is a reversal in year $t + 1$, zero otherwise. Vorst (2016) includes controls for non-reversing cuts and four control variables (*BTM*, *Size*, *Z-Score*, and *Return*). We provide definitions for these variables in Appendix 2.

Descriptive statistics

Table 2 provides descriptive statistics. Sample firms are large (mean total assets of about 1.6 billion and mean market capitalization of about 2.1 billion), profitable (median *ROA* is 3.7%), and covered by one analyst (*NumAnalysts*). About 82% of sample firms are audited by a Big N auditor (*BigNAuditor*) and about 11% of firm-years are characterized by a restatement (*Restate*). Mean earnings manipulation is zero, as expected, since these measures are residuals from regression models or a principal component, which is standardized to have a mean of zero (*AbnLowExp*, *AbnLowCFO*, *AbnHighProd*, *REMComp*).

4. Empirical design and results

Future operating performance consequences

We design our tests to overcome two challenges in providing legitimate evidence: (i) correlated omitted variables influence both earnings manipulation and also future operating performance and (ii) the consequences of earnings manipulation might not be isolated to the first year following the manipulation. First, to reduce concerns over correlated omitted variables, we include firm fixed effects and numerous time-varying control variables.[16] Second, because the consequences of earnings manipulation can occur over multiple years, we examine future operating performance over both one and three subsequent years.

To examine the relation between the earnings manipulation metrics and future operating performance, we estimate the following model:

15. Moreover, if we use industry-adjusted performance measures in our main results in Tables 3 and 4, our inferences are unchanged (results not tabulated). We refrain from using industry-adjusted measures in our main analysis due to concerns raised by Gormley and Matsa (2014) over industry-adjusting dependent variables.

16. In results not tabulated, we explore several alternative specifications. Our Table 4 inferences are unchanged if we exclude firm and year fixed effects. When we do not include firm and year fixed effects in Table 3, panel B, *REM SUM1* is positively associated with future performance and negatively associated with restatements while inferences for *REM SUM2* are unchanged. The results in Table 3, panel A, excluding firm and year fixed effects are similar, except that *AbnLowExp* has a negative association with future ROA and future cash flows.

$$FuturePerformance_{t+1,t+x} = \alpha + \alpha_1 AbnRealEarningsManagment_t + \alpha_2 Accruals\ Comp_t + \alpha_3 ROA_t + \alpha_4 AbnormalReturn_t + \alpha_5 SalesGrowth_t + \alpha_6 MB_t + \alpha_7 Size_t + \alpha_8 NumAnalysts_t + \alpha_9 BigNAuditor_t + \alpha_{10} RDMissing_t + \sum_{i=1 \text{to} j} \beta_i Firm_i + \sum_{i=1 \text{to} j} \beta_i Year_i + \varepsilon_{t+1,t+x}. \quad (4)$$

We measure future operating performance with future earnings (*FutureROA*, *FutureROA3*), future operating cash flows (*FutureCFO*, *FutureCFO3*) and subsequent restatements of year *t* financial information (*Restate*). We measure *AbnRealEarningsManagement* using the individual metrics (*AbnLowExp*, *AbnLowCFO*, and *AbnHighProd*), summary metrics derived from the individual metrics *(REM SUM1* and *REM SUM2*) or the component (*REMComp*). To control for economy-wide fluctuations in performance, we include year fixed effects (see Bowen et al. 2008; Core et al. 1999). Finally, we include firm fixed effects to control for all other persistent firm characteristics.[17] For equation (4), we report robust standard errors and cluster by firm. Even though we measure *FutureROA3* and *FutureCFO3* over multiple years, the firm fixed effects and firm-clustered standard errors account for serial correlation and heteroscedasticity in the error term, thereby providing robust statistical inferences (Cameron and Miller 2015; Petersen 2009).[18]

To test H1 that the REM metrics associate negatively with future operating performance, Table 3 reports the results from estimating equation (4) using the individual metrics (*AbnLowExp*, *AbnLowCFO*, and *AbnHighProd*) and the corresponding summary metrics (*REM SUM1*, *REM SUM2*). Panel A does not reveal significant associations between *AbnLowExp* and future operating performance. In contrast, *AbnLowCFO* and *AbnHighProd* associate negatively with future ROA (*FutureROA*, *FutureROA3*) and future CFO (*FutureCFO*, *FutureCFO3*). In addition, *AbnLowCFO* is associated with an elevated risk of restatement.[19] For brevity, we include all three metrics concurrently, but conclusions are similar if we only include one metric at a time (see additional discussion in the "Calibration tests" section). In panel B, we observe that *REM SUM1* associates positively with future one-year forward ROA and negatively with restatements and future cash flows. Thus, *REM SUM1* does not appear to measure earnings manipulation as it associates with superior future earnings performance and reduced restatement rates. *REM SUM2*, however, negatively associates with future performance and positively associates with restatements.

Table 4 reports the results from estimating equation (4) using the component (*REMComp*) and presents evidence on H2. The results suggest that the component associates negatively with future ROA (*FutureROA*, *FutureROA3*) and future CFO (*FutureCFO*, *FutureCFO3*). In addition, it associates positively with *Restate*. Finally, in an untabulated analysis, we find similar results as those reported in Tables 3 and 4 when we examine future earnings before discretionary expenses. This evidence suggests that negative future operating performance costs of REM do not come solely in the form of doubling up on future discretionary expenses, but also in the form of incremental costs. Overall, *REMComp* performs consistently better than *AbnLowExp, AbnHighProd*, and *REM SUM1* and performs similarly to *AbnLowCFO* and *REM SUM2*.

17. In untabulated analyses discussed in the "Calibration tests" section, we substitute industry fixed effects for firm fixed effects and find that inferences are unchanged.
18. As a practical matter, our results are not sensitive to clustering by year or not clustering standard errors. As discussed in the "Calibration tests" section, we also estimate Fama and MacBeth annual regressions, which yield similar results (results not tabulated).
19. In Table 3, column (5), we estimate restatements using a logit model and excluding firm fixed effects. In Table 3, column (6), we estimate the restatement model using a conditional logistic regression because Allison (2005, 2009) provides evidence that the conditional logistic model estimated with firm fixed effects is least susceptible to omitted variable bias and produces consistent estimates. Note that the sample size is reduced because firm fixed effects models with binary dependent variables require both outcomes of the dependent variable for estimation of coefficients. Thus, the model is not estimable for firms that have never had a restatement.

TABLE 3
Abnormal real earnings management metrics and future operating performance

$$FuturePerformance_{t+1,t+x} = \alpha + \alpha_1 AbnRealEarningsManagment_t + \alpha_2 Accruals\ Comp_t + \alpha_3 ROA_t + \alpha_4 AbnormalReturn_t + \alpha_5 Sales\ Growth_t + \alpha_6 MB_t + \alpha_7 Size_t + \alpha_8 NumAnalysts_t + \alpha_9 BigNAuditor_t + \alpha_{10} RDMissing_t + \sum_{i=1toj} \beta_i Firm_i + \sum_{i=1toj} \beta_i Year_i + \varepsilon_{t+1,t+x}$$

Panel A: Individual abnormal real earnings management metrics and future operating performance

	Dependent variable					
	FutureROA$_{t+1}$ (1)	*FutureROA3*$_{t+1,t+3}$ (2)	*FutureCFO*$_{t+1}$ (3)	*FutureCFO3*$_{t+1,t+3}$ (4)	*Restate*(logit) (5)	*Restate*(clogit) (6)
AbnLowExp	**−0.009**	**−0.003**	**−0.012**	**−0.007**	**−0.012**	**0.080**
	(−1.102)	**(−0.313)**	**(−1.390)**	**(−0.865)**	**(−0.103)**	**(0.421)**
AbnHighProd	**−0.050*****	**−0.047*****	**−0.068*****	**−0.051*****	**0.069**	**0.050**
	(−6.201)	**(−5.403)**	**(−8.544)**	**(−6.653)**	**(0.619)**	**(0.302)**
AbnLowCFO	**−0.117*****	**−0.074*****	**−0.062*****	**−0.036*****	**0.549*****	**0.239***
	(−18.702)	**(−12.150)**	**(−9.201)**	**(−6.096)**	**(5.606)**	**(1.942)**
Accruals Comp	0.014***	−0.004	0.023***	0.013***	−0.379***	−0.218***
	(3.282)	(−0.941)	(5.320)	(3.475)	(−5.819)	(−2.700)
ROA	0.298***	0.168***	0.235***	0.169***	0.091	−0.646***
	(23.243)	(13.043)	(20.011)	(14.801)	(0.691)	(−3.128)
AbnormalReturn	0.053***	0.026***	0.014***	0.008***	0.043**	0.054**
	(27.019)	(16.569)	(8.878)	(6.572)	(2.205)	(2.171)
SalesGrowth	0.017***	0.008***	0.004	0.005**	−0.037	−0.013
	(5.634)	(3.034)	(1.323)	(1.969)	(−0.978)	(−0.232)
MB	0.003***	0.002***	0.002***	0.001***	−0.002	0.002
	(8.115)	(5.103)	(5.350)	(4.099)	(−0.591)	(0.493)
Size	−0.059***	−0.079***	−0.008**	−0.012***	0.129***	0.381***
	(−19.032)	(−20.318)	(−2.459)	(−3.199)	(7.119)	(5.147)
NumAnalysts	0.034***	0.027***	0.024***	0.027***	−0.079***	−0.099
	(10.695)	(7.285)	(7.727)	(7.393)	(−2.684)	(−1.461)

(The table is continued on the next page.)

TABLE 3 (continued)

$$FuturePerformance_{t+1,t+x} = \alpha + \alpha_1 AbnRealEarningsManagment_t + \alpha_2 Accruals\ Comp_t + \alpha_3 ROA_t + \alpha_4 AbnormalReturn_t + \alpha_5 Sales\ Growth_t + \alpha_6 MB_t + \alpha_7 Size_t + \alpha_8 NumAnalysts_t + \alpha_9 BigNAuditor_t + \alpha_{10} RDMissing_t + \sum_{i=1toj} \beta_i Firm_i + \sum_{i=1toj} \beta_i Year_i + \varepsilon_{t+1,t+x}$$

Panel A: Individual abnormal real earnings management metrics and future operating performance

	Dependent variable					
	FutureROA$_{t+1}$ (1)	*FutureROA3*$_{t+1,t+3}$ (2)	*FutureCFO*$_{t+1}$ (3)	*FutureCFO3*$_{t+1,t+3}$ (4)	*Restate*(logit) (5)	*Restate*(clogit) (6)
BigNAuditor	−0.001	−0.000	−0.009	−0.010	0.093	0.244*
	(−0.123)	(−0.042)	(−1.248)	(−1.249)	(1.214)	(1.723)
RDMissing	0.011	0.011	0.017**	0.016*	−0.136**	0.033
	(1.310)	(1.104)	(2.113)	(1.653)	(−2.177)	(0.175)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	No	Yes
N	53,180	53,180	53,180	53,180	35,808	14,995
Adj./Pseudo R^2	0.529	0.615	0.514	0.671	0.0211	0.0237

Panel B: Combined metrics based on the individual abnormal real earnings management metrics and future operating performance

	Dependent variable					
	FutureROA$_{t+1}$ (1)	*FutureROA3*$_{t+1,t+3}$ (2)	*Future CFO*$_{t+1}$ (3)	*Future CFO3*$_{t+1,t+3}$ (4)	*Restate*(logit) (5)	*Restate*(clogit) (6)
REM SUM1	**0.030*****	**0.008**	**−0.027****	**−0.029*****	**−0.519*****	**−0.124**
	(2.798)	**(0.718)**	**(−2.486)**	**(−2.861)**	**(−3.410)**	**(−0.557)**
REM SUM2	**−0.112*****	**−0.071*****	**−0.052*****	**−0.026*****	**0.676*****	**0.402****
	(−11.463)	**(−7.579)**	**(−5.087)**	**(−2.952)**	**(4.396)**	**(2.038)**
Accruals Comp	−0.008**	−0.019***	0.009**	0.004	−0.306***	−0.208***
	(−2.020)	(−4.815)	(1.974)	(0.981)	(−4.887)	(−2.651)

(The table is continued on the next page.)

TABLE 3 (continued)

Panel B: Combined metrics based on the individual abnormal real earnings management metrics and future operating performance

	Dependent variable					
	FutureROA$_{t+1}$ (1)	*FutureROA3*$_{t+1,t+3}$ (2)	*Future CFO*$_{t+1}$ (3)	*Future CFO3*$_{t+1,t+3}$ (4)	*Restate*(logit) (5)	*Restate*(clogit) (6)
ROA	0.376***	0.223***	0.288***	0.203***	−0.216**	−0.746***
	(29.534)	(17.328)	(24.139)	(17.423)	(−1.966)	(−3.836)
AbnormalReturn	0.054***	0.026***	0.015***	0.009***	0.044**	0.053**
	(27.194)	(16.734)	(9.148)	(6.815)	(2.188)	(2.147)
SalesGrowth	0.009***	0.003	−0.001	0.001	−0.018	0.004
	(3.261)	(1.060)	(−0.451)	(0.656)	(−0.508)	(0.067)
MB	0.003***	0.002***	0.002***	0.001***	−0.004	0.002
	(8.145)	(5.134)	(5.342)	(4.097)	(−1.068)	(0.482)
Size	−0.059***	−0.079***	−0.008**	−0.012***	0.131***	0.373***
	(−18.648)	(−20.157)	(−2.523)	(−3.263)	(7.332)	(5.049)
NumAnalysts	0.037***	0.029***	0.026***	0.028***	−0.090***	−0.099
	(11.347)	(7.752)	(8.114)	(7.627)	(−3.037)	(−1.473)
BigNAuditor	−0.006	−0.004	−0.012*	−0.012	0.128*	0.248*
	(−0.809)	(−0.453)	(−1.680)	(−1.490)	(1.656)	(1.752)
RDMissing	0.014	0.013	0.019**	0.017*	−0.128**	0.025
	(1.626)	(1.306)	(2.268)	(1.748)	(−2.040)	(0.130)
Year FE	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Firm FE	*Yes*	*Yes*	*Yes*	*Yes*	*No*	*Yes*
N	53,180	53,180	53,180	53,180	35,808	14,995
Adj./Pseudo R^2	0.523	0.612	0.511	0.670	0.0190	0.0238

Notes: We present t-statistics or z-statistics below the coefficients. See Appendix 2 for variable definitions. ***, **, and * denote two-tailed p-values of less than 0.01, 0.05, and 0.10, respectively, based on robust standard errors that are clustered by firm. Bold font denotes an earnings manipulation variable.

TABLE 4
REM component

$$FuturePerformance_{t+1,t+x} = \alpha + \alpha_1 REMComp_t + \alpha_2 AccrualsComp_t + \alpha_3 ROA_t + \alpha_4 AbnormalReturn_t + \alpha_5 SalesGrowth_t + \alpha_6 MB_t + \alpha_7 Size_t + \alpha_8 NumAnalysts_t + \alpha_9 BigNAuditor_t + \alpha_{10} RDMissing_t + \sum_{i=1toj} \beta_i Firm_i + \sum_{i=1toj} \beta_i Year_i + \varepsilon_{t+1,t+x}$$

	Dependent variable					
	FutureROA$_{t+1}$ (1)	*Future ROA3*$_{t+1,t+3}$ (2)	*FutureCFO*$_{t+1}$ (3)	*Future CFO3*$_{t+1,t+3}$ (4)	*Restate*(logit) (6)	*Restate*(clogit) (7)
REMComp	**−0.085*****	**−0.071*****	**−0.073*****	**−0.050*****	**0.093**	**0.331****
	(−10.677)	**(−7.851)**	**(−9.286)**	**(−6.085)**	**(1.068)**	**(2.011)**
Accruals Comp	−0.019***	−0.025***	0.004	0.002	−0.223***	−0.174**
	(−4.821)	(−6.623)	(1.081)	(0.612)	(−3.765)	(−2.344)
ROA	0.380***	0.223***	0.287***	0.201***	−0.278**	−0.756***
	(29.533)	(17.251)	(24.009)	(17.199)	(−2.575)	(−3.870)
AbnormalReturn	0.054***	0.027***	0.015***	0.009***	0.041**	0.051**
	(27.440)	(16.922)	(9.274)	(6.881)	(2.038)	(2.081)
SalesGrowth	0.013***	0.005**	0.001	0.003	−0.028	−0.008
	(4.450)	(2.033)	(0.500)	(1.376)	(−0.789)	(−0.146)
MB	0.003***	0.002***	0.002***	0.001***	−0.004	0.002
	(8.175)	(5.130)	(5.350)	(4.103)	(−1.147)	(0.494)
Size	−0.058***	−0.078***	−0.008**	−0.012***	0.134***	0.361***
	(−17.995)	(−19.695)	(−2.252)	(−3.089)	(7.474)	(4.864)
NumAnalysts	0.037***	0.029***	0.026***	0.028***	−0.097***	−0.099
	(11.386)	(7.775)	(8.150)	(7.648)	(−3.293)	(−1.471)
BigNAuditor	−0.007	−0.004	−0.012*	−0.012	0.123	0.251*
	(−0.905)	(−0.521)	(−1.702)	(−1.497)	(1.589)	(1.767)
RDMissing	0.014*	0.014	0.019**	0.017*	−0.121*	0.031
	(1.668)	(1.355)	(2.262)	(1.743)	(−1.924)	(0.164)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	No	Yes
N	53,180	53,180	53,180	53,180	35,808	14,995
Adjusted (Pseudo) R^2	0.523	0.612	0.511	0.670	0.0179	0.0238

Notes: We present t-statistics or z-statistics below the coefficients. See Appendix 2 for variable definitions. ***, **, and * denote two-tailed p-values of less than 0.01, 0.05, and 0.10, respectively, based on robust standard errors that are clustered by firm. Bold font denotes an earnings manipulation variable.

Earnings persistence

An additional implication of earnings manipulation is a change in the intertemporal relations of cash flows, accruals, and earnings (e.g., changes in the persistence of cash flows). To examine differential persistence, we follow prior research (Casey et al. 2017; Richardson et al. 2005; Sloan 1996) and estimate equations (5) to (8) using the Fama and MacBeth approach (i.e., we estimate regressions by year) with two lags for the Newey-West standard error correction:

$$FutureROA_{t+1} = \gamma_0 + \gamma_1 CFO_t + \gamma_2 TotalAccruals_t + \nu_{t+1}. \tag{5}$$

In equation (5), γ_1 reflects the persistence of cash flows and γ_2 reflects the persistence of accruals. In a similar vein, we estimate the following model:

TABLE 5
The influence of REM on earnings persistence

$$FutureROA_{t+1} = \gamma_0 + \gamma_1 CFO_t + \gamma_2 TotalAccruals_t + \gamma_3 AbnRealEarningsManagment_t + \gamma_4 CFO_t \times AbnRealEarningsManagment_t + \nu_{t+1}$$

	Dependent variable = $FutureROA_{t+1}$		
	(1)	(2)	(3)
CFO	1.027***	1.063***	1.096***
	(27.384)	(27.651)	(31.809)
TotalAccruals	0.577***	0.615***	0.597***
	(15.291)	(15.697)	(15.888)
REM SUM1		−0.005	
		(−0.273)	
REM SUM2		−0.061***	
		(−3.854)	
CFO×REM SUM1		**−0.027**	
		(−0.343)	
CFO×REM SUM2		**−0.087**	
		(−1.307)	
REMComp			−0.068***
			(−10.298)
CFO×REMComp			**−0.176*** **
			(−5.377)
N	53,180	53,180	53,180
Adjusted (Pseudo) R^2	0.350	0.359	0.358

Notes: This table presents the results from an estimation of future ROA on current period CFO, total accruals, and real activities manipulation. We estimate the models using the Fama-MacBeth procedure with two lags for the Newey-West standard error correction. We present t-statistics below the coefficients. ***, **, and * denote two-tailed p-values of less than 0.01, 0.05, and 0.10, respectively. We provide variable definitions in Appendix 2. Bold font denotes an earnings manipulation variable.

$$FutureROA_{t+1} = \gamma_0 + \gamma_1 CFO_t + \gamma_2 TotalAccruals_t + \gamma_3 AbnRealEarningsManagment_t + \gamma_4 AbnRealEarningsManagment_t \times CFO_t + \nu_{t+1}. \tag{6}$$

In equation (6), the coefficient on CFO reflects the persistence of the cash portion of earnings not achieved via real activities manipulation, γ_3 reflects the persistence of REM and γ_4 reflects the differential persistence of the cash portion of earnings when REM is high. We measure *AbnRealEarningsManagement* using the simple summations of the standard metrics (*REM SUM1*, *REM SUM2*) or the component (*REMComp*). We expect to observe negative coefficients for γ_4 if the metrics measure REM and REM reduces the persistence of the cash flow portion of earnings.

We report the results from estimating equations (5) and (6) in Table 5. In the first specification, the coefficient for *CFO* is 1.027 (t = 27.38) and the coefficient for *TotalAccruals* is 0.577 (t = 15.29). The signs and magnitudes of the coefficients are similar to prior research (see Sloan 1996, table 3, panel B). Column (2) of Table 5 reports the results for *REM SUM1* and *REM SUM2*. *REM SUM1* is not associated with earnings persistence and does not influence the persistence of CFO (i.e., *REM SUM1×CFO* is not significantly different from zero). Similarly, *REM SUM2* does not influence the persistence of CFO (i.e., *REM SUM2×CFO* is not significantly different from zero). Column (3) reports the results for *REMComp*. *REMComp* is associated with decreased future earnings (i.e., *REMComp* significantly negative) and reduces the persistence of CFO (i.e., *REMComp*×*CFO* is significantly negative).

The tests reported in Table 5 provide evidence that the REM component is negatively associated with earnings persistence, as we would expect if the component measures real activities manipulation. We note that in some circumstances, an additional characteristic of a good earnings manipulation measure is a difference in the associations between the predicted level of discretionary expenditures and production cuts and their abnormal levels that allow us to measure REM. That is, a partition of operating and investing activities into economic-driven versus manipulation-driven portions should lead to variables that behave differently. In Table 6, we use equation (7) to explore how the predicted versus abnormal portions of operating activities influence earnings persistence.

$$\begin{aligned} FutureROA_{t+1} = \alpha + \alpha_1 CFO_t + \alpha_2 TotalAccruals_t + \alpha_3 ExpectedPortion_t \\ + \alpha_4 AbnRealEarningsManagment_t + \alpha_5 (CFO_t \times ExpectedPortion_t) \\ + \alpha_6 (CFO_t \times AbnRealEarningsManagment_t) + \varepsilon_{t+1}. \end{aligned} \quad (7)$$

The results reported in Table 6 suggest that the predicted and abnormal portions of the activities underlying *REM SUM1* are not associated with the persistence of the cash flow portion of earnings (i.e., *CFO*×*Expected REM SUM1* and *CFO*×*REM SUM1* are not significantly different from zero). These results again suggest that *REM SUM1* does not reflect earnings manipulation. In contrast, the predicted activities underlying *REMComp* increase CFO persistence (i.e., *CFO*×*Expected REMComp*) whereas the abnormal activities decrease CFO persistence (*CFO*×*REMComp*). We find similar results for *REM SUM2*. To the extent that different consequences of the predicted and abnormal activities are relevant dimensions of a good measure of REM, the results in Table 6 suggest that the component approach outperforms the individual metrics—notably, *REM SUM1*.

Sales growth persistence

Roychowdhury (2006, 339) notes that one particularly costly REM activity is sales manipulation which involves “accelerating the timing of sales and/or generating additional unsustainable sales through increased price discounts or more lenient credit terms.” He also notes that the increased sales volume is likely to disappear is subsequent periods. This line of reasoning suggests that a possible consequence of REM is reduced persistence of sales growth due to sales manipulation. To explore how the REM metrics influence the persistence of sales growth we estimate models similar to equations (6) and (7), but replacing *CFO* with *SalesGrowth*. If REM reduces the persistence of sales growth, then we expect to observe a negative coefficient for the interaction of REM and *SalesGrowth*. We report the results in Table 7. The first two columns present the results from estimating the equivalent of equation (6) and the last two columns present the results from estimating the equivalent of equation (7). Across all estimations, the results indicate a negative coefficient on *SalesGrowth*×*REMComp*, suggesting that *REMComp* reduces the persistence of sales growth. We find similar results for *REM SUM1*. In contrast to our other tests, we find that *REM SUM2* is not negatively associated with the persistence of sales growth (see the coefficient for *SalesGrowth*×*REM SUM2*). To the extent that reduced sales growth persistence is an expected outcome from REM, this test again corroborates the use of *REMComp* rather than *REM SUM1* and *REM SUM2*.

Simulation analysis

Prior research has also validated measures of earnings management using simulation analyses that intentionally seed errors to evaluate the power of the metrics to detect known errors (Cohen

TABLE 6
The influence of the predicted versus abnormal portions of REM on earnings persistence

$$FutureROA_{t+1} = \alpha + \alpha_1 CFO_t + \alpha_2 TotalAccruals_t + \alpha_3 ExpectedPortion_t + \alpha_4 AbnRealEarningsManagment_t + \alpha_5 CFO_t \times ExpectedPortion_t + \alpha_6 CFO_t \times AbnRealEarningsManagment_t + \varepsilon_{t+1}$$

Dependent variable = $FutureROA_{t+1}$

Variables	(1)	(2)
CFO	0.992***	0.850***
	(21.151)	(23.113)
TotalAccruals	0.589***	0.682***
	(14.788)	(17.790)
Expected REM SUM1	0.086***	
	(9.226)	
REM SUM1	−0.029*	
	(−1.714)	
Expected REM SUM2	0.013	
	(1.284)	
REM SUM2	−0.036**	
	(−2.410)	
CFO×Expected REM SUM1	−0.112	
	(−1.420)	
CFO×REM SUM1	**0.030**	
	(0.476)	
CFO×Expected REM SUM2	0.275***	
	(3.430)	
CFO×REM SUM2	**−0.120***	
	(−1.998)	
Expected REMComp		−0.086***
		(−5.438)
REMComp		−0.049***
		(−5.871)
CFO×Expected REMComp		1.099***
		(14.958)
CFO×REMComp		**−0.085***
		(−3.592)
N	53,180	53,180
Adjusted (Pseudo) R^2	0.378	0.387

Notes: This table presents the results from an estimation of future ROA on current period CFO, expected operating activities, abnormal operating activities and the interaction of CFO and expected and abnormal activities. We estimate the models using the Fama-MacBeth procedure with two lags for the Newey-West standard error correction. The expected components denoted with the prefix *Expected* are the predicted values from the applicable earnings management models. The expected component is the resulting component from a PCA that uses as its inputs *ExpectedCFO*, *ExpectedProd*, and *ExpectedDisExp*. We present t-statistics below the coefficients. ***, **, and * denote two-tailed p-values of less than 0.01, 0.05, and 0.10, respectively. We provide other variable definitions in Appendix 2. Bold font denotes an earnings manipulation variable.

et al. 2019; Kothari et al. 2005). The simulation approach takes advantage of the fact that measures of abnormal real activities are residuals from regression analyses performed on a population of firms, and thus are mean zero. If the mean residual value is significantly different from zero for a given subsample of observations, the inference is that the subsample has engaged in real earnings management, on average. The simulation approach differs from our prior tests using

TABLE 7
The influence of REM on sales growth persistence

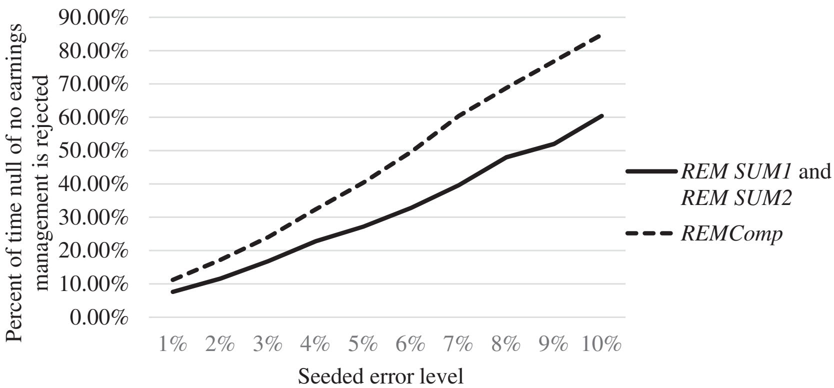
$$SalesGrowth_{t+1} = \gamma_0 + \gamma_1 SalesGrowth_t + \gamma_2 TotalAccruals_t + \gamma_3 AbnRealEarningsManagment_t + \gamma_4 SalesGrowth_t \times AbnRealEarningsManagment_t + \nu_{t+1}$$

$$SalesGrowth_{t+1} = \gamma_0 + \gamma_1 SalesGrowth_t + \gamma_2 TotalAccruals_t + \gamma_3 Expected\ Portion_t + \gamma_4 AbnRealEarningsManagment_t + \gamma_5 SalesGrowth_t \times Expected\ Portion_t + \gamma_6 SalesGrowth_t \times AbnRealEarningsManagment_t + \nu_{t+1}$$

	Dependent variable = $SalesGrowth_{t+1}$			
	(1)	(2)	(3)	(4)
SalesGrowth	0.106***	0.106***	0.042**	0.075***
	(6.891)	(7.165)	(2.061)	(5.624)
TotalAccruals	0.196***	0.163***	0.195***	0.159***
	(5.430)	(4.956)	(5.720)	(4.827)
Expected REM SUM1			−0.082***	
			(−7.496)	
REM SUM1	0.083***		0.087***	
	(6.192)		(7.448)	
Expected REM SUM2			0.033	
			(1.604)	
REM SUM2	−0.075***		−0.085***	
	(−5.161)		(−7.071)	
SalesGrowth×Expected REM SUM1			0.092***	
			(3.881)	
SalesGrowth×REM SUM1	**−0.072****		**−0.085****	
	(−2.428)		**(−2.434)**	
SalesGrowth×Expected REM SUM2			0.080***	
			(4.024)	
SalesGrowth×REM SUM2	**0.008**		**0.034**	
	(0.467)		**(1.452)**	
Expected REMComp				−0.031*
				(−1.863)
REMComp		0.029***		0.029***
		(3.535)		(3.798)
SalesGrowth×Expected REMComp				0.114***
				(4.705)
SalesGrowth×REMComp		**−0.063*****		**−0.065*****
		(−3.114)		**(−3.255)**
N	53,034	53,034	53,034	53,034
Adjusted (Pseudo) R^2	0.026	0.023	0.046	0.037

Notes: This table presents the results from an estimation of future sales growth on current period sales growth, total accruals, and real activities manipulation. We estimate the models using the Fama-MacBeth procedure with two lags for the Newey-West standard error correction. ***, **, and * denote two-tailed p-values of less than 0.01, 0.05, and 0.10, respectively. We provide variable definitions in Appendix 2. We present t-statistics below the coefficients. For this table only, sales growth is *not* industry adjusted consistent with our approach throughout of not industry-adjusting dependent variables (see footnote 15). Because we do not industry-adjust future sales growth, in this table only, we do not industry-adjust current period sales growth. In all other tables, when we use sales growth as a control variable, we do use industry-adjusted values as noted in Appendix 2. Bold font denotes an earnings manipulation variable.

Figure 1 Simulation analysis



Notes: This figure presents results from a simulation analysis. We randomly select 100 observations from our sample and seed income-increasing errors in discretionary expenses, operating cash flow and production from 1% to 10% in 1% increments. In the full sample that includes 100 observations with seeded errors and 53,080 firm-years without any seeded errors, we estimate equations (1)–(3) (the first-stage Roychowdhury models) and save the residuals from those regressions. Next, we calculate *REM SUM1**, *REM SUM2**, and *REMComp** using the residuals from those regressions. We retain only the 100 observations with seeded errors. Within that sample of 100 seeded error observations, we use a one-tailed *t*-test to determine whether *REMComp** is significantly greater than zero. We also test in that same sample of 100 observations whether both *REM SUM1** and *REM SUM2** reject the null that the mean is zero (one-tailed test at the 5% level again). For each given error level between 1% and 10%, we repeat this process 250 times and graph the percentage of the 250 times that the null is rejected for *REMComp** (dashed line in this figure) and *REM SUM1** and *REM SUM2** (the solid line in this figure).

associations with future operating performance or earnings persistence because it abstracts from assumptions about what earnings would have been absent the earnings manipulation. To the extent that our future operating performance, earnings persistence, and simulation analyses produce similar results, we consider this result to be strong evidence that the components approach is a valid way to measure REM.

Following Kothari et al. (2005), we randomly select 100 firm-year observations without replacement from our sample of firms (see Table 1). For those 100 observations, we seed income-increasing errors in *Discretionary Expenses*, *CFO*, and *Production Costs* (see Appendix 2 for variable definitions) in 1% increments from 1% to 10%. For example, the 10% seeded-error version of *Production Costs* is calculated as $Production\ Costs_{Error} = Production\ Costs + (Production\ Costs \times 0.10)$. To achieve income-increasing errors, we add errors for *Production Costs* and subtract errors for *CFO* and *Discretionary Expenses*. After seeding errors in 100 firm-years, we then perform first-stage regressions using our full sample of firm-years. We use the errors from the unseeded firm-years (*AbnLowExp*, *AbnLowCFO*, and *AbnHighProd*) and the seeded firm-years ($AbnLowExp_{Error}$, $AbnLowCFO_{Error}$, and $AbnHighProd_{Error}$) to calculate the summary metrics *REM SUM1** and *REM SUM2** and as inputs into the PCA analysis to obtain our component. We repeat the process of introducing error into 100 observations, modeling abnormal real activities and computing the REM variables (*REM SUM1**, *REM SUM2**, *RemComp**) 250 times. For each of the 250 iterations, we then retain the 100 observations with seeded errors (the error iterations). We then calculate the percentage of times (out of the 250 iterations) that a one-tailed *t*-test rejects the null hypothesis that the REM measure is mean zero at the 5% level in the error iterations and present the results in Figure 1.

Figure 1 indicates that for all seeded error levels from 1% to 10%, the PCA measure rejects the null at a greater frequency than the rejection rates for *REM SUM1* and *REM SUM2*. For seeded errors of 1% of assets, the rejection rate for *REMComp* is 11.20% versus 7.6% for the simple sums. At a 10% error level, *REMComp* is associated with a rejection rate of 84.80% versus 60.40% for the simple sums. Overall, these tests provide additional evidence that the PCA approach provides meaningful improvement over simple combinations of the individual metrics for seeded error levels between 1% and 10% of assets.

Vorst replication

The results thus far suggest that the PCA approach is better able to identify real activities manipulation than summary metrics derived from the standard REM metrics. We note, however, that there have been recent advancements in the measurement of REM. Vorst (2016) provides one salient innovation. Thus, it is important to consider whether our approach is incremental to Vorst's (2016) and also whether our method can be used in different (or additional) settings where his approach might not be appropriate. If so, this evidence would suggest that we contribute to this literature by providing researchers with a refined way to measure REM that is incremental to Vorst's (2016) approach.

We perform two tests. First, we examine how our component performs in Vorst's (2016) setting. A benefit of this test, in addition to replicating Vorst's (2016) results, is that we conduct another PCA using his alternate sampling selection procedures, which yields different component loadings. This evidence is beneficial because it allows us to investigate whether the benefit of the component approach is specific to particular component loadings or if the benefit simply stems from considering the activities concurrently (regardless of particular loadings). Stated differently, it allows us to explore whether the component approach works in a different setting.

Following Vorst (2016), we identify a sample of firm-years between 1983 and 2012 and follow his sample selection criteria. That is, we exclude regulated industries defined as 2-digit SIC codes 40–59 or 60–63. We require 15 firms in each industry-year and require non-zero R&D for the R&D sample and non-zero SG&A for the SG&A sample. Finally, we require firms to have sufficient data available to estimate the REM models, which are based on Gunny (2010) and include market capitalization and Tobin's Q (i.e., publicly traded equity). These screens yield a final sample that has (at most) 41,588 firm-years.

Using this sample, we perform a second PCA and retain the first component, which has an eigenvalue of 1.71 (see panel A, Table 8). In panels B and C of Table 8, we replicate Vorst's (2016) table 7, which examines the relation between earnings manipulation and future abnormal operating performance. We consider reversing cuts in R&D in panel B and reversing SG&A cuts in panel C. We supplement Vorst's models with our REM component. We find results similar to Vorst's (2016) that reversing cuts in R&D (panel B) and reversing cuts in SG&A (panel C) are associated with lower future abnormal ROA and CFO, and that non-reversing cuts are positively (or not) associated with future abnormal ROA and CFO. We note that the PCA measure associates negatively with future abnormal ROA and CFO. These results suggest that the PCA approach identifies earnings manipulation that is incremental to the manipulation identified by Vorst (2016).

Next, we consider how controlling for reversing and non-reversing cuts in R&D and SG&A influences our main results. That is, we control for Vorst's (2016) manipulation variables in our sample. We report the results from this analysis in panel A (R&D) and panel B (SG&A) of Table 9. First, we find that our component still associates negatively with future ROA and CFO. That is, in our setting, our results are incremental to the effect that Vorst (2016) finds in his analyses. Second, we do not observe the negative association between reversing cuts and future operating performance in our setting. This result may occur for numerous reasons. The key takeaway is that the PCA approach works in both settings and is incremental to Vorst's (2016) approach.

TABLE 8

REM components in the Vorst (2016) setting

Panel A: REM component in the Vorst (2016) setting

	REMComp
AbnLowExp	0.923
AbnHighProd	−0.284
AbnLowCFO	0.755
LitigiousIndustry	0.230
LitigiousCircuit	0.170
PastAM	−0.346
Eigenvalue	1.714
Proportion of variance explained	0.286

Panel B: Vorst (2016) settings and models (reversing R&D cuts) supplemented with the REM component

Dependent variable	$AbnROA_{t+2}$ (1)	$AbnROA_{t+3}$ (2)	$AbnROA_{t+4}$ (3)	$AbnROA_{t+5}$ (4)	$AbnCFO_{t+2}$ (5)	$AbnCFO_{t+3}$ (6)	$AbnCFO_{t+4}$ (7)	$AbnCFO_{t+5}$ (8)
REMComp	−0.079***	−0.073***	−0.072***	−0.070***	−0.061***	−0.056***	−0.054***	−0.049***
	(−7.292)	(−6.148)	(−5.745)	(−4.855)	(−7.412)	(−6.377)	(−5.933)	(−4.785)
Non-reversing R&D cut	0.009	0.015*	0.005	0.005	0.015***	0.013**	0.010	0.013*
	(1.238)	(1.718)	(0.492)	(0.423)	(2.931)	(2.362)	(1.547)	(1.848)
Reversing R&D cut	−0.139***	−0.157***	−0.190***	−0.152***	−0.099***	−0.119***	−0.120***	−0.106***
	(−7.061)	(−7.257)	(−6.827)	(−5.242)	(−9.134)	(−8.280)	(−7.292)	(−6.571)
AbnROA	0.280***	0.239***	0.220***	0.223***	0.209***	0.185***	0.171***	0.161***
	(7.263)	(5.858)	(5.229)	(4.662)	(6.933)	(6.190)	(5.821)	(4.923)
BTM	−0.000	0.000	0.010**	0.012**	−0.001	−0.001	0.001	0.003
	(−0.069)	(0.016)	(2.008)	(2.435)	(−0.544)	(−0.563)	(0.406)	(1.088)
Size	0.024***	0.028***	0.029***	0.031***	0.022***	0.023***	0.023***	0.023***
	(15.496)	(14.890)	(13.720)	(12.992)	(18.584)	(17.445)	(16.019)	(14.922)
Z-score	0.010***	0.010***	0.012***	0.011***	0.007***	0.007***	0.007***	0.007***
	(5.456)	(4.370)	(4.276)	(3.410)	(4.772)	(4.537)	(4.058)	(3.760)
Return	0.009***	0.008***	0.008**	0.005	0.007***	0.006***	0.006**	0.003
	(3.121)	(2.795)	(2.362)	(1.055)	(3.205)	(2.799)	(2.370)	(1.104)

(The table is continued on the next page.)

TABLE 8 (continued)

Panel B: Vorst (2016) settings and models (reversing R&D cuts) supplemented with the REM component

Dependent variable	$AbnROA_{t+2}$ (1)	$AbnROA_{t+3}$ (2)	$AbnROA_{t+4}$ (3)	$AbnROA_{t+5}$ (4)	$AbnCFO_{t+2}$ (5)	$AbnCFO_{t+3}$ (6)	$AbnCFO_{t+4}$ (7)	$AbnCFO_{t+5}$ (8)
N	33,493	31,345	29,288	27,400	33,457	31,307	29,253	27,369
Adj. R^2	0.254	0.212	0.195	0.172	0.314	0.269	0.233	0.217
Year FE and Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Panel C: Vorst (2016) setting and models (reversing SG&A cuts) supplemented with the REM component

Dependent variable	$AbnROA_{t+2}$ (1)	$AbnROA_{t+3}$ (2)	$AbnROA_{t+4}$ (3)	$AbnROA_{t+5}$ (4)	$AbnCFO_{t+2}$ (5)	$AbnCFO_{t+3}$ (6)	$AbnCFO_{t+4}$ (7)	$AbnCFO_{t+5}$ (8)
REMComp	−0.075***	−0.069***	−0.071***	−0.067***	−0.063***	−0.057***	−0.056***	−0.052***
	(−7.649)	(−6.470)	(−6.281)	(−5.053)	(−8.629)	(−7.256)	(−6.963)	(−5.634)
Non-reversing SG&A cut	−0.009**	−0.010*	−0.008	−0.004	−0.010***	−0.009**	−0.008**	−0.006
	(−1.977)	(−1.829)	(−1.449)	(−0.558)	(−3.051)	(−2.482)	(−2.052)	(−1.246)
Reversing SG&A cut	−0.120***	−0.097***	−0.097***	−0.101***	−0.081***	−0.089***	−0.075***	−0.076***
	(−6.906)	(−6.448)	(−4.893)	(−4.652)	(−7.679)	(−7.449)	(−6.144)	(−5.166)
AbnROA	0.277***	0.237***	0.215***	0.232***	0.212***	0.183***	0.169***	0.164***
	(7.405)	(5.745)	(5.185)	(4.916)	(7.286)	(6.258)	(6.051)	(5.151)
BTM	−0.001	−0.002	0.007*	0.010**	−0.002	−0.003	−0.001	0.001
	(−0.399)	(−0.483)	(1.793)	(2.480)	(−0.982)	(−1.477)	(−0.423)	(0.201)
Size	0.022***	0.024***	0.026***	0.028***	0.019***	0.020***	0.021***	0.021***
	(15.503)	(14.900)	(14.144)	(13.434)	(18.258)	(17.507)	(16.577)	(15.347)
Z-score	0.011***	0.012***	0.013***	0.011***	0.008***	0.008***	0.008***	0.008***
	(5.666)	(5.048)	(4.471)	(3.616)	(5.011)	(5.206)	(4.496)	(4.059)
Return	0.009***	0.007***	0.008**	0.005	0.007***	0.006***	0.005***	0.003
	(3.575)	(2.714)	(2.453)	(1.122)	(3.405)	(2.813)	(2.608)	(1.396)
N	41,588	38,835	36,236	33,848	41,544	38,789	36,193	33,810
Adj. R^2	0.260	0.208	0.189	0.163	0.304	0.256	0.225	0.203
Year FE and Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table utilizes a setting similar to Vorst (2016) and the Vorst (2016) models (see Vorst 2016, table 7) to explore how the REM components perform in his setting and after controlling for reversing R&D cuts (panel A) and reversing SG&A cuts (panel B). We present *t*-statistics below the coefficients. See Appendix 2 for variable definitions. ***, **, and * denote two-tailed *p*-values of less than 0.01, 0.05, and 0.10, respectively, based on robust standard errors that are clustered by firm.

TABLE 9
Vorst (2016) earnings manipulation variables in our setting

Panel A: Our setting and models supplemented with the Vorst (2016) earnings manipulation variables (reversing and non-reversing R&D cuts)

Dependent variable	*Future ROA*$_{t+1}$ (1)	*Future ROA3*$_{t+1,t+3}$ (2)	*Future CFO*$_{t+1}$ (3)	*Future CFO3*$_{t+1,t+3}$ (4)	*Restate*(logit) (5)	*Restate*(clogit) (6)
REMComp	−0.082***	−0.080***	−0.061***	−0.045***	0.028	0.455**
	(−7.583)	(−6.556)	(−5.788)	(−3.970)	(0.248)	(2.044)
Non-reversing R&D cut	−0.053***	−0.038***	−0.026***	−0.023***	−0.243**	−0.145
	(−9.158)	(−6.863)	(−4.942)	(−4.665)	(−2.576)	(−1.147)
Reversing R&D cut	0.018***	0.008*	0.012**	0.010**	0.002	0.086
	(3.371)	(1.838)	(2.226)	(2.447)	(0.023)	(0.899)
Controls and Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	No	Yes
N	30,173	30,173	30,173	30,173	20,880	8,168
Adjusted (Pseudo) R^2	0.552	0.642	0.533	0.689	0.020	0.023

Panel B: Our setting and models supplemented with the Vorst (2016) earnings manipulation variables (reversing and non-reversing SG&A cuts)

Dependent variable	*Future ROA*$_{t+1}$ (1)	*Future ROA3*$_{t+1,t+3}$ (2)	*Future CFO*$_{t+1}$ (3)	*Future CFO3*$_{t+1,t+3}$ (4)	*Restate*(logit) (5)	*Restate*(clogit) (6)
REMComp	−0.088***	−0.079***	−0.061***	−0.040***	0.159	0.520**
	(−8.067)	(−6.536)	(−5.725)	(−3.521)	(1.261)	(2.259)
Non-reversing SG&A cut	−0.023***	−0.004	−0.003	0.009	0.153*	0.126
	(−3.429)	(−0.651)	(−0.480)	(1.469)	(1.663)	(0.887)
Reversing SG&A cut	0.036**	0.037***	0.019	0.019*	0.411***	0.266
	(2.401)	(3.123)	(1.397)	(1.885)	(2.730)	(1.259)
Controls and Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	No	Yes
N	30,173	30,173	30,173	30,173	20,880	8,168
Adjusted (Pseudo) R^2	0.550	0.641	0.532	0.689	0.020	0.023

Notes: This table present results from estimating our models supplemented with the Vorst (2016) manipulation variables. We present *t*-statistics below the coefficients. See Appendix 2 for variable definitions. ***, **, and * denote two-tailed *p*-values of less than 0.01, 0.05, and 0.10, respectively, based on robust standard errors clustered by firm.

Moreover, these tests suggest that even though the specific component loadings from the PCA differ between Vorst's (2016) setting and our setting, the resulting components both work well in predicting future operating performance costs. Therefore, it appears that the PCA approach is effective because it considers multiple REM tools used concurrently along with other signals of REM. Thus, we expect the components approach to work well across settings without the need to impose strict data requirements such as a long time-series of data or non-missing discretionary expense values.

Results by firm life cycle

Agency conflicts and the associated effect on earnings management strategies vary across firm life cycles (Diamond 1991; Dickinson 2011). For example, agency conflicts for mature firms often manifest from managerial pursuit of growth over stockholder wealth (Mueller 1972) and are intensified by the presence of free cash flows (Jensen 1986). Introductory and growth life-cycle firms, however, typically have less cash than mature firms and are more reliant on debt financing (Diamond 1991; Dickinson 2011; Barclay and Smith 2005), both of which limit managers' ability to pursue value-destroying growth. Consequently, early life-cycle firms typically encounter different types of agency conflicts than mature life-cycle firms and these differences create challenges in measuring earnings manipulation. Yet, researchers aim to measure earnings manipulation in a variety of settings. Consequently, we next explore how *REMComp*, *REM SUM1*, and *REM SUM2* perform across firm life cycles, noting that ideally a measure of earnings manipulation would work well in a variety of settings including across firm life cycles.

Table 10 presents descriptive statistics by firm life cycle. First, we observe that most firms are classified as mature life-cycle firms ($N = 21{,}806$). The second largest group is growth life-cycle stage firms ($N = 14{,}586$), followed by introductory life-cycle firms ($N = 7{,}472$). Together these three life-cycle stages comprise 82% of the sample. Second, the decline life-cycle group, which is the smallest group ($N = 3{,}953$ firm-years) has a lower restatement rate than the rest of the sample (*t*-test untabulated, significant at the 5% level, two-tailed test). This evidence is consistent with the notion that decline life-cycle firms have less incentive to engage in earnings manipulation than firms in other life-cycle stages.

Table 11 reports the results from estimating equation (4) by life-cycle stage (panels A–E) and considers *REM SUM1* and *REM SUM2* (columns (1)–(4)) and the PCA measure, *REMComp* (columns (5)–(8)). The last two rows of each panel report the percentage of the earnings manipulation coefficients that are significantly negative and the percentage that are significantly positive, which is contrary to what we expect for REM measures. First, we do not observe any contrary results for the PCA measure. That is, in no case does the PCA measure exhibit significantly positive associations with future operating performance. For mature life-cycle firms, which is the group with the most firms, 100% of the coefficients for *REMComp* are significant and negative whereas only 50% of the coefficients for *REM SUM1* and *REM SUM2* are significant and negative. For growth firms, the second largest subset of firms, the coefficients for *REMComp* are significantly negative 100% of the time compared to only 38% of the time for *REM SUM1* and *REM SUM2*. In the introductory life-cycle stage 50% of the coefficients for *REMComp* are significantly negative whereas only 13% of coefficients for *REM SUM1* and *REM SUM2* are negative. For shakeout life-cycle stage firms, *REMComp* is significantly negative 100% of the time and *REM SUM1* and *REM SUM2* are only significant and negative 38% of the time. Decline life-cycle stage firms are the only subset where the individual metrics appear to perform better on the dimension of future operating performance, but even then, both approaches yield weak results. This subset is also the smallest and has the lowest restatement rate of any life-cycle stage, suggesting that earnings manipulation is less prevalent in this life-cycle stage. Overall, these results are promising because they suggest that the PCA

TABLE 10
Descriptive statistics for the REM metrics by life-cycle stage

	N	Mean	Median	SD
Introductory life-cycle firms				
REM SUM1	7,472	0.018	0.079	0.499
REM SUM2	7,472	0.069	0.096	0.277
REMComp	7,472	0.130	0.265	1.217
Restate	4,775	0.105	0.000	0.306
Growth life-cycle firms				
REM SUM1	14,586	0.005	0.034	0.372
REM SUM2	14,586	−0.012	0.014	0.222
REMComp	14,586	0.019	0.150	0.955
Restate	9,693	0.119	0.000	0.323
Mature life-cycle firms				
REM SUM1	21,806	−0.012	0.014	0.366
REM SUM2	21,806	−0.029	−0.009	0.208
REMComp	21,806	−0.061	0.040	0.921
Restate	14,729	0.112	0.000	0.316
Decline life-cycle firms				
REM SUM1	3,953	−0.008	0.062	0.471
REM SUM2	3,953	0.058	0.088	0.267
REMComp	3,953	0.018	0.182	1.139
Restate	2,913	0.094	0.000	0.292
Shakeout life-cycle firms				
REM SUM1	5,363	0.017	0.045	0.382
REM SUM2	5,363	0.010	0.033	0.225
REMComp	5,363	0.000	0.096	0.961
Restate	3,698	0.120	0.000	0.324

Notes: For ease of interpretation, we present unranked earnings manipulation variables in this table (i.e., *REM SUM1*, *REM SUM2*, and *REMComp* are not ranked). We define firm life-cycle stage following Dickinson (2011). Introductory life-cycle firms have negative operating and investing cash flow, and positive financing cash flow. Growth life-cycle firms have negative investing cash flow, and positive operating and financing cash flows. Mature life-cycle firms have negative investing and financing cash flows, and positive operating cash flow. Decline life-cycle firms have negative operating cash flow and positive investing cash flow. We classify all other firms as Shakeout firms. We provide other variable definitions in Appendix 2.

approach works well in four of the five firm life-cycle stages and that the PCA approach works better than the summary metrics derived from the individual REM metrics for mature and growth life-cycle firms, which are the largest subsets of firms.

Financial reporting incentives

Dechow and Skinner (2000) suggest that an increase in earnings management in response to manager incentives provides strong evidence of intentional earnings manipulation. Thus, we consider how the measures of REM vary with incentives. Specifically, we examine three incentives that can influence managers' likelihood of managing earnings that have been used in extant research: closeness to the threshold of violating a debt covenant (*Tight*—see Demerjian and Owens 2016), "last-chance earnings management" in which managers just beat last year's earnings by cutting tax expense (*LastChance*—see Dhaliwal et al. 2004), and

TABLE 11

Abnormal REM components and future operating performance by firm life-cycle stage

Panel A: Introductory life-cycle stage firms ($N = 7{,}472$)

Dependent variable	$FutureROA_{t+1}$ (1)	Future $ROA3_{t+1,t+3}$ (2)	Future CFO_{t+1} (3)	Future $CFO3_{t+1,t+3}$ (4)	Future ROA_{t+1} (5)	Future $ROA_{t+1,t+3}$ (6)	Future CFO_{t+1} (7)	Future $CFO3_{t+1,t+3}$ (8)
REM SUM1	0.003	−0.017	−0.004	−0.024				
	(0.134)	(−0.790)	(−0.202)	(−1.299)				
REM SUM2	−0.046**	−0.018	−0.003	0.021				
	(−2.151)	(−0.923)	(−0.127)	(1.212)				
REMComp					−0.044***	−0.044***	0.001	−0.004
					(−2.641)	(−2.764)	(0.075)	(−0.291)
Controls, Firm FE, and Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Percent significantly negative		13%				50%		
Percent significantly positive		0%				0%		

Panel B: Growth life-cycle stage firms ($N = 14{,}586$)

Dependent variable	Future ROA_{t+1} (1)	Future $ROA3_{t+1,t+3}$ (2)	Future CFO_{t+1} (3)	Future $CFO3_{t+1,t+3}$ (4)	Future ROA_{t+1} (5)	Future $ROA_{t+1,t+3}$ (6)	Future CFO_{t+1} (7)	Future $CFO3_{t+1,t+3}$ (8)
REM SUM1	0.044*	−0.004	−0.051**	−0.047**				
	(1.798)	(−0.178)	(−2.009)	(−2.180)				
REM SUM2	−0.123***	−0.035	−0.014	0.019				
	(−4.915)	(−1.579)	(−0.578)	(0.874)				
REMComp					−0.087***	−0.052***	−0.078***	−0.037**
					(−5.226)	(−2.987)	(−4.747)	(−2.356)
Controls, Firm FE, and Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Percent significantly negative		38%				100%		
Percent significantly positive		13%				0%		

(The table is continued on the next page.)

TABLE 11 (continued)

Panel C: Mature life-cycle stage firms (N = 21,806)

Dependent variable	*Future* ROA_{t+1} (1)	*Future* $ROA3_{t+1,t+3}$ (2)	*Future* CFO_{t+1} (3)	*Future* $CFO3_{t+1,t+3}$ (4)	*Future* ROA_{t+1} (5)	*Future* $ROA_{t+1,t+3}$ (6)	*Future* CFO_{t+1} (7)	*Future* $CFO3_{t+1,t+3}$ (8)
REM SUM1	0.018	0.017	−0.082***	−0.046**				
	(0.933)	(0.865)	(−3.982)	(−2.534)				
REM SUM2	−0.059***	−0.038**	0.025	0.007				
	(−3.129)	(−2.152)	(1.259)	(0.412)				
REMComp					−0.062***	−0.042***	−0.069***	−0.040***
					(−5.009)	(−3.069)	(−5.503)	(−3.214)
Controls, Firm FE, and Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Percent significantly negative		50%				100%		
Percent significantly positive		0%						

Panel D: Decline life-cycle stage firms (N = 3,953)

Dependent variable	*Future* ROA_{t+1} (1)	*Future* $ROA3_{t+1,t+3}$ (2)	*Future* CFO_{t+1} (3)	*Future* $CFO3_{t+1,t+3}$ (4)	*Future* ROA_{t+1} (5)	*Future* $ROA_{t+1,t+3}$ (6)	*Future* CFO_{t+1} (7)	*Future* $CFO3_{t+1,t+3}$ (8)
REM SUM1	0.033	−0.002	0.031	0.006				
	(1.079)	(−0.079)	(1.058)	(0.250)				
REM SUM2	−0.048*	−0.039	−0.057**	−0.040**				
	(−1.780)	(−1.423)	(−1.999)	(−1.979)				
REMComp					−0.016	−0.046*	−0.004	−0.026
					(−0.656)	(−1.725)	(−0.183)	(−1.354)
Controls, Firm FE, and Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Percent significantly negative		38%				25%		
Percent significantly positive		0%				0%		

(The table is continued on the next page.)

TABLE 11 (continued)

Panel E: Shakeout life-cycle stage firms (N = 5,363)

Dependent variable	*Future* ROA_{t+1} (1)	*Future* $ROA3_{t+1,t+3}$ (2)	*Future* CFO_{t+1} (3)	*Future* $CFO3_{t+1,t+3}$ (4)	*Future* ROA_{t+1} (5)	*Future* $ROA_{t+1,t+3}$ (6)	Future CFO_{t+1} (7)	Future $CFO3_{t+1,t+3}$ (8)
REM SUM1	0.009	−0.057	−0.080**	−0.093***				
	(0.241)	(−1.558)	(−2.077)	(−2.678)				
REM SUM2	−0.077**	−0.023	0.007	0.009				
	(−2.140)	(−0.726)	(0.199)	(0.320)				
REMComp					−0.065**	−0.078***	−0.056**	−0.065**
					(−2.502)	(−2.756)	(−2.166)	(−2.312)
Controls, Firm FE, and Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Percent significantly negative	38%				100%			
Percent significantly positive	0%				0%			

Notes: We present t-statistics below the coefficients. See Appendix 2 for variable definitions. ***, **, and * denote two-tailed p-values of less than 0.01, 0.05, and 0.10, respectively, based on robust standard errors that are clustered by firm. See the note accompanying Table 10 for life-cycle stage definitions.

TABLE 12
The REM components and earnings manipulation reporting incentives

Panel A: *Tight* and REM

		Dependent variable	
	Pred. sign	*REMComp Top Quartile* (1)	*REM SUM1* and *REM SUM2 Top Quartile* (2)
Tight	+	0.207***	0.077
		(2.751)	(0.927)
N		53,180	53,180
Pseudo R^2		0.007	0.013

Panel B: *LastChance* and REM

		Dependent variable	
	Pred. sign	*REMComp Top Quartile* (1)	*REM SUM1* and *REM SUM2 Top Quartile* (2)
LastChance	+	0.026	−0.043
		(1.003)	(−1.526)
N		53,180	53,180
Pseudo R^2		0.007	0.013

Panel C: *JustBeat* and REM

		Dependent variable	
	Pred. sign	*REMComp Top Quartile* (1)	*REM SUM1* and *REM SUM2 Top Quartile* (2)
JustBeat	+	0.328***	0.244***
		(4.368)	(2.958)
N		45,994	45,994
Pseudo R^2		0.008	0.014

Notes: We present z-statistics below the coefficients. See Appendix 2 for variable definitions. ***denotes a two-tailed p-value of less than 0.01 based on robust standard errors that are clustered by firm. *Tight* reflects years with relatively little slack in meeting their private debt contracts. *LastChance* reflects years where in the fourth quarter the firm is likely to have managed earnings to report fourth-quarter earnings that exceed last year's fourth-quarter earnings. *JustBeat* reflects just beating zero earnings. The primary control variables used in our main analyses are not included in these models. The first-stage control variables used in equations (1)–(3) are included in the models. The control variables in equations (1)–(3) are sales and lagged sales (*Sales*), and the change in sales and lagged change in sales (*Sales Change*). We estimate the models using a logit model.

JustBeat for firm-years that just beat zero earnings. Appendix 2 provides details on how we calculate these variables. We expect that *Tight*, *LastChance*, and *JustBeat* will increase real earnings management incentives.

Table 12 reports the results. We exclude firm and year fixed effects and we add the first-stage control variables used in the earnings management models (i.e., the control variables included in

TABLE 13
Summary of future performance results

Measures of REM	*REM SUM1*	*REM SUM2*	*REMComp*
Future performance measures			
Future ROA_{t+1}	0	1	1
Future $ROA3_{t+1,t+3}$	0	1	1
Future CFO_{t+1}	1	1	1
Future $CFO3_{t+1,t+3}$	1	1	1
Earnings Persistence (2 estimations)	0	1	2
SalesGrowth (2 estimations)	2	0	2
Lifecycle Stages Future ROA_{t+1} and $ROA3_{t+1,t+3}$ (10 estimations)	0	6	9
Lifecycle Stages Future CFO_{t+1} and $CFO3_{t+1,t+3}$ (10 estimations)	6	2	6
Count of negative associations	10	13	23
Proportion of tests with negative associations (28)	36%	46%	82%
Total number of significantly positive coefficients	2	0	0

Notes: This table summarizes the results presented in Tables 3–7 and 11—the "future performance" tests. In each of the tables, we count the number of negative associations between the three REM measures and future performance. This amounts to a total of 28 regressions, of which *REMComp* most frequently associates negatively with future performance (82% of the tests) relative to *REM SUM1* (36% of tests) and *REM SUM2* (46% of tests).

equations (1)–(3)) as suggested by Chen et al. (2018).[20] To provide a more powerful test, we create indicator variables that are set to one when the component is in the top quartile for the year (*REMComp Top Quartile*) and when both simple sums are in the top quartile for the year. We examine the association between each incentive (presented in panels A–C) and presence in the top quartile of REM for the year. The results from these analyses provide evidence that *REMComp* changes in the presence of elevated earnings manipulation incentives in the direction expected for all but the *LastChance* incentive, though it is directionally consistent with expectations. In contrast, the incentives are only significantly associated with *REM SUM1* and *REM SUM2* for *JustBeat*, but not for *Tight* or *LastChance*.

Summary of results and calibration tests

Table 13 summarizes the results by tallying the proportion of times that each measure (*REM SUM1*, *REM SUM2*, and *REMComp*) exhibits the expected negative relation with future performance. The results point to *REMComp* as the measure that most consistently associates negatively with future earnings and future cash flows, and most consistently reduces earnings and sales growth persistence. *REMComp* generates the expected negative relation in 82% of our tests and never exhibits a positive association with future performance. In contrast, *REM SUM2* generates the expected negative relation in 46% of tests. Similarly, *REM SUM1* associates negatively with future performance in 36% of tests and positively with future performance in two tests (7%).

To ensure robustness of our conclusions, we perform serval additional calibration tests. In untabulated analyses, we examine the robustness of the Table 4 results to the following adjustments and in all cases, we find similar results. First, we exclude the accrual component from

20. Chen et al. (2018) suggest that a simple way to mitigate potential bias when residuals from first-stage manipulation models are used as dependent variables is to include the independent variables from the first stage models in subsequent regression analyses.

the equation. Second, we control for lagged ROA rather than contemporaneous ROA. Third, we exclude all control variables except for year and firm fixed effects. Fourth, we replace firm fixed effects with industry fixed effects. Fifth, we estimate equation (4) using annual Fama and MacBeth annual regressions. Sixth, we replace ROA with operating income (OIADP) scaled by average assets and include a control for special items (also scaled by average assets). Seventh, we also replicate Table 3 but include the individual metrics or simple sums one at a time rather than concurrently and find similar results except that *AbnLowExp* is inconsistently negatively associated with future performance. Eighth, we replicate Table 4 but include *REMComp*, *REM SUM1*, and *REM SUM2*. The inferences for *REMComp* are consistent with those reported in Table 4. Ninth, we limit the observations used in Tables 3 and 4 to observations where analysts predict earnings growth thus reducing the likelihood that cuts to expenses reflect prudent business decisions. Results for *REMComp* are stronger than those tabulated in Table 4.

In summary, our analyses provide evidence that the PCA approach results in an REM measure that possesses characteristics expected of earnings manipulation and varies with incentives. Without adjustment, *REMComp* can be used as a control variable in a variety of settings without concern over multicollinearity. Recent research highlights additional steps that researchers must take to avoid incorrect inferences when using residuals from regression models as dependent variables in subsequent analyses (Chen et al. 2018). Because we derive the components from residuals, the concerns and solutions articulated by Chen et al. (2018) may be relevant to consider when the component is used as a dependent variable.

5. Conclusion

Although researchers have invested considerable energy into the development of accrual-based earnings manipulation measures, REM measures remain relatively underdeveloped. One challenge in identifying REM is that companies can change their operating and investing decisions for strategic reasons unrelated to earnings manipulation, and it is difficult to distinguish strategic changes from earnings manipulation. Recent studies develop improvements to standard REM metrics that better distinguish manipulation from strategic changes (Cohen et al. 2019; Srivastava 2019; Vorst 2016). However, because these new methodologies are impractical in some settings and empiricists still frequently use the standard REM proxies in varying contexts, additional innovations are warranted.

We develop a REM measure that leverages the insight that managers can simultaneously employ multiple strategies to manage real activities (Roychowdhury 2006) and the observation that real earnings management is more likely to occur in specific settings. To the extent that managers concurrently use multiple tools to manage earnings and that abnormal activities are more likely to reflect manipulation than strategy in some situations, considering activities and signals will better separate REM from changes in firm strategy. We use PCA to develop a measure of REM that reflects the concurrent use of multiple REM activities, and increases when these activities are more likely to reflect manipulation. To validate our PCA measure, we examine the association between the PCA measure and both future performance and earnings persistence, how the PCA measure performs in simulation analyses, and how it performs across firm life cycles. Moreover, in each of these tests we compare the performance of the PCA component to similar summary metrics commonly used in the literature. Across all tests we find that the PCA component outperforms the standard metrics. Thus, we present a simple approach to measure REM that can be used in a variety of settings and performs better than the standard metrics. Our approach, together with other recent innovations (Cohen et al. 2019; Srivastava 2019; Vorst 2016) arm researchers with REM metrics that better distinguish earnings manipulation from changes stemming from shifts in operating and investing strategies.

Appendix 1: Number of real earnings management studies in the top-five accounting journals

This appendix tabulates the number of studies over the period 2015–2020 that consider real earnings management (row 1), the number of studies that measure real earnings management using the standard individual metrics developed by Roychowdhury (2006) of Gunny (2010) (*AbnLowExp*, *AbnLowCFO*, or *AbnHighProd*, rows 2 and 3), the number of studies examining abnormal discretionary expenses (*AbnLowExp)* individually (row 4), the number of studies using a summary REM measure calculated as the sum of two or three of the standard individual metrics (*AbnLowExp*, *AbnLowCFO*, or *AbnHighProd*), and the number of subsequent papers citing the papers in row 3 (i.e., the number of papers that cite research that uses the standard individual metrics to measure REM). Citations are from Google Scholar and obtained on June 28, 2022. We identify earnings management studies by searching each article in the publication outlet over the last five years and reviewing the articles to determine (i) whether the study examines real earning management and (ii) if so, whether real earnings management is measured using the standard metrics. We provide variable definitions for the earnings management metrics in Appendix 2.

Row		*The Accounting Review*	*Journal of Accounting Research*	*Journal of Accounting and Economics*	*Contemporary Accounting Research*	*Review of Accounting Studies*	Total
1	Total real earnings management studies	14	3	9	17	11	54
2	Total real earnings management studies using:						
	AbnExp	9	2	6	13	7	37
	AbnCFO	4	2	4	12	3	25
	AbnProd	7	1	3	13	5	29
3	Total unique studies using at least one of the REM metrics (% based on row 1)	9 (64.3%)	3 (100%)	7 (78%)	15 (88%)	7 (64%)	41 (76%)
4	Total unique studies examining abnormal discretionary expenses individually (% based on row 3)	1 (11.1%)	2 (66.7%)	4 (57.1%)	4 (26.7%)	6 (85.7%)	17 (41.5%)
5	Total unique studies using a summary REM measure calculated as the sum of two or three of the standard individual metrics (% based on row 3)	6 (66.7%)	0 (0%)	2 (28.6%)	9 (60.0%)	2 (28.6%)	19 (46.3%)
6	Citations: Papers citing the studies using at least one of the REM metrics	1,065	408	1,483	796	560	4,312

Appendix 2: Variable definitions

Variable	Definition
REM measures	
AbnLowExp	Real activities-based earnings management: The residual from the model of normal discretionary expenses (Roychowdhury 2006). We replace missing values of R&D with zero. We multiply the residual by negative one so that it is increasing in the real earnings management activity. We decile rank *AbnLowExp* by year, where ranks range from zero to one, when used in regression analyses
AbnLowCFO	Real activities-based earnings management: The residual from the model of normal cash flow from operations (Roychowdhury 2006). We multiply the residual by negative one so that it is increasing in the real earnings management activity. We decile rank *AbnLowCFO* by year, where ranks range from zero to one, when used in regression analyses
AbnHighProd	Real activities-based earnings management: The residual from the model of normal production (Roychowdhury 2006). We decile rank *AbnHighProd* by year, where ranks range from zero to one, when used in regression analyses
REM SUM1	Sum of *AbnLowExp* and *AbnHighProd*. We decile rank *REM SUM1* by year, where ranks range from zero to one, when used in regression analyses
REM SUM2	Sum of *AbnLowExp* and *AbnLowCFO*. We decile rank *REM SUM2* by year, where ranks range from zero to one, when used in regression analyses
REMComp	First component from a PCA of *AbnLowCFO*, *AbnHighProd*, *AbnLowExp*, *LitigiousIndustry*, *LitigiousCircuit*, and *PastAM*. We decile rank *REMComp* by year, where ranks range from zero to one, when used in regression analyses
Inputs to REM models and PCA	
Sales	Sales (SALE) divided by average total assets
Sales Change	Change in sales from the prior year scaled by average total assets (SALE_t – SALE_{t-1})/Average Assets
CFO	Cash flow from operations less extraordinary items and discontinued operations (OANCF – XIDOC) divided by average total assets
Production Costs (PROD)	Cost of goods sold plus the change in inventory ($\text{COGS}_t + (\text{INVT}_t - \text{INVT}_{t-1})$) scaled by average total assets
Discretionary Expenses (EXP)	Sum of R&D expenses (XRD) scaled by average total assets and SG&A expenses (XSGA) scaled by average total assets. We replace missing values of XRD and XSGA with 0
LitigiousIndustry	Takes on the value of one for firms with SICH codes in highly litigious industries, namely SICH 2833–2836, 8731–8734, 3570–3577, 7370–7374, 3600–3674, and 5200–5961, and zero otherwise
LitigiousCircuit	Takes on the value of one for firm-years in the ninth district before 2000 and zero otherwise
PastAM	Takes on the value of one for firm-years with net operating assets above their FF48 industry median for that year and zero otherwise. Net operating assets are defined as operating assets (AT-CHE) less operating liabilities (AT-DLC-DLTT-MIBT-PSTK-CEQ) scaled by lagged sales (SALE)
Dependent variables	
$FutureROA_{t+1}$	Net income (NI) in year $t + 1$ scaled by the firm's average assets. We decile rank *FutureROA* by year, where ranks range from zero to one, when used in regression analyses

(The table is continued on the next page.)

(continued)

Variable	Definition
FutureROA3$_{t+1,t+3}$	Average of *FutureROA* over year $t+1$ to $t+3$. We decile rank *FutureROA3* by year, where ranks range from zero to one, when used in regression analyses
FutureCFO$_{t+1}$	Cash flow from operating activities (OANCF) scaled by the firm's average assets in year $t+1$. We decile rank *FutureCFO* by year, where ranks range from zero to one, when used in regression analyses
FutureCFO3$_{t+1,t+3}$	Average of *FutureCFO* over year $t+1$ to $t+3$. We decile rank *FutureCFO3* by year, where ranks range from zero to one, when used in regression analyses
Restate	Indicator variable set equal to one in financial reporting periods that are subsequently restated. We identify restatement for years 1997–2006 from the Hennes et al. (2008) restatement data set and for years 2007 forward from the Audit Analytics database. For the Hennes et al. (2008) data, the restatement announcement date is known, but not the years being restated. Thus, for this data only, we assume that low-quality financial reporting in the year that a firm announced a restatement and in the prior two years. Our results are not sensitive to this decision as we find similar results if we assume that only the announcement year and prior year or only the announcement year are low-quality reporting periods. For Audit Analytics restatements, the restatements period is provided. Note that restatement data are not available for the earlier years in our sample (prior to 1997) and thus the sample size is reduced for this test
Control variables	
Accruals Comp	Resulting component from a principal component analysis of four accruals-based earnings management variables: (i) the cross-sectional version of the Jones model (Jones 1991), (ii) the modified version of the cross-sectional Jones model (Dechow et al. 1995), (iii) a modified version of the cross-sectional Jones model supplemented with additional controls for current, future, and past operating cash flows (Dechow and Dichev 2002), and (iv) the residual from the accounts receivable model of Stubben (2006, 2010). We decile rank *Accruals Comp* by year, where ranks range from zero to one, when used in regression analyses. The component has an eigenvalue of 2.772 and explains variance of 0.924
ROA	Net income (NI) in year t scaled by the firm's average assets
AbnormalReturn	Value-weighted market-adjusted, buy-and-hold return over the year t
SalesGrowth	Industry-adjusted sales growth. Sales growth is defined as sales during the year ($SALE_t$) less prior year's sales ($SALE_{t-1}$) less the increase in receivables all scaled by the prior year's sales. We industry adjust sales growth each year based on the Fama and French (1997) industry median except when used in Table 7 (see the note accompanying Table 7)
MB	Market-to-book ratio defined as the firm's market capitalization (PRCC_F $\times$ CSHO) divided by book value (SEQ) at year t
Size	Natural log of total assets in year t
NumAnalysts	Log of 1 plus the number of analysts following the firm in year t, based on the most recent forecast before year-end
BigNAuditor	Indicator variable set equal to one for firms audited by Big N firms in year t; zero otherwise. Firms with Compustat "AU" codes between 1 and 8, inclusive, are classified as Big N auditors. Specifically, we classify Arthur Andersen, Ernst & Young, Deloitte & Touche, KPMG, and PricewaterhouseCoopers as Big N auditors
RDMissing	Indicator taking on the value of one for firm-years where R&D expenditures (XRD) in Compustat are missing, and zero otherwise
TotalAccruals	Total accruals are defined as net income (NI) − operating cash flows (OANCF) scaled by average total assets (AT)

(The table is continued on the next page.)

(continued)

Variable	Definition
Additional variables used in Vorst replication	
AbnROA	Earnings before extraordinary items (IB), scaled by lagged total assets (AT). We adjust ROA for the industry-year median, where industries are defined using 2-digit SIC codes
AbnCFO	Cash flow from operating activities (OANCF) scaled by lagged total assets (AT). We adjust CFO for the industry-year median, where industries are defined using 2-digit SIC codes
Non-reversing R&D cut	Indicator variable that equals one if Abnormal R&D in year *t* is in the bottom quintile and there is no reversal, zero otherwise. Abnormal R&D is the residual from the model of normal R&D (Gunny 2010)
Reversing R&D cut	Indicator variable that equals one if Abnormal R&D in year *t* is in the bottom quintile and there is a reversal, zero otherwise. Abnormal R&D is the residual from the model of normal R&D (Gunny 2010)
Non-reversing SG&A cut	Indicator variable that equals one if Abnormal SG&A in year *t* is in the bottom quintile and there is no reversal, zero otherwise. Abnormal SG&A is the residual from the model of normal SG&A (Gunny 2010)
Reversing SG&A cut	Indicator variable that equals one if Abnormal SG&A in year *t* is in the bottom quintile and there is a reversal, zero otherwise. Abnormal SG&A is the residual from the model of normal SG&A (Gunny 2010)
BTM	Book value (SEQ) divided by the firm's market capitalization (CSHO×PRCC_F) in year *t*
Z-score	Altman's *z*-score (Altman 1968), calculated as 3.3 × (Net Income/Lagged Assets) + (Sales/Lagged Assets) + 1.4 × (Retained Earnings/Lagged Assets) + 1.2 × ((ACT-LCT)/Lagged Assets)
Return	Value-weighted market-adjusted, buy-and-hold return over the year *t*
Financial reporting incentive variables	
Tight	Following Demerjian and Owens (2016, table 4), we examine seven financial statement covenants that use earnings in their definition. We calculate covenant slackness for each firm-quarter-covenant by scaling the Compustat standard definition (see table 4 in Demerjian and Owens 2016) by the covenant threshold reported in DealScan. We then rank for each covenant the slackness measure by quarter. *Tight* takes a value of one when the slackness measure in any quarter of the fiscal year falls into the lowest decile of the covenant ranking for that quarter and zero otherwise
LastChance	Similar to Dhaliwal et al. (2004), we identify firms (i) where fourth quarter tax expense (TXTQ) is less than tax expense reported in the third quarter and (ii) fourth quarter pre-tax income (PIQ), scaled by common shares outstanding (CSHO), is less than last year's fourth quarter pre-tax income scaled by CSHO. *LastChance* takes a value of one in firm-years where both criteria are met
JustBeat	Similar to Zhang (2012), *JustBeat* equals one for firm-years with net income over lagged total assets between 0% and 0.5% and zero for firm-years that miss or beat zero earnings benchmark by 2.5% of lagged total assets. Following Zang (2012), other firm-years close to zero earnings have *JustBeat* set to missing and are thus omitted from regression results

Notes: Unless we use ranked values in regression analyses, we winsorize all continuous variables at 1% and 99% by year.

References

Allee, K. D., C. Do, and F. G. Raymundo. 2022. Factor analysis and principal component analysis in accounting research. *Journal of Financial Reporting*, https://doi.org/10.2308/JFR-2021-005

Allison, P. 2005. *Fixed Effects Regression Methods for Longitudinal Data Using SAS*. Cary, NC: SAS Institute.

Allison, P. 2009. *Fixed Effects Regression Models. Quantitative Applications in Social Sciences Series*. Thousand Oaks, CA: Sage Publications.

Altman, E. 1968. Financial ratios, discriminate analysis and the prediction of corporate bankruptcy. *Journal of Finance* 23 (4): 589–609.

Badertscher, B. 2011. Overvaluation and its effect on management's choice of alternative earnings management mechanisms. *The Accounting Review* 86 (5): 1491–518.

Badertscher, B., D. Collins, and T. Lys. 2012. Discretionary accounting choices and the predictive ability of accruals with respect to future cash flows. *Journal of Accounting and Economics* 53: 330–52.

Barclay, M., and C. Smith Jr. 2005. The capital structure puzzle: The evidence revisited. *Journal of Applied Corporate Finance* 17 (1): 8–17.

Barton, J., and P. Simko. 2002. The balance sheet as an earnings management constraint. *The Accounting Review* 77: 1–27.

Barua, A., S. Lin, and A. Sbaraglia. 2010. Earnings management using discontinued operations. *The Accounting Review* 85 (5): 1485–509.

Beatty, A., S. Chamberlain, and J. Magliolo. 1995. Managing financial reports of commercial banks: The influence of taxes, regulatory capital, and earnings. *Journal of Accounting Research* 33 (2): 231–61.

Becker, C., M. DeFond, J. Jiambalvo, and K. Subramanyam. 1998. The effect of audit quality on earnings management. *Contemporary Accounting Research* 15 (1): 1–24.

Beneish, M. D. 1997. Detecting GAAP violations: Implications for assessing earnings management among firms with extreme financial performance. *Journal of Accounting and Public Policy* 16: 271–309.

Bereskin, F., P. Hsu, and W. Rotenberg. 2018. The real effects of real earnings management: Evidence from innovation. *Contemporary Accounting Research* 35 (1): 525–57.

Bhojraj, S., P. Hribar, M. Picconi, and J. McInnis. 2009. Making sense of cents: An examination of firms that marginally miss or beat analyst forecasts. *Journal of Finance* 64 (5): 2361–88.

Bowen, R., S. Rajgopal, and M. Venkatachalam. 2008. Accounting discretion, corporate governance, and firm performance. *Contemporary Accounting Research* 25 (2): 351–405.

Cameron, C., and D. Miller. 2015. A practitioner's guide to cluster-robust inferences. *Journal of Human Resources* 50 (2): 317–72.

Casey, R., F. Gao, M. Kirschenheiter, S. Li, and S. Pandit. 2017. Articulation-based accruals. *Review of Accounting Studies* 22: 288–319.

Chan, L., K. Chen, T. Chen, and Y. Yu. 2015. Substitution between real and accruals-based earnings management after voluntary adoption of compensation clawback provisions. *The Accounting Review* 90 (1): 147–74.

Chen, W., P. Hribar, and S. Melessa. 2018. Incorrect inferences when using residuals as dependent variables. *Journal of Accounting Research* 56 (3): 751–96.

Cheng, Q., J. Lee, and T. Shevlin. 2016. Internal governance and real earnings management. *The Accounting Review* 91 (4): 1051–85.

Christensen, T., A. Huffman, M. Lewis-Western, and R. Scott. 2022. Accruals earnings management proxies: Prudent business decisions or earnings manipulation? *Journal of Business Finance & Accounting* 49: 536–87.

Cohen, D., A. Dey, and T. Lys. 2008. Real and accrual-based earnings management in the pre- and post-Sarbanes-Oxley period. *The Accounting Review* 83 (3): 757–87.

Cohen, D., S. Pandit, C. Wasley, and T. Zach. 2019. Measuring real activity management. *Contemporary Accounting Research* 37 (2): 1172–98.

Cohen, D., and P. Zarowin. 2010. Accrual-based and real earnings management activities around seasoned equity offerings. *Journal of Accounting and Economics* 50 (1): 2–19.

Core, J., R. Holthausen, and D. Larcker. 1999. Corporate governance, chief executive officer compensation, and firm performance. *Journal of Financial Economics* 51: 371–406.

Correia, S. 2015. *Singletons, Cluster-Robust Standard Errors and Fixed Effects: A Bad Mix*. Durham, NC: Duke University.

Curtis, A., S. McVay, and S. Toynbee. 2020. The changing implications of research and development expenditures for future profitability. *Review of Accounting Studies* 25: 405–37.

Dechow, P., and I. Dichev. 2002. The quality of accruals and earnings: The role of accrual estimation errors. *The Accounting Review* 77 (4): 35–59.

Dechow, P., W. Ge, C. Larson, and R. Sloan. 2011. Predicting material accounting misstatements. *Contemporary Accounting Research* 28 (1): 17–82.

Dechow, P., W. Ge, and C. Schrand. 2010. Understanding earnings quality: A review of the proxies, their determinants and their consequences. *Journal of Accounting and Economics* 50: 344–401.

Dechow, P., A. Hutton, J. Kim, and R. Sloan. 2012. Detecting earnings management: A new approach. *Journal of Accounting Research* 50 (2): 275–334.

Dechow, P., and D. Skinner. 2000. Earnings management: Reconciling the views of accounting academics, practitioners, and regulators. *Accounting Horizons* 14 (2): 235–50.

Dechow, P., R. Sloan, and A. Sweeney. 1995. Detecting earnings management. *The Accounting Review* 70 (2): 193–225.

Demerjian, P., B. Lev, M. Lewis, and S. McVay. 2013. Managerial ability and earnings quality. *The Accounting Review* 88 (2): 463–98.

Demerjian, P., M. Lewis-Western, and S. McVay. 2020. How does intentional earnings smoothing vary with managerial ability? *Journal of Accounting, Auditing & Finance* 35 (2): 406–37.

Demerjian, P., and E. Owens. 2016. Measuring financial covenant strictness in private debt contracts. *Journal of Accounting and Economics* 61 (2): 433–47.

Dhaliwal, D., C. Gleason, and L. Mills. 2004. Last-chance earnings management: Using tax expense to meet analysts' forecasts. *Contemporary Accounting Research* 21 (2): 431–59.

Diamond, D. 1991. Monitoring and reputation: The choice between bank loans and directly placed debt. *Journal of Political Economy* 99 (4): 689–721.

Dickinson, V. 2011. Cash flow patterns as a proxy for firm life cycle. *The Accounting Review* 86 (6): 1969–94.

Efendi, J., A. Srivastava, and E. Swanson. 2007. Why do corporate managers misstate financial statements? The role of in-the-money options and other incentives. *Journal of Financial Economics* 85 (3): 667–708.

Ettredge, M., S. Scholz, K. Smith, and L. Sun. 2010. How do restatements begin? Evidence of earnings management preceding restated financial reports. *Journal of Business and Finance & Accounting* 37 (3): 332–55.

Fama, E., and K. French. 1997. Industry costs of equity. *Journal of Financial Economics* 43 (2): 153–93.

Fama, E., and J. MacBeth. 1973. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy* 81 (3): 607–36.

Fields, T., T. Lys, and L. Vincent. 2001. Empirical research on accounting choice. *Journal of Accounting and Economics* 31 (1–3): 255–308.

Gleason, C. A., N. T. Jenkins, and W. B. Johnson. 2008. The contagion effects of accounting restatements. *The Accounting Review* 83 (1): 83–110.

Gormley, T., and D. Matsa. 2014. Common errors: How to (and not to) control for unobserved heterogeneity. *Review of Financial Studies* 27 (2): 617–61.

Gow, I., D. Larcker, and P. Reiss. 2016. Casual inferences in accounting research. *Journal of Accounting Research* 54 (2): 477–523.

Graham, J., C. Harvey, and S. Rajgopal. 2005. The economic implications of corporate financial reporting. *Journal of Accounting and Economics* 40: 3–73.

Gunny, K. 2010. The relation between earnings management using real activities manipulation and future performance: Evidence from meeting earnings benchmarks. *Contemporary Accounting Research* 27 (3): 855–88.

Healy, P., and J. Wahlen. 1999. A review of the earnings management literature and its implications for standard setting. *Accounting Horizons* 13 (4): 365–83.

Hennes, K., A. Leone, and B. Miller. 2008. The importance of distinguishing errors from irregularities in restatement research: The case of restatements and CEO/CFO turnover. *The Accounting Review* 83 (6): 1487–519.

Hirshleifer, D., P. Hsu, and D. Li. 2018. Innovative originality, profitability and stock returns. *Review of Financial Studies* 31 (7): 2553–605.

Huang, S., S. Roychowdhury, and E. Sletten. 2020. Does litigation deter or encourage real earnings management? *The Accounting Review* 95 (3): 251–78.

Jensen, M. C. 1986. Agency cost of free cash flow, corporate finance, and takeovers. *American Economic Review* 76 (2): 323–9.

Jones, J. 1991. Earnings management during import relief investigations. *Journal of Accounting Research* 29: 193–228.

Kedia, S., K. Koh, and S. Rajgopal. 2015. Evidence on contagion in earnings management. *The Accounting Review* 90 (6): 2337–73.

Koh, P., and D. Reeb. 2015. Missing R&D. *Journal of Accounting and Economics* 60 (1): 73–94.

Kothari, S. P., A. Leone, and W. Charles. 2005. Performance matched discretionary accrual measures. *Journal of Accounting and Economics* 39 (1): 163–97.

Kothari, S. P., N. Mizik, and S. Roychowdhury. 2016. Managing for the moment: The role of earnings management via real activities versus accruals in SEO valuation. *The Accounting Review* 91 (2): 559–86.

Larcker, D., and T. Rusticus. 2010. On the use of instrumental variables in accounting research. *Journal of Accounting and Economics* 49: 186–205.

Laux, C. 2014. Pay convexity, earnings manipulation, and project continuation. *The Accounting Review* 89 (6): 2233–59.

McNichols, M. 2002. Discussion of the quality of accruals and earnings: The role of accrual estimation errors. *The Accounting Review* 77 (S-1): 61–9.

Mueller, D. 1972. A life cycle theory of the firm. *Journal of Industrial Economics* 20 (3): 199–219.

Petersen, M. 2009. Estimating standard errors in financial panel data sets: Comparing approaches. *Review of Financial Studies* 22 (1): 435–80.

Porter, M. E. 1980. *Competitive Strategy: Techniques for Analyzing Industries and Competitors*. New York: Free Press.

Richardson, S. A., R. G. Sloan, M. T. Soliman, and I. Tuna. 2005. Accrual reliability, earnings persistence and stock prices. *Journal of Accounting and Economics* 39 (3): 437–85.

Roychowdhury, S. 2006. Earnings management through real activities manipulation. *Journal of Accounting and Economics* 42 (3): 335–70.

Roychowdhury, S., N. Shroff, and R. Verdi. 2019. The effects of reporting and disclosure on corporate investment: A review. *Journal of Accounting and Economics* 68 (1–2): 1–27.

Schipper, K. 1989. Commentary on earnings management. *Accounting Horizons* 3 (4): 91–102.

Sloan, R. G. 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review* 71 (3): 289–315.

Srivastava, A. 2019. Improving the measures of real earnings management. *Review of Accounting Studies* 24: 1277–316.

Stubben, S. 2006. The use of discretionary revenues to meet earnings and revenue targets. Doctoral dissertation, Stanford University.

Stubben, S. 2010. Discretionary revenues as a measure of earnings management. *The Accounting Review* 85 (2): 695–717.

Vorst, P. 2016. Real earnings management and long-term operating performance: The role of reversals in discretionary investment cuts. *The Accounting Review* 91 (4): 1219–56.

Zang, A. 2012. Evidence on the trade-off between real activities manipulation and accrual-based earnings management. *The Accounting Review* 87 (2): 675–703.