Accounting Restatements: Malfeasance and/or Optimal Incompetence?

Jacquelyn Gillette
MIT Sloan School of Management

Sudarshan Jayaraman
Simon Business School, University of Rochester

Jerold Zimmerman
Simon Business School, University of Rochester

May 2017

Abstract

While accounting research often uses restatements as a measure of earnings quality (EQ), often overlooked is the explicit recognition that EQ arises endogenously as a by-product of the firm's value-maximizing investment in accounting resources. This paper develops a model of restatements that incorporates both the optimal investment in accounting systems (i.e., "optimal incompetence") and incentives to intentionally misrepresent the firm's financial performance (i.e., "malfeasance"). Our framework predicts that smaller, financially weaker firms will optimally invest less in accounting resources than their larger, more profitable counterparts. To identify our reduced-form empirical model, we propose two novel measures of the firm's investment in accounting systems: filing timeliness of the firm's 10-K and spelling errors in the 10-K. Our empirical results validate the predictions of the model. First, larger and more profitable firms are associated with greater accounting resources. Second, accounting resources are negatively associated with the likelihood of a restatement in the subsequent year. Third, including accounting resources in the restatement prediction model diminishes the explanatory power of covariates identified in prior studies as capturing misreporting incentives, especially in the case of restatements flagged as irregularities (i.e., fraud). In cross-sectional tests, we find that the association between accounting resources and future restatements is stronger during booms as compared to downturns – which questions prior inferences on the role of the business cycle on accounting fraud. Our results are robust to controlling for the endogeneity of accounting resources using an instrumental design that exploits regulatory changes in firms' filing deadlines.

We thank Jeff Burks (discussant) and workshop participants at the Duke/UNC Fall Camp, MIT Sloan School of Management, SMU, and the UIUC Young Scholars Conference for helpful comments.

1. Introduction

A large body of accounting research focuses on the determinants and consequences of earnings quality (EQ) (see Dechow, Ge, and Schrand (DGS), 2010).[1] While some prior studies recognize that innate firm characteristics are correlated with observed levels of EQ and that low EQ is not necessarily the result of malfeasance (e.g., Hennes et al., 2008), the mechanisms through which firm fundamentals affect EQ have received relatively less attention compared to the role of managerial opportunism (DGS, 2010).[2] For example, managers of poorly performing companies are often assumed to engage in accounting schemes to improve their earnings and hence lower EQ. Further, bond covenants or capital requirements create earnings management incentives, thereby affecting EQ. In the restatement literature, a large number of studies examine whether future restatements are associated with incentives arising from managerial compensation, non-audit fees, weak corporate governance, low quality auditors, and so forth (DGS, 2010). Often overlooked is the explicit recognition that managers choose the quality and quantity of accounting resources optimally and that observed EQ is a by-product of this investment decision. This paper develops a conceptual framework and provides novel empirical evidence demonstrating that both the firm-value-maximizing investment in accounting systems (i.e., "optimal incompetence") and incentives to intentionally misrepresent the firm's financial performance (i.e., "malfeasance") jointly determine EQ.

Like any other investment decision such as R&D or marketing, firms optimally invest in accounting resources (including hardware, software, and accounting staff human capital) and the

[1] Dechow, Ge, and Schrand (2010) review this literature and list six determinants of EQ: (1) firm characteristics (e.g., performance, debt, growth, investment, and size), (2) financial reporting practices, (3) governance and controls (e.g., characteristics of the board of directors, internal control procedures, managerial share ownership, and compensation), (4) auditors, (5) equity market incentives, and (6) external factors (e.g., capital requirements, political processes, tax and non-tax regulation).

[2] See, for example, Ashbaugh-Skaife et al., 2007; Dechow and Dichev, 2005; Francis et al., 2005.

resulting EQ of the firm is the outcome of this (optimal) investment decision. Like any other firm-value maximizing resource allocation, the optimal amount to invest in accounting resources should equate the returns to the last increment in these resources to the costs of generating that last increment. In the absence of any agency problem, cross-sectional and time-series variation in EQ will still exist due to plausible benefit and cost behavior patterns in the investment in accounting resources (which could be uncorrelated with malfeasance incentives).

We build on prior work that shows that firm size is positively correlated with EQ (DGS, 2010) and that accounting and control systems have a large fixed cost component (Ashbaugh-Skaife,et al., 2007) by developing a simple conceptual framework to flesh out the idea that firms optimally invest in accounting resources. The framework assumes that (1) managers maximize firm value, (2) the benefits to investments in accounting resources increase with firm size (due to inherent fixed costs of investment in accounting resources), (3) the benefits to investments in accounting resources are increasing, but at a decreasing rate, (4) the benefits of accounting resources increase with firm profitability, but at a decreasing rate, and (5) the fixed cost of providing EQ (through an increase in accounting resources) increases in firm size, but at a decreasing rate. Our framework predicts that absent any managerial incentive other than to maximize firm value, smaller, less profitable firms will have relatively smaller and/or less competent accounting staffs. Having less accounting human capital (and systems) will lead to lower EQ (ex-post): more mistakes and hence more accounting restatements, and less timely SEC 10-K and 10-Q filings. Evidence of lower EQ, does not, by itself, indicate that such investments in accounting resources are sub-optimal. To the contrary, the optimal investment in accounting resources is what leads to observed levels of EQ. Stated differently, cross-sectional and time-series variation in observed EQ could be due to managers choosing an optimal level of accounting incompetence rather than (or in addition to) malfeasance.

Some studies view EQ as consisting of innate EQ such as fundamental economics (e.g., profitability, volatility and firm complexity) and discretionary EQ chosen by the managers (Francis et al., 2005). Although these papers recognize that EQ is correlated with innate firm fundamentals, they propose a different mechanism. They suggest that it is harder for managers to estimate accruals when the firm operates in a more volatile or uncertain operating environment (Dechow and Dichev, 2002; Francis et al., 2005). Our paper suggests a different mechanism, which is that so-called innate factors such as size and profitability likely influence the optimal investment in accounting systems, which then affect EQ (i.e., smaller, less profitable firms will invest less in their accounting systems, and hence will have lower EQ). As a result, we expect that holding the uncertainty of the operating environment constant, firms investing less in accounting resources will make less accurate accrual estimates because they have fewer and less competent staff and/or systems. And, these accounting-resource constrained firms will make more mistakes unrelated to accruals estimation. Finally, we propose that firms that invest less in accounting resources are more likely to misrepresent firm performance because managers have more incentives to commit fraud due to a lower likelihood of detection. Simply classifying EQ into “innate” versus “discretionary” misses the important point that so-called innate factors such as size and profitability likely influence the optimal investment in accounting systems, which then affect EQ (i.e., smaller, less profitable firms will invest less in their accounting systems, and hence will have lower EQ). In summary, while prior literature has recognized that innate firm factors are correlated with EQ, we propose a different mechanism (i.e., the investment in accounting resources), which explains why firm fundamentals are related to both errors and malfeasance.

Our conceptual framework demonstrates that the usual reduced-form empirical model employed in many studies suffers from an identification problem. In particular, the same variables – namely size and profitability – are underlying drivers of both the optimal investment in accounting systems and the incentives to misreport. To help solve this problem, we derive a reduced form empirical model that is identified using two novel measures of the investment in accounting resources and show that both optimal incompetence and malfeasance jointly explain EQ.

We focus on restatements as our measure of EQ for several reasons. First, it is a widely used proxy for EQ in the literature. In particular, studies point to the limitations of discretionary accruals to measure EQ, but go on to use restatements as a rather unambiguous measure of EQ (e.g., Desai et al., 2006; McNichols and Stubben, 2008). We wish to emphasize that observed variation in restatements could also be driven by underlying firm-level characteristics of the firm through their effect on the optimal investment in accounting resources. Second, there are a large number of restatements, which increases power in the empirical tests. Third, it does not suffer from the further selection bias induced by focusing on SEC Accounting and Auditing Enforcement Actions (AAERs) (i.e., the SEC chooses which restatements to investigate and the penalties to impose).

To identify our reduced form empirical model, we propose two measures of the firm's investment in accounting systems: filing timeliness of the firm's 10-K and spelling errors in the 10-K. We create a composite measure of accounting resources based on these two measures. Filing timeliness has been used in previous papers, but is usually measured from the end of the fiscal period until the 10-K is filed (or the audit opinion signed). The measure currently used in the literature assumes all firms face the same filing deadlines. In 2002 and again in 2005, the SEC changed the filing deadlines for firms based on their public float. We measure filing timeliness as the number of days prior to its mandated deadline that a firm files its 10-K. So firms filing before their deadline make more timely filings, presumably because they have more accounting resources. Spelling errors in the 10-K is a novel measure of the quality and quantity of the accounting staff. Between the end of the fiscal year and the SEC filing deadline the accounting staff must consolidate and prepare the financial statements, address any issue arising from the annual audit, draft the 10-K, circulate the drafts to senior management, the auditors, the audit committee, and legal counsel. Then, all the comments received must be reconciled and a revised version checked for errors. Finally, the 10-K must be XBRL tagged and filed. As of 2005, all this work had to be done within 90 days for non-accelerated filers, 75 days for accelerated filers, and 60 days for large accelerated filers.[3]

The empirical results are consistent with our predictions. First, the two proxies for accounting resources – filing timeliness and spelling errors are negatively correlated, as expected. Second, our composite measure of accounting resources is positively correlated with firm size and profitability – as predicted. Turning to our main tests, we find that accounting resources are negatively associated with the likelihood of a restatement in the subsequent year. In terms of economic significance, the probability of a restatement in firms at the bottom decile of accounting resources is 0.073 (i.e., 1 in every 14 10Ks filed) as compared to 0.057 in the top decile (1 in every 18 10Ks filed). This result is not only statistically significant across the various specifications we employ, but is also robust to including covariates shown by prior studies to be associated with a restatement. Further, the explanatory power of covariates deemed to capture misreporting incentives diminishes when we include accounting resources in the prediction model. For example,

[3] If the firm is unable to file on time, it must file SEC form 12b-25 (notification of late filing) and state that the 10-K will be filed within 15 days of the deadline.

the negative association between ROA and restatement likelihood could be interpreted as evidence that poorly performing firms are more likely to misreport financial statements. However, including accounting resources diminishes the explanatory power of ROA by around 13%, consistent with our theoretical framework where profitable firms are more likely to invest more in accounting resources (and as a result have fewer restatements).

The above results are merely suggestive as they do not address the endogeneity of accounting resources. We pursue two strategies to alleviate this concern. First, we use an instrumental variable analysis, using the filing window (number of days that a firm has to file its financial statements after the fiscal period end) as our instrument. The SEC requires the filing window to be based on firms' filing status (e.g., large accelerated filer, accelerated filer, small reporting company). The rationale is that a shorter filing window imposes greater stress on the firm's accounting resources. Results indicate that our instrument is not only significantly associated with accounting resources, but also comfortably exceeds the threshold to qualify as a valid instrument (Stock, Wright and Yogo, 2002). Results from this two-stage analysis again support our primary evidence – accounting resources are negatively associated with the likelihood of a future restatement, with this likelihood being 0.082 (1 in every 12 10Ks) for firms in the bottom decile as compared to 0.052 (1 in every 19 10Ks) for those in the top decile of accounting resources.

Second, we exploit changes in filing deadlines as a function of changes in the SEC requirements and a firm's filing status. Specifically, we employ a matched-sample design (similar to a discontinuity design) around changes in the filing deadline using a narrow window around the cutoff for accelerated and large accelerated filers (i.e., $75 million and $700 million in public float, respectively). The idea is that while firm performance in general would increase public float, it is only when firms exceed the threshold of $75 million that the filing deadline change applies. Thus, by comparing firms that experienced a filing deadline change with other firms that experienced public float increases of a similar magnitude but fell just short of the $75 million and $700 million thresholds, we can estimate the (causal) impact of the deadline change. The advantage of this design is that the control group also experiences similar changes in (and levels of) public float as the affected firms, and the only thing different about the latter is the accompanying filing deadline change. This helps us to ascribe a causal interpretation to our findings. Results from these tests corroborate our earlier results that accounting resources are negatively associated with the likelihood of a future restatement. Specifically, the likelihood of a restatement for firms in the bottom decile of accounting resources is 0.082 (1 in every 12 10Ks), and for those in the top decile it is 0.045 (1 in every 22 10Ks).

We perform two cross-sectional tests to link our findings to the accounting fraud literature. First, we examine variation in our findings across the business cycle. Anecdotal evidence (e.g., *The Economist*, 2002) as well as academic research (e.g., Povel et al., 2007; Kedia and Philippon, 2009; Wang et al., 2010) purport that firms have greater incentives to perpetuate accounting fraud during good times, which is subsequently revealed during downturns. A missing piece not acknowledged in prior work is that accounting resources are also likely to show disparities between booms and busts, and that these disparities (rather than fraud incentives) could drive the uptick in misreporting. For example, boom times are when firms take on new projects, make more sales, attract new customers, launch new products, undertake more acquisitions, all of which put greater constraints on the accounting staff to reflect these activities in the financial statements accurately, and also file these reports on time. We contend that the marginal cost of devoting resources to increasing the accounting function is greater during boom periods when these resources have more productive alternatives. As a result, accounting staffs are likely to be stretched thin during booms, as compared to busts. Thus, it is likely that extant associations between the business cycle and accounting fraud could be driven by variation in accounting resources across the cycle. Our results provide evidence consistent with our hypothesis – (i) accounting resources are indeed stretched thin during boom periods as compared to bust, and (ii) the association between accounting resources and subsequent restatements is more pronounced during booms as compared to busts. Moving from the bottom decile to the top decile of accounting resources reduces the likelihood of a restatement by 4.6% during booms but by only 2.3% during busts.

Second, we split restatements into accounting errors and irregularities. Since accounting resources influence misreporting incentives (as per our model), the predictive ability of misreporting variables for predicting irregularities should weaken once we control for accounting resources. This is precisely what we find – the explanatory power of the incentive variables drops from 5% significance (based on a partial *F* test) to insignificance once we control for accounting resources in the model that predicts accounting irregularities. Further, accounting resources do not have an independent effect on irregularities, controlling for the incentive variables. Turning to restatements due to errors, we find a strong predictive power for accounting resources – consistent with our model where low accounting quality manifests in more frequent reporting errors. We also find, consistent with our prediction, that controlling for accounting resources does not diminish the explanatory power of the incentive variables in the errors subsample.

Using filing timeliness and spelling errors to capture the firm's investment in accounting resources has advantages over other commonly used metrics for EQ such as internal control deficiencies over financial reporting (ICDs) or restatements. While the literature has documented a statistical relation between internal control deficiencies over financial reporting (ICDs) and firm size, adequacy of the accounting staff, and financial constraints (Ashbaugh-Skaife et al., 2007; Doyle et al., 2007), interpreting these associations is confounded by the sample selection bias of first detecting the ICD and then disclosing it (see Ashbaugh-Skaife et al., 2007). In other words, ICDs must first exist, then managers must discover them, and finally decide to disclose them. So the disclosure of an ICD is a noisy and possibly biased measure for the quality and quantity of the accounting staff because managers exert considerable discretion in how much to invest searching for ICDs and in disclosing ICDs. Spelling errors and filing timeliness are more direct measures of firms' investments in accounting resources that are not confounded by managers' and auditors' incentives to discover and report the ICDs. The same criticism applies to using restatements as a measure of the investment in accounting systems.

This paper makes three contributions to the EQ literature in general and the restatements literature specifically. First, we introduce explicitly the notion that the choice of EQ is firm-value maximizing and offer a simple framework of that choice in the absence of managerial opportunism. Second, we provide a structural model of how restatements depend on both the firm-value-maximizing investment in EQ and managerial incentives to misreport. And third, we offer two novel measures of the optimum investment in EQ – 10-K filing timeliness and spelling errors.

2. Literature Review

Restatements can be triggered by the SEC, the firm, or the firm's auditor. A large literature examines the causes and consequences of accounting restatements (see Dechow et al., 2010). A variety of factors have been offered to explain the increasing frequency of restatements including an increased focus on meeting analyst forecasts, management and auditors have become more conservative in their restatement decisions, the increasing number and complexity of accounting standards, SOX 404 reviews uncover more errors, and more complex business transactions lead to more restatements (Plumlee and Yohn, 2010). To this list, we would add the increasing competition, lower expected profitability, and shortened firm life expectancies that cause firms to optimally invest less in accounting resources.[4] Supporting our view, Kinney and McDaniel (1989) report that restating companies are smaller, less profitable, are more leveraged, and are growing slower than their non-restating industry-matched peers. But unlike us, Kinney and McDaniel (1989, p. 72) attribute these correlations to earnings management by financially weak firms, and/or auditors requiring correction of errors of firms with greater risk of failure. Choudhary et al. (2016) examine errors deemed immaterial by management, and they find that immaterial errors predict future immaterial errors, material errors, and material weakness assessments. Their evidence supports our hypothesis that some firms optimally invest less in their accounting systems.

The accounting literature generally recognizes that accounting restatements result from intentional misreporting or unintentional errors or ambiguities in accounting rules that lead to errors (Dechow et al. 2010). Based on their reading of individual restatement disclosures reported in 8-Ks filed between 2002 and 2005, Hennes et al. (2008) classify restatements that arise from intentional misreporting if the restatement is associated with a subsequent investigation by the audit committee, Department of Justice, or the SEC, or if the disclosure contained the words "fraud" or "irregularity." All other restatements are classified as "errors." Their final sample is 630 firms. They classify 24 percent of the restatements as intentional misstatements and 76 percent as unintentional errors. Further, Hennes et al. (2008) document that unintentional restatements (i.e., errors) are associated with smaller negative announcement returns than intentional misrepresentations.

[4] See Fama and French (2004), Irvine and Pontiff (2009), and Owens et al. (2015).

Based on data between 2003 and 2006, Plumlee and Yohn (2010) use the restating companies' disclosures about the restatements, rather than 8-K filings like Hennes et al. (2008), and find that 57 percent of the 3,744 restatements are attributed to internal errors. Three percent result from the complexity of the transaction. Another 37 percent are attributed to the standard lacked clarity or required the use of judgment in applying the standard. Based on their findings, 97 percent of all restatements can be attributed to low levels of accounting staff human capital – insufficient accounting resources either made internal errors, were unable to apply GAAP to record complex transactions, or did not properly apply an accounting standard. Only 3 percent were due to intentional manipulation. Plumlee and Yohn (2010) conclude, "the majority of restatements are attributed to basic books and record deficiencies within the company and to simple misapplications of generally accepted accounting standards."[5] Plumlee and Yohn (2010) document fewer intentional restatements than Hennes et al. (2008) because Hennes et al. (2008) eliminate restatements with no net income effect.

After correcting for various sources of bias in previous studies Lobo and Zhao (2013) find a negative relation between audit fees and the restatement of the subsequent annual report. They use all restatements and do not control for intentional vs. unintentional restatements because their research question focuses on whether audit effort affects misreporting, both intentional and unintentional. The authors employ the Dechow et al. (2011) misstatement detection model to test whether audit effort, measured using audit fees, is negatively correlated with restatements. Implicit in their tests is that the auditor chooses the amount of audit effort, proxied by audit fees, and more auditor effort generates more restatements. But the firm chooses the auditor (Big N, a regional auditor, or local firm) and negotiates the scope of the audit and audit fees. To the extent

[5] Plumlee and Yohn (2010) recognize that their findings are based on managements' disclosures about what caused the restatement and these disclosures may be strategic.

managers choose the optimum amount to invest in accounting resources, including audit services, then lower audit fees are associated with the likelihood of subsequent restatements, consistent with our prediction.[6]

Aier et al. (2005) find that more qualified CFOs, those with CPAs, MBAs, and more experience, are less likely to restate earnings. No reason is given as to why some firms choose more qualified and others choose less qualified CFOs. Nor do the authors argue that the qualification of the CFO is chosen to maximize firm value. Nonetheless, their results are consistent with our model that firms choose the optimum investment in accounting resources. Presumably, less qualified CFOs are less expensive than more qualified CFOs, but make more mistakes (i.e., "you get what you pay for").

Studies that examine the determinants of restatements employ a variety of "control" variables.[7] For example Burns and Kedia (2006), whose primary variables of interest are equity incentives, use lagged market value of equity, market-to-book, leverage and industry dummies as control variables. Unlike Kinney and McDaniel (1989), they find that larger firms are more likely to restate. Consistent with Kinney and McDaniel (1989), Burns and Kedia (2006) report more leveraged firms are more likely to restate. Market-to-book is insignificant in the Burns and Kedia logistic regressions predicting restatements. Efendi et al. (2007) find that restatements are more likely for firms having more CEO in-the-money stock options, facing more binding interest-coverage debt covenants, raising new debt or equity capital, or having the CEO serve as board

[6] One conceptual problem with the Lobo and Zhao (2013) study is that most restatements are identified by the company, not the auditor. Taub (2006) reports anecdotal data that about half of the restatement disclosures made by firms did not report how the error was found. But when information was provided, about half of the disclosures indicate that the company identified the error and only about 15% cited the external audit. http://www.sec.gov/news/speech/2006/spch111706sat.htm

[7] Dechow et al. (2011) point out that COMPUSTAT backfills misstated numbers when a company files an amended 10-K. So we need to use the "as reported" numbers for restating firms.

chair. Richardson et al. (2002) report that restatements are associated with new debt and equity financings, higher price-earnings and market-to-book multiples, leverage, and longer strings of consecutive positive earnings growth.

Boland et al. (2015) examine whether the regulation requiring accelerated filing deadlines increases the likelihood of a future restatement, and find evidence of a temporary increase in the likelihood of a restatement for accelerated filers when the deadline decreased from 90 to 75 days. They do not find a similar increase in the likelihood of a restatement for large accelerated filers when the filing deadline decreased from 75 to 60 days. Similar to the Boland et al. (2015) paper, we predict that the regulatory change in filing deadlines is associated with an increase in the likelihood of a restatement. However, our paper differs by using filing timeliness as our primary variable of interest. In addition, Arif et al. (2016) find that accelerating the filing deadlines delayed the release of voluntary firm disclosures. Specifically, they find that in response to shorter filing deadlines, firms delayed their earnings announcements in order to coincide with filing the 10-K.

The extant literature often assumes the restatement proxies for earnings management (i.e., intentional misstatements) and then posit a variety of variables including CEO compensation, composition of the board of directors, and non-audit fees to capture managers' incentives to misreport. After reviewing the literature, Dechow et al. (2010, p 374) find the results mixed and conclude, "The generally weak and mixed evidence across the determinants of restatements suggests that they are not a reliable indicator of intentional misstatements." The authors then suggest that "careful screening of the sample has the potential to yield more powerful tests of the determinants and consequences of unintentional versus intentional misstatements" (p. 374-5).

In summary, while the literature recognizes that restatements can results from unintentional errors or intentional misstatements, most of the control variables proposed are ad hoc. Also, a variety of variables are associated with restatements such as size, leverage, and profitability, but inconsistent signs are reported across studies. Finally, there is no widely used empirical model of the determinants of accounting restatements.

3. A Conceptual and Corresponding Empirical Model of Restatements

Based on the prior restatements literature and our maintained assumption that managers choose the firm-value-maximizing investment in accounting systems, this section develops a conceptual framework and corresponding empirical model of the determinants of accounting restatements. It is common in the literature for researchers to estimate logistic models of restatements that have the following general form:[8]

$$R_t = v_0 + v_1 SIZE_{t-1} + v_2 ROA_{t-1} + v_3 MTB_{t-1} + v_4 LEV_{t-1} + v_5 NCAP_{t-1} + v_6 GROW_{t-1} + v_7 INCENT_{t-1} \quad (1)$$

Where R=1 if the firm restates in year t, 0 otherwise. SIZE is firm size in year t-1, ROA captures firm profitability in year t-1, MTB is market to book in year t-1, LEV is leverage in year t-1, NCAP =1 if new capital is raised in year t-1, GROW is growth in year t-1, and INCENT represents other incentive variable(s) of particular interest to the study, such as equity compensation incentives. The lag structure in model (1) implicitly assumes that innate firm characteristics and incentives in year t-1 cause intentional and unintentional errors in year t-1 financial statements that get detected and reported as restatements in year t.

Besides INCENT, the other independent variables in equation (1) capture other misreporting incentives. Large firms (SIZE) are assumed to have better internal controls that deter managerial malfeasance (Kinney and McDaniel, 1989). Less profitable firms (ROA) have incentives to boost

[8] See Burns and Kedia (2006) and Armstrong et al. (2010) for a critique of the existing literature.

profits (Kinney and McDaniel, 1989). Firms with high valuations (MTB) have incentives to maintain these high values (Burns and Kedia, 2006). Highly leveraged firms are assumed to be closer to their debt covenants and hence more likely to misrepresent; and firms issuing new capital have incentives to misreport (Burns and Kedia, 2006).

Equation (1) suffers from several theoretical problems. First, it ignores the firm's optimal investment in accounting systems, creating a potential correlated omitted variables bias. Second, the incentive to misreport, INCENT, is endogenous. Managers are more likely to misreport when they know the firm has weak financial controls, and hence the likelihood of detection is lower. Third, some of the independent variables are subject to alternative interpretations. Low growth and less profitable firms have less incentive to invest in accounting systems because they have shorter expected life spans to recoup the initial investments in these systems. High quality accountants have less incentive to accept jobs in less profitable firms with short horizons. So, finding low growth, less profitable firms are more likely to restate could be due to either managerial malfeasance or less investment in accounting systems causing more unintentional errors.

In this paper, we argue that firms choose the firm value-maximizing investment in accounting systems. Appendix A offers a simple model whereby firm-value maximizing managers in smaller and less profitable firms invest less in accounting systems and staff leading to more errors, restatements, and frauds. Incorporated into this cost-benefit analysis is the expected cost of accounting fraud. Firms that optimally choose lower investments in their accounting systems expect more frauds, including intentional misreporting not just from senior executives but lower level employees. Some financial misrepresentations are committed by divisional presidents, in other words, not all financial misrepresentations are promulgated by CEOs and CFOs.

We propose the following conceptual framework of the relations between the optimal investment in accounting systems (ACC_RES), spelling errors (SPELL_ERR), filing timeliness (FILE_TIME), incentives to misreport (INCENT) and restatements, R.[9]

$$ACC\_RES_t = a_0 + a_1SIZE_t + a_2ROA_t + a_3CMPLX_t + a_4GROW_t \quad (2)$$

$$FILE\_TIME_t = b_0 + b_1ACC\_RES_t \quad (3)$$

$$SPELL\_ERR_t = c_0 + c_1ACC\_RES_t \quad (4)$$

$$INCENT_t = d_0 + d_1ACC\_RES_t + d_2ROA + d_3NCAP_t + d_4LEV_t + d_5MTB_t + d_6GROW_t \quad (5)$$

$$R_t = e_0 + e_1ACC\_RES_{t-1} + e_2INCENT_{t-1} \quad (6)$$

Equation (2) is based on the model in Appendix A and states that the optimum investment in accounting resources (ACC_RES), depends on firm size (SIZE), profitability (ROA), firm complexity (CMPLX), and growth (GROW).[10] We expect, a_1, a_2, a_3, and $a_4 > 0$. Equations (3) and (4) posit that filing timeliness (FILE_TIME) and spelling errors (SPELL_ERR) are related to the amount invested in accounting systems (ACC_RES). We expect $b_1 > 0$ and $c_1 < 0$.

Based on the restatements literature, Equation (5) states that the incentives to misreport (INCENT) are a function of the amount invested in the accounting system (ACC_RES), profitability (ROA), market-to-book (MTB), leverage (LEV), new capital raised (NCAP), and growth (GROW).[11] From the prior literature, we predict d_3, d_4, d_5 and $d_6 > 0$ and d_1 and $d_2 < 0$.

Finally, equation (6) explicitly recognizes that restatements in year t can arise from either unintentional errors or intentional misrepresentations in year t-1. Unintentional errors arise from

[9] To simplify the notation, all the error terms in equations (2) – (6) are omitted.

[10] For parsimony the model in Appendix A considers just size and profitability. In equation (2) we introduce two dimension of firm size: scale and complexity. Likewise, equation (2) considers two dimensions of profitability: ROA and growth.

[11] Excluded from equation (5) are misreporting incentives arising from compensation. We omit these incentives because (i) compensation data would limit sample size and thereby reduce power, (ii) introduce potential sample selection bias, and (iii) cause other methodological problems (Armstrong, et al. 2010). However, we include these in additional tests to verify that our inferences are robust to controlling for compensation-based reporting incentives.

low investments in accounting systems (ACC_RES); whereas, intentional misrepresentations arise when managers have strong incentives to misreport (INCENT). In other words, equation (6) assumes that the quantity of the accounting systems (ACC_RES) in year t-1 and the misreporting incentives (INCENT) in year t-1 drive intentional misrepresentations and unintentional errors in the firm's fiscal year t-1 reports, which are then detected and restated in year t. We expect $e_1 < 0$ and $e_2 > 0$.

Unfortunately, this system of equations cannot be directly estimated because ACC_RES and INCENT are unobservable. One can derive a model of restatements containing only observable variables by first substituting equation (2)'s RHS variables into equation (5) for ACC_RES, and then substituting that equation's RHS variables into equation (6) for INCENT and again substituting equation (2)'s RHS variables of into equation (6) for ACC_RES. These straightforward algebraic manipulations result in the following empirical model:

$$R_{t+1} = w_0 + w_1 SIZE_t + w_2 ROA_t + w_3 CMPLX_t + w_4 GROW_t + w_3 NCAP_t + w_4 LEV_t + w_5 MTB_t \quad (7)$$

where:

$$w_1 = e_1a_1 + e_2d_1a_1 > \text{or} < 0 \quad (8)$$

$$w_2 = e_1a_2 + e_2(d_1a_2+d_2) < 0 \quad (9)$$

$$w_3 = e_1a_3 + e_2d_1a_3 < 0 \quad (10)$$

$$w_4 = e_1a_4 + e_2(d_1a_4+d_6) > \text{or} < 0 \quad (11)$$

If one were to estimate equation (7), the predicted signs on SIZE (w_1) and on GROW (w_4) can be positive or negative. The ambiguity arises from mixing unintentional errors with intentional misstatements. If all restatements were intentional misrepresentations, then $e_1 = 0$ in equation (6), and the predicted sign on the SIZE coefficient (w_1) in equation (7) is negative. However, the predicted sign on GROW coefficient (w_4) in equation (7) remains ambiguous. The basic point of this analysis is that mixing intentional and unintentional accounting errors in the sample (and assuming investment in accounting resources is endogenous) one cannot draw inferences about managerial malfeasance from estimating simple reduced form regressions like equations (1) or (7) because variables that are used to capture managerial incentives to misreport also capture managerial incentives to choose the optimum investment in accounting resources.

One way to address the inherent identification problem that exists with estimating equation (7) is to use spelling errors (SPELL_ERR) and filing timeliness (FILE_TIME) in equations (3) and (4), respectively, as instruments for the optimum investment in accounting systems (ACC_RES). Based on equations (3) and (4) we can assume that:

$$ACC\_RES_t = f_0 + f_1FILE\_TIME_t + f_2SPELL\_ERR_t \tag{12}$$

where $f_1 > 0$ (based on $b_1 > 0$ in equation (3)) and $f_2 < 0$ (based on $c_1 < 0$ in equation (4)). By substituting equation (12) into equations (5) and (6) for ACC_RES we can derive the empirical model:

$$R_{t+1} = g_0 + g_1FILE\_TIME_t + g_2SPELL\_ERR_t + g_3ROA_t + g_4NCAP_t + g_5LEV_t + g_6MTB_t + g_7GROW_t \tag{13}$$

where: $g_1 = f_1(e_1 + e_2d_1) < 0$, $g_2 = f_2(e_1 + e_2d_1) > 0$, $g_3 = e_2d_2 < 0$, $g_4 = e_2d_3 > 0$, $g_5 = e_2d_4 > 0$, $g_6 = e_2d_5 > 0$, and $g_7 = e_2d_6 > 0$. Using equation (13) as the empirical model to estimate restatements has the advantage of generating unambiguous predicted signs on the coefficients for all the independent variables in the model.

Another testable empirical prediction from our structural system of equations (2) – (6) is that for restatements caused entirely by low investments in accounting systems (ACC_RES) e_1 in equation (6) is greater than zero and $e_2 = 0$. In regressions containing only unintentional errors the coefficients on ACC_RES should be statistically significant and the coefficients on the other variables in equation (13) that capture incentives to misreport (ROA, NCAP, LEV, MTB, GROW) should be statistically insignificant. To implement this test, we rely on Hennes et al. (2008) who find that restatements they classify as intentional misstatements have much larger negative stock returns in the 14 days surrounding the restatement announcement than restatements they classify as unintentional misstatements. Stated differently, we assume that the market on average can separate restatements caused by errors ($e_1 > 0$ and $e_2 = 0$ in equation (6)) from restatements caused by intentional misstatements ($e_1 = 0$ and $e_2 > 0$ in equation (6)).

We first calculate the CAR (-7, +7) (following Hennes et al., 2008) centered on the restatement announcement date (CAR = firm return – market). All restatements are sorted into terciles based on their CAR. The tercile with the largest negative CARs likely reflects fraud/malfeasance ($e_2 > 0$). Equation (13) is estimated separately for tercile 1 (which likely capture frauds) and the other terciles (which likely capture errors). We expect to observe the predictive ability of spelling errors and filing timeliness to be higher in the errors sample as compared to the fraud sample. We implement these tests in Table 7 and find consistent evidence. Likewise, the statistical significance of the coefficients on the misstatement incentives (*ROA*, *RET*, *NCAP*, *LEV*, *MTB*, *GROW*, *SIZE*) should decline across the terciles with progressively less negative CARs.

4. Data

4.1 Sample Selection

The sample selection process is described in Table 1. We download all 10Ks available on the SEC EDGAR website between August 31, 1993 and May 31, 2014, a total of 135,120 10Ks (see Appendix C for a detailed description of this procedure). We use textual analysis in Python to analyze the number of spelling errors in each 10K, and we successfully perform this procedure for

134,300 10Ks (see Appendix D for a detailed description of the calculation of spelling errors using Python). We join the sample of 10Ks to Audit Analytics (AA) to obtain the filing status for each firm-year, resulting in a sample of 117,356 10Ks with filing status data. We join the sample of 10Ks to AA by matching the CIK number and filing date on the front of each 10K (obtained via Python) to the company_fkey and filing date of each 10K in AA. We obtain accounting data and business segments data from Compustat using the historical CIK to GVKEY linktable available in WRDS SEC Analytics. The sample of 10Ks with accounting and business segment data consists of 94,701 observations. Because coverage in the AA non-reliance restatement data is limited prior to 2000 and we capture restatements in year t+1, we further reduce the sample to 81,631 10Ks filed between 1998-2014. We eliminate duplicate 10Ks in the same fiscal year, obtain stock return and volatility data from CRSP, and eliminate 10Ks belonging to asset backed securities, REITs, shell companies, blank check companies, non-operational companies, funds, and trusts (as provided by AA), resulting in 41,284 10Ks. Finally, we truncate observations at the 5% level of filing timeliness (i.e., between -15 days and 24 days), resulting in a final sample of 38,185 firm-year observations.

4.2 Filing Timeliness

We calculate filing timeliness as the difference between the firm's filing deadline date and the date the firm filed its 10K. The firm's filing deadline date is based on its filing status (e.g., large accelerated filer, accelerated filer, small reporting company). The SEC requires firms to file their financial statements within a filing window which is based on the firm's filing status, and this window has changed over time. Prior to December 15, 2003, all firms had 90 days after their fiscal year end to file their financial statements with the SEC. From December 15, 2003 to December

15, 2006, large accelerated filers and accelerated filers had 75 days, and non-accelerated filers had 90 days. After December 15, 2006, large accelerated filers have 60 days, accelerated filers have 75 days, and non-accelerated filers have 90 days.[12]

4.3 Research design

Given equation (13), which shows the relation between the likelihood of a restatement and accounting resources, we estimate the following logistic regression:

$$\begin{aligned} Pr\left[RESTATE_{i,t+1}=1\right] = {} & \alpha_0 + \beta_1 ACC\_RES_{i,t} + \beta_2 ROA_{i,t} + \beta_3 RET_{i,t} + \beta_4 NCAP_{i,t} + \beta_5 LEV_{i,t} \\ & + \beta_6 SGR_{i,t} + \beta_7 RETVOL_{i,t} + \beta_8 SIZE_{i,t} + \beta_9 SIZE\_SQ_{i,t} + \sum_i Ind_i \\ & + \sum_t Year_t + \varepsilon_{i.t+1} \end{aligned} \tag{14}$$

where, *RESTATE* is an indicator variable that takes the value of 1 if firm *i* restates its financial statements in year *t+1*; *ACC_RES* represents accounting resources measured in year *t*, *ROA* denotes return on assets in year *t*, *RET* indicates the firm's excess market return in year *t*, *NCAP* is an indicator that is set to one when the firm issued debt or equity greater than 5% of the firm's total assets in year *t*; *LEV* denotes the firm's book leverage in year *t*; *SGR* denotes annual sales growth in year *t*, *RETVOL* represents stock return volatility in year *t*; *SIZE* is (the log of) total sales and *SIZE_SQ* is the squared term of *SIZE* to capture size-related nonlinearities. Industry fixed effects (defined at the 2-digit SIC industry level) and year fixed effects are also included to control for cross-industry and inter-temporal effects. We cluster the robust standard errors by firm.

Our framework predicts a negative coefficient on β_1 since firms with greater investments in accounting resources are associated with a lower likelihood of a restatement.

[12] For further information, please see the SEC's website, http://edgar.sec.gov/answers/form10k.htm.

4.4 Descriptive statistics

Table 2 presents descriptive statistics of the sample. The mean *RESTATE* of 0.065 indicates that 1 in every 15 10Ks is restated. The median firm files its financial statements 2 days before the filing deadline (*FILE_TIME*). This variable has been truncated at the 5% tails (i.e., -15 and 24 days) to ensure that we exclude shell companies and other pseudo entities. Close to one-third of the sample makes either a debt or an equity issuance during the year (the mean *NCAP* = 0.338), which is consistent the positive sales growth in these firms (mean *SGR* = 0.169) and the presence of growth opportunities (mean *MTB* = 2.034). The mean value of 5.731 for *SIZE* indicates annual sales of $308 million ($e^{5.731}$).

Within the sample of Execucomp firms, the mean *CEODELTA* of 5.318 corresponds to a $203,000 increase in wealth ($e^{5.318}$-1) for a 1% increase in the company's stock price. The sensitivity of the CEO's wealth to a 1% increase in the firm's stock return volatility (*CEOVEGA*) is much lower at $39,000 ($e^{3.657}$). Comparable numbers for the CFO are $35,000 for the stock return and $13,000 for the volatility. These numbers compare closely with those in Chava and Purnanandam (2010).

Table 3 presents correlations among the variables. Most notably, *FILE_TIME* and *RESTATE* are significantly negatively correlated (-0.045), and *SPELL_ERR* and *RESTATE* are positively (but not significantly) correlated. The composite measure of accounting resources (*ACC_RES*) is negatively and significantly correlated with *RESTATE* (-0.036, 1% significance). *RESTATE* is significantly negatively correlated with *ROA* (-0.048) and *MTB* (-0.050) and significantly positively correlated with *LEV* (0.033) and *RETVOL* (0.048). *ACC_RES* is significantly positively correlated with *ROA* (0.054), *RET* (0.054), *CMPLX* (0.020), and *SIZE* (0.052) and significantly negatively correlated with *RETVOL* (-0.019).

5. Results

5.1 Graphical evidence

We begin with graphical evidence. In particular, we create a dichotomous classification of our sample firms (Late vs. Timely) depending on whether they file their regulatory reports after or before the filing deadline. We then plot the mean likelihood of a restatement in each of these bins. These results are presented in Figure 1. Consistent with our main hypothesis, the occurrence of restatements is more frequent in the sample of late filers as compared to timely ones. In particular, the mean occurrence of restatements within the group of late filers is 0.085 (1 in every 12 10Ks) as opposed to 0.062 (1 in every 16 10Ks) for timely filers.

5.2 Multivariate evidence

5.2.1 Logistic regression

Table 4 presents results of our logistic regression of eq. (14). We begin in model (1) with the dichotomous variable (*TIMELY*) which takes the value of 1 if firms file their financial statements on or before the filing deadline and 0 if they are late. Consistent with the graphical evidence, the coefficient on *TIMELY* is negative and significant at the 1% significance level. As discussed above, the likelihood of a restatement is 0.085 (1 in every 12 10Ks) for late filers as opposed to 0.062 (1 in every 16 10Ks) for timely filers.

Models (2) and (3) use the two proxies for accounting resources – filing timeliness and spelling errors, respectively. As indicated by the univariate correlations, the coefficient on *FILE_TIME* is negative and significant at the 1% level while that on *SPELL_ERR* is negative but insignificant. In terms of marginal effects, the likelihood of a restatement in firms that are five days late is 0.076 (1 in every 13 10Ks) as compared to 0.064 (1 in every 16 10Ks) in firms that are

5 days early. Model (4) uses the composite measure of accounting resources and finds a negative and significant coefficient on *ACC_RES* (-0.073 with a *p* value<0.01).

The next set of models examines how the explanatory power of the misreporting variables changes once we control for *ACC_RES*. Model (5) includes the covariates used in prior studies viz., profitability (*ROA*, *RET*), capital issuance (*NCAP*), leverage (*LEV*), growth opportunities (*MTB*, *SGR*), operating complexity/uncertainty (*CMPLX*, *RETVOL*) and firm size (*SIZE*) and also year and industry fixed effects (defined at the 2-digit SIC code level). We regard *ROA*, *RET*, *NCAP*, *LEV*, *MTB*, *SGR*, and *SIZE* as the misreporting incentive variables (based on our theoretical model – see equation (5)). Several of these variables are significantly associated with restatements in the direction consistent with prior studies. For example, *ROA* is negative and significant (although *RET* is not) indicating that more profitable firms are less likely to manipulate their financial statements. Similarly, *NCAP*, *LEV* and *SGR* are all positive and significant. The partial *F* state of these incentive variables is 60.71, and significant at the 1% level.

Model (6) presents results after including *ACC_RES* as an additional variable. The coefficient on *ACC_RES* remains negative and significant, indicating that the explanatory power of accounting resources is not subsumed by the other variables. More relevant to our theoretical model, the explanatory power of several of the misreporting incentive variables diminishes in significance once we control for *ACC_RES*. For example, the coefficient on *ROA* decreases from -0.295 in model (5) to -0.257 in model (6), which represents a 13% decrease in economic magnitude. Similarly, *LEV* and *SGR* each decreases by 5% (0.536 to 0.511, and 0.098 to 0.093 respectively). The overall explanatory power of the incentive variables also decreases as seen by the lower partial *F* stat of 55.37 in model (6) (still significant at the 1% level). We interpret this evidence as consistent with our theoretical model where firms optimally choose accounting resources as a function of profitability, growth and size. Thus, the interpretation of misreporting attributed to these firm-characteristics might be misleading as these variables also determine firms' optimal investment in accounting resources (which in turn drive restatement likelihoods).[13] Overall, these results indicate a robust inverse association between a firm's accounting resources and the likelihood of an earnings restatement.

5.2.2 Ruling out audit verification effects

While the results in Table 4 indicate a negative association between firms' investments in accounting resources and the likelihood of future restatements, it is possible that some of this effect is driven by the role of audit verification or SOX 404 audits.[14] We address this concern by controlling for the quality of audit verification in Table 5. Specifically, we control for the log of audit fees (scaled to millions) in Model (1), and unexplained audit fees (defined as the residual of a regression of log audit fees on firm size) in Model (2). In both specifications, we continue to find that *ACC_RES* is negatively and significantly associated with the likelihood of a future restatement. Similar to Table 4, the coefficient on *ACC_RES* translates into a mean likelihood of a restatement of 0.074 (1 in every 14 10Ks) in firms at the bottom decile of accounting resources as compared to 0.058 (1 in every 17 10Ks) for those in the top decile. In sum, our results are robust to controlling for the role of auditor attestation and SOX 404 compliance.

5.2.3 Endogeneity of filing timeliness

The results thus far, while suggestive, do not provide a conclusive link from accounting resources to restatement likelihood as they ignore the endogeneity of accounting resources. It could be that omitted firm-level factors correlated with malfeasance are also correlated with accounting

[13] Our results are robust to including a non-linear size effects.
[14] The choice of auditor and the amount invested in external audit fees is part of the optimal investment in accounting resources.

resources. For example, one could argue that firms that are unable to file their financial statements in a timely fashion might choose to acquire more accounting resources, leading to reverse-causality. In addition, one could argue that firms with “nothing to hide” are more prompt in filing their financial statements, and are also less likely to restate these financials, thereby causing the association that we detect. While we do not observe a significant effect on the misreporting incentives variables (which weakens this alternative interpretation), it could be that the incentives variables are noisy.

A potential solution to this endogeneity problem is to build a structural model. While the advantage of this approach is the ability to test detailed predictions of the theory, the shortcoming is the difficulty in constructing such a structural model: one would have to include all possible channels through which other variables jointly affect both filing timeliness and restatement likelihood, and this is clearly not possible. We pursue two alternative strategies to mitigate endogeneity. First, we use an instrumental variables approach. The validity of this approach, however, hinges on finding an instrument that not only explains accounting resources (i.e., fulfils the relevance criterion) but also does not affect restatement likelihood directly through any alternative channels (i.e., the exclusion criterion). We use the filing window as our instrument, and present these 2SLS results in Models (1) and (2) of Table 6. As noted previously, the filing window is the number of days that a firm has to file their financial statements after the fiscal period end. The SEC mandates firms’ filing window based on the firms’ filing status (e.g., large accelerated filer, accelerated filer, small reporting company), and this window has decreased over time for large accelerated and accelerated filers (see Arif et al., 2016; Boland et al., 2015; Gao, 2015; Lambert et al., 2016). Our rationale is that a longer filing window is tantamount to greater accounting resources as it puts less strain on the firm’s accounting staff. Conversely, shorter windows stretch the firm's accounting resources to a greater extent. The other advantage of the filing window is that it is not endogenously chosen by the firm.[15]

Model (1) estimates the first-stage of the 2SLS where we regress *ACC_RES* on the filing window (*WINDOW*) and all the other control variables (including the fixed effects) from the restatement prediction model. The coefficient on *WINDOW* is not only positive and significant at the 1% level, but also comfortably exceeds the threshold to qualify as a valid instrument as seen by the *p* value of the partial *F* stat on *WINDOW* (Stock, Wright and Yogo, 2002). Model (2) presents the second-stage restatement prediction model where we replace *ACC_RES* with its predicted value from the first-stage (*ACC_RES_PRED*). The coefficient on *ACC_RED_PRED* is negative and significant (at the 10% level), indicating the robustness of our results. In terms of economic significance, firms in the bottom decile of accounting resources have a restatement likelihood of 0.083 (1 in every 12 10Ks) as compared to 0.052 (1 in every 19 10Ks) those in the top decile.

The second method we use to address endogeneity is to exploit changes in filing deadlines as a function of changes in the SEC requirements and a firm's filing status. By using changes in the filing deadline as an instrument, we exploit cross-sectional variation in filing deadlines across firms within each year as well as time-series variation within each firm where certain firms "cross-over" to an accelerated filing status when they fulfil certain conditions. As described above, the SEC changed the filing deadline beginning in December 15, 2003. As of December 15, 2003, accelerated filers have to file their 10K within 75 days (previously 90 days), and as of December 15, 2006, large accelerated filers have to file their 10K in 60 days. Although the definition of accelerated and large accelerated filers contains several different criteria, the primary condition

[15] Controlling for firm size in all the specifications ensures that we capture changes in filing deadlines that are likely to be caused by changes in market cap.

for firms to be classified as accelerated and large accelerated filers is based on their public float as of the end of the second fiscal quarter (see Gao, 2015; Lambert et al., 2016). Specifically, firms with public float greater than $75 million are categorized as accelerated filers, and firms with public float greater than $700 million are categorized as large accelerated filers.

We employ a matched-sample design (similar to a discontinuity design) around changes in the filing deadline using a narrow window around the cutoff for accelerated and large accelerated filers (i.e., $75 million and $700 million in public float, respectively). The idea is that while firm performance in general would increase public float, it is only when firms exceed the threshold of $75 million that the filing deadline change applies. Thus, by comparing firms with a filing deadline change with other firms that experienced public float increases of a similar magnitude but fell just short of the $75 million threshold, we can estimate the (causal) impact of the deadline change. The advantage of this design is that the control group also experiences similar changes in public float as the affected firms, and the only thing different about the latter is the accompanying filing deadline change. This helps us to ascribe a causal interpretation to our findings.

Models (3)-(6) of Table 6 present these results. We restrict the sample only to firms with changes in their filing deadlines and a matched sample based on similar market values of equity (MVE). To ensure that we have the correct counter-factual, we include control firms with similar changes or levels of MVE as the affected firms, but those without a similar change to their filing deadlines. (Following prior studies, we use MVE as a proxy for public float). Treated firms are defined as follows: (i) accelerated filers in the first year of compliance with shorter deadlines as of December 15, 2003, (ii) large accelerated filers in the first year of compliance with shorter deadlines as of December 15, 2006, (iii) firms that change from non-accelerated filers to accelerated filers after December 15, 2004, and (iv) firms that change from accelerated to large accelerated filers after December 15, 2007. Matched control firms for each of these four types of treated firms are as follows: (i) firms that are not accelerated filers with similar MVE (MVE within five percent of an accelerated filer) in the first year of shorter deadlines, (ii) firms that are not large accelerated filers with similar MVE (MVE within five percent of a large accelerated filer) in the first year of compliance with shorter deadlines, (iii) firms that are not accelerated filers with a similar change in MVE (defined as the yearly change in MVE in the second fiscal quarter that is within five percent of the change in MVE for a firm that became an accelerated filer) after December 15, 2004, (iv) firms that are not large accelerated filers with a similar change in MVE (MVE that this within five percent of the change in MVE for a firm that became a large accelerated filer) after December 15, 2007. In summary, Models (3)-(6) in Table 6 contain all of the treated and control firms described above.

Model (4) of Table 5 presents the matched-sample design tests based on changes in filing deadlines. The negative coefficient on *ACC_RES_PRED* remains significant at the 10% level. Firms in the bottom decile of accounting resources have a greater likelihood of a future earnings restatement of 0.083 (1 in every 12 10Ks) as compared to 0.044 (1 in every 22 10Ks). Under the maintained assumption that firms cannot choose the exact change in their public float (and their filing status), we can attach a causal interpretation to the effect of filing timeliness on lowering the likelihood of a restatement.

The last specification, Model (6), includes the yearly change in MVE as an additional covariate to control for any residual differences between the affected and unaffected firms due to imperfect matching. Results from this two-stage analysis again confirm our primary evidence – accounting resources are negatively associated with the likelihood of a future restatement. All of the specifications in Table 6 include unexplained audit fees, meaning that our findings are incremental to the effects documented in Lambert et al. (2016). The similarity in inferences across the two endogeneity-correction methods is reassuring.

5.3 Cross-sectional tests

An alternative strategy to address causality focuses on the underlying mechanisms through which the theory operates and document that they are working (Rajan and Zingales, 1998). We attempt to do that using cross-sectional tests in this section. We use two partitioning variables – one based on booms versus busts and the other based on whether the restatement is on account of an error or an irregularity.

5.3.1 Accounting manipulation across the business cycle

Anecdotal evidence (e.g., The Economist, 2002) as well as academic research (e.g., Kedia and Philippon, 2009; Povel et al., 2007; Wang et al., 2010) purports that firms have greater incentives to perpetuate accounting fraud during good times, which is subsequently revealed during downturns. Povel et al. (2007) present a theoretical model where firms' incentives to commit fraud are higher during booms as compared to busts. Investors in their model receive financing requests from firms with good investment opportunities as well as those with bad ones. To reduce this adverse selection problem, investors can either screen based on financial reports (which are noisy) or invest in monitoring these firms. The model assumes that the marginal benefit of monitoring firms with positive public information is lower in good times, as it merely confirms investors' priors. In contrast, good public information is likely to receive greater scrutiny from investors during bad times. Consequently, investors monitor less, and firms indulge in more manipulation during booms as compared to busts.

What is missing from the above studies is that accounting resources also likely show disparities between booms and other times, and that these disparities (rather than fraud incentives) could explain variation in misreported financial numbers. In particular, during periods of growth, firms take on new projects, make more sales, attract new customers, launch new products, undertake more acquisitions, all of which put greater stress on the accounting staff to not only reflect these activities in the financial statements accurately, but also file these reports on time. The marginal cost of devoting resources to increasing the accounting function is greater during boom periods when these resources have more productive alternatives (such as increasing output to meet high demand). Moreover, given the high fixed costs and high adjustment costs of accounting resources, most firms do not adjust accounting resources during booms and busts. As a result, accounting staffs are likely to be stretched thin during booms versus busts.

If our conjecture is valid, then prior studies' hypothesized link between the business cycle and accounting fraud could be driven by variation in accounting resources across the cycle. In other words, if accounting resources are stretched more during booms than other times, then it could be that these resources are the missing link that connects the business cycle with the greater proportion of misreported financial statements. It is, however, conceivable that accounting resources are in fact stretched thin during downturns and not booms as the former is when firms cut costs to survive, and that these cost containment programs constrain accounting resources more during tough times. In such a case we should observe less accounting resources during downturns than booms. Since these countervailing arguments can only be addressed empirically, we begin with an examination of how accounting resources vary across the business cycle.

Panel A of Table 7 presents graphical evidence. The x-axis splits the sample period into periods of high (i.e., above sample median) sales growth versus low sales growth. We refer to the latter as Boom periods and the former as Non-boom periods. The first histogram indicates that the mean sales-growth (*SGR*) during non-boom periods is 5% as compared to 13% during booms. The second histogram indicates that accounting resources are lower during booms, as compared to other periods. This evidence is consistent with accounting resources being stretched more during good times relative to other times. Given this evidence, we next examine whether the association between accounting resources and the likelihood of a future earnings restatement is stronger during booms (as we predict) as compared to other times. To do so, we modify the single-stage (2SLS) design by splitting the accounting resources variable *ACC_RES* (*ACC_RES_PRED*) into *ACC_RES_BOOM* and *ACC_RES_NON_BOOM* (*ACC_RES_PRED_BOOM* and *ACC_RES_PRED_NON_BOOM)* to denote accounting resources during booms and other periods respectively.[16]

Panel B of Table 7 presents the results, where we tabulate only the coefficients on the accounting resources variables although the regression includes all the incentives variables and also the fixed effects. Model (1) presents results for the *ACC_RES* split, while model (2) splits the *ACC_RES_PRED* variable. As predicted, the explanatory power of accounting resources for restatements is more pronounced for booms as compared to other periods. In particular, the coefficient on *ACC_RES_BOOM* is -0.166 while that on *ACC_RES_NON_BOOM* is -0.075 with these coefficients being significantly different from each other at the 5% level. Similarly, only the coefficient on *ACC_RES_PRED_BOOM* is negative (-0.334) and significant at the 5% level while that on *ACC_RES_PRED_NON_BOOM* is negative but insignificant.[17] In terms of economic significance, moving from the bottom decile to the top decile of accounting resources reduces the likelihood of a restatement by 4.6% (9.1% minus 4.5%) during booms while such a move during

[16] An alternative design is to estimate separate regressions for booms and busts. This is tantamount to estimating a single specification with interaction terms of the boom/non-boom indicator with all the controls and the fixed effects. Our results are robust to this alternative specification. In particular, the coefficient on *ACC_RES* is economically larger during booms (-0.165) as compared to busts (-0.075).

[17] However, these coefficients are not statistically different from each other at conventional levels.

non-boom periods reduces the restatement likelihood by only 2.3% (7.8% minus 5.5%). Model (3) presents the results for the 2SLS design restricted to the matched sample. The results are similar but less significant using this restricted sample.

5.3.2 Restatements due to errors versus irregularities

We examine how the explanatory power of accounting resources for restatements varies depending on the underlying reason for the restatement – i.e., errors versus irregularities. Our model predicts that misreporting incentives depend on the amount of accounting resources in the firm. Thus, we expect the explanatory power of the incentive variables to weaken (especially for the subsample of irregularities) once we control for accounting resources. In addition, we expect the direct effect of accounting resources to be more important in predicting errors as compared to irregularities since low accounting resources are likely to result in more accounting mistakes that need to be corrected through restatements.

Table 8 presents these results. We follow Hennes et al. (2008) and partition the sample based on the magnitude of the 15-day market reaction to the announcement of the restatement to differentiate between errors and irregularities – where negative market reactions are assumed to indicate restatements due to irregularities (i.e., malfeasance). We split our sample of restatements into two groups – Low versus Medium/high based on terciles of the market reaction. The groups correspond to a median 15-day market reaction of -6.4% and 0.8% respectively. We classify the "Low" tercile as indicating irregularities and the Medium/high groups to capture errors.[18] We estimate our restatement prediction model within each group.

Models (1) and (2) present results for the irregularities subsample while models (3) and (4) present those for the errors subsample. We begin with assessing the explanatory power of the

[18] Our results are robust to examining each of the three groups individually.

covariates without *ACC_RES*. Model (1) depicts a partial *F* stat of 14.26 for the incentive variables indicating joint significance at the 5% level (*p* value = 0.047). In contrast to the insignificant coefficient on *SIZE* for the entire sample in Table 4, *SIZE* is now negative and significant (not tabulated), consistent with prior studies' interpretation that larger firms are less likely to manipulate financial statements (presumably on account of greater scrutiny). Model (1) also presents results that include *ACC_RES* as an additional variable. Two inferences emerge – first, the coefficient on *ACC_RES* remains negative and significant, indicating that accounting resources are inversely correlated with the likelihood of a restatement due to an irregularity; and second, the partial *F* stat of the incentive variables now becomes insignificant (*p* value = 0.125), indicating that these variables no longer provide any additional explanatory power for predicting restatements. This is exactly what one would expect if accounting resources jointly determine misreporting incentives and restatement likelihood. Further, the coefficient on the predicted value of accounting resources (*ACC_RES_PRED*) from the 2SLS specification is insignificant in model (2) indicating that accounting resources do not have an independent effect on predicting irregularities (beyond their indirect effect through misreporting incentives).

We repeat these analyses for errors in models (3) and (4). The partial *F* stats do not differ by much (37.26 versus 36.93 respectively) indicating that accounting resources do not diminish the explanatory power of the incentive variables in the errors subsample. Finally, consistent with our theoretical model, the coefficient on *ACC_RES_PRED* in model (4) is negative and significant indicating an independent effect of accounting resources in predicting restatements due to errors. In terms of economic significance, moving from the bottom decile of accounting resources to the top decile reduces the likelihood of an error-based restatement from 0.057 (1 in 18 10Ks) to 0.033 (1 in 31 10Ks).

5.4 Managerial equity-based incentives

In our final test, we include CEO (and CFO) equity-based incentives as additional determinants. This follows a long stream of research that seeks to document the association between equity-based incentives and financial misreporting. While some studies document a positive association (e.g., Burns and Kedia, 2006; Bergstresser and Philippon, 2006; Denis, Hanouna and Sarin, 2006; Efendi, Srivastava and Swanson, 2007; Harris and Bromiley, 2007), others fail to find such an association (e.g., Jayaraman and Milbourn, 2015; Armstrong, Jagolinzer and Larcker, 2010; Baber et al., 2007; Erickson et al., 2006).

We perform this analysis as an additional test rather than in the main tables because the requirement of equity-based compensation data from Execucomp restricts the sample to larger firms and consequently shrinks the sample size to 10,115 observations. The equity-based incentives variables we include are CEO delta (*CEODELTA*) defined as (the log of) the increase in CEO wealth for a 1% increase in the stock price, and CEO vega (*CEOVEGA*) defined analogously as (the log of) the increase in CEO wealth for a 1% increase in stock price volatility. We also include delta and vega computed for the CFO (*CFODELTA* and *CFOVEGA*).[19]

Table 9 presents these results. We begin in model (1) by estimating the restatement likelihood model with the existing set of covariates (excluding *ACC_RES*) and including the compensation variables. All of the four compensation variables are insignificant.[20] Model (2) includes *ACC_RES* as an additional variable to help gauge how the existing variables change when the former is included. Several variables (e.g., *ROA*, *LEV*, *SGR*, *RETVOL*) decrease their economic magnitudes once *ACC_RES* is added as an additional explanatory variable. Further, we continue to find a negative and significant coefficient on *ACC_RES* (-0.138, with a p value <0.01). Further,

[19] We obtain these data from Coles et al. (2006) who use the methodology in Core and Guay (2002).
[20] Additionally, we are unable to reject the null that all the compensation variables are jointly insignificant.

all four compensation variables remain insignificant. Models (3) and (4) present the 2SLS results. Model (3) presents the first-stage results where, in addition to the usual determinants, we include the four compensation variables. Except for *CEODELTA* (which is negative and significant) and *CEOVEGA* (which is positive at a 10% level), the other compensation variables are insignificantly associated with *ACC_RES*. The negative association between *ACC_RES* and *CEODELTA* is suggestive of the cash-constraint channel that we discussed previously. The instrument (*WINDOW*) continues to be positively and significantly associated with *ACC_RES* and also comfortably clears the weak-instrument threshold. Model (4) shows that the association between the likelihood of a future restatement and the (predicted value of) accounting resources remains robust – the coefficient on *ACC_RES_PRED* remains negative and significant. Overall, our prior results are robust to the inclusion of controls for managerial compensation.

6. Conclusion

Prior studies acknowledge that not all restatements signify instances where managers fraudulently misreport financial statements, yet the theoretical framework and empirical measures for understanding the drivers of unintentional mistakes are not well identified in the literature. We contribute to this literature by positing that firms optimally choose how much to invest in accounting resources (including hardware, software and accounting staff human capital) and the resulting earnings quality of the firm is the outcome of this optimal investment in accounting resources. We emphasize that since this optimal choice is determined by many of the same firm characteristics that lead to misreporting incentives, an identification problem exists. Our conceptual framework and empirical models offer guidance on resolving this identification problem and the inherent endogeneity issues.

We present a simple conceptual framework where the optimal investment in accounting resources depends on characteristics such as size and profitability, and this investment in turn drives financial misreporting incentives and future restatements. Using two novel measures for accounting resources, filing timeliness and 10K spelling errors, we find evidence consistent with our model's predictions. First, accounting resources are increasing in size and profitability. Second, accounting resources are negatively correlated with the likelihood of a future restatement. Third, including accounting resources in the restatement prediction model diminishes the explanatory power of variables purported to capture misreporting incentives. Our results are robust to controlling for the endogeneity of accounting resources using alternative designs.

In additional tests, we show that accounting resources diminish the explanatory power of the misreporting variables more so in the case of irregularities as compared to errors. We also find that the explanatory power of accounting resources for restatements is stronger during booms as opposed to downturns, which in turn questions prior inferences on fraud across the business cycle. Our results raise an alternative interpretation, i.e., accounting resources also vary systematically across the business cycle and these resources (rather than fraud) could drive the higher incidence of restatements during booms as compared to busts.

APPENDIX A: A Model of the Optimal Investment in Earnings Quality

Let $A \geq 1$ represent the firm's investment decision in accounting resources where A=1 is the minimum investment to comply with GAAP and SEC filing requirements. S is the size of the firm, and $S > 1$. Π is an index of long-run expected profitability where $\Pi > 0$. $\Pi \approx 0$ implies the firm is about to shut down, and in the long-run profits must be greater than zero because the owners have an abandonment option (Hayn, 1995). We assume that the benefits of higher investments in accounting resources, conditional on firm size, S, and profitability of Π is:

$$\text{Benefits}(A|S,\Pi) = bS \ln(A) \ln(\Pi) \qquad \text{(A1)}$$

where $b > 0$. The benefits of higher investments in accounting resources increases in firm size, expected profits, and A. The marginal returns to higher levels of A are decreasing in A. Likewise, higher levels of profitability yield higher benefits but the marginal benefits of higher profits are decreasing as Π increases. The benefits of larger A increases in Π because the likelihood of firm survival is larger and hence the discounted present value of future benefits from accounting investments today are larger. Alternatively, financially constrained firms (i.e., very low profits) find the opportunity cost of investing in accounting resources to be high.

The costs of higher investments in accounting, conditional on firm size, S, and expected profitability of Π is

$$\text{Costs}(A|S,\Pi) = \text{Fixed Costs}(A|S,\Pi) + \text{Variable Costs}(A|S,\Pi) \qquad \text{(A2)}$$

$$\text{Fixed Costs}(A|S,\Pi) = A \ln(S) \qquad \text{(A3)}$$

$$\text{Variable Costs}(A|S,\Pi) = vAS \qquad \text{(A4)}$$

where $v > 0$. The fixed costs of investing in accounting resources is increasing with A and in firm size, but at a decreasing rate. The variable costs are linear in firm size. The net benefits, $NB(A|S, \Pi)$, can be written as:

$$NB(A|S, \Pi) = bS \ln(A) \ln(\Pi) – A \ln(S) – vAS \tag{A5}$$

Maximizing NB(A|S,Π) with respect to A yields the following first order condition:

$$\delta NB(A|S, \Pi) / \delta A = [bS \ln (\Pi)]/A – \ln(S) – vS = 0 \tag{A6}$$

Or,

$$A^* = [bS \ln(\Pi)] / [\ln(S) + vS] \quad : \tag{A7}$$

Here we see that the firm-value maximizing level of investment in accounting resources, A*, depends on firm size and profitability.

Taking the partial derivative of A* with respect to S yields the following equation:

$$\delta A^* / \delta S = \{[\ln(S) + vS] b\ln(\Pi) – bS \ln(\Pi) [1/S + v]\} / [\ln(S) + vS]^2 \tag{A8}$$

$$= b \ln(\Pi) [\ln(S) – 1] / [\ln(S) + vS]^2 \tag{A9}$$

$\delta A^* / \delta S > 0$ since $\ln(S) > 1$.

Taking the partial derivative of A* with respect to Π yields the following equation:

$$\delta A^* / \delta \Pi = bS / \{\Pi \; [\ln(S) + vS]\} \tag{A10}$$

$\delta A^* / \delta \Pi > 0$ because both the numerator and denominator are greater than zero.

This simple model predicts that larger, more profitable firms will optimally invest more in accounting resources than its smaller, less profitable peers.

Appendix B: Variable Definitions

ACC_RES	The composite measure of accounting resources estimated by combining *FILE_TIME* and *SPELL_ERR* using principal components.
ACC_RES_BOOM	The composite measure of accounting resources estimated by combining *FILE_TIME* and *SPELL_ERR* using principal components during boom periods. Boom periods are defined as periods with sales growth above the sample median.
ACC_RES_NON_BOOM	The composite measure of accounting resources estimated by combining *FILE_TIME* and *SPELL_ERR* using principal components during non-boom periods. Non-boom periods are defined as periods with sales growth below the sample median.
ACC_RES_PRED	The predicted value of *ACC_RES* estimated from the first-stage regression using the firm's filing window as an instrumental variable.
CEO_DELTA	The log of the increase in CEO wealth for a 1% increase in the company's stock price.
CEO_VEGA	The log of the increase in CEO wealth for a 1% increase in the company's stock price volatility.
CFO_DELTA	The log of the increase in CFO wealth for a 1% increase in the company's stock price.
CFO_VEGA	The log of the increase in CFO wealth for a 1% increase in the company's stock price volatility.
CMPLX	The natural log of the number of business segments reported in Compustat Segments.
FILE_TIME	The difference between the firm's filing deadline date and the date the firm filed its 10K.
LEV	Book leverage defined as total short-term and long-term debt scaled by total assets.
MTB	The market-to-book ratio defined as the ratio of market value of equity plus the book value of debt scaled by the book value of assets.
NCAP	An indicator variable equal to one when the firm issued debt or equity greater than 5% of the firm's total assets, and 0 otherwise.
RESTATE	An indicator variable equal to 1 if the firm restates their financial statements in the subsequent year, and 0 otherwise.
RET	Excess stock returns defined as annual firm returns minus size-adjusted returns.
RETVOL	Stock return volatility defined as the annual standard deviation of daily returns.
ROA	Return on assets defined as income before extraordinary items scaled by lagged total assets.
SGR	The annual growth in sales over the previous year.
SIZE	The natural log of total sales.
SIZE_SQ	The squared term of firm size.
SPELL_ERR	Spelling errors defined as the number of spelling errors in the 10K (times 10^3).
TIMELY	An indicator variable that takes the value of 1 if the firm files its 10K on or before the filing deadline, and 0 otherwise.
WINDOW	The 10K filing window defined as the difference between the filing deadline date and the date of the fiscal year end.

Appendix C: Downloading 10Ks from EDGAR

The following details the procedure to download 10Ks from the SEC EDGAR website using the FTP server.[21]

1. **General Process**
   - **1.1.** Download Daily Index Files from the SEC FTP Server using Wget.exe.
   - **1.2.** Unzip Daily Index files using WinRAR.
   - **1.3.** Run Python code called "10K_Collect.py." This code runs through each year of daily index files and writes a text file that lists the location of the 10K forms to be downloaded from EDGAR.
   - **1.4.** Download 10Ks from EDGAR using Wget.exe and the text file created in 1.3 above.
   - **1.5.** Run Python code called "10K_Parse.py." This parses each individual 10K file, extracts the company information, cleans the HTML/XBRL formatting, and provides a spelling error score for each 10K using Microsoft Word (see Appendix D).

2. **Detailed Procedures**

   2.1. Download Daily Index Files from the SEC FTP Server using Wget.exe.
   - 2.1.1. Download Wget.exe from GNU.
   - 2.1.2. Save Wget.exe in the folder where you want to store all of the daily index files.
   - 2.1.3. Use the command prompt to run Wget.exe to get the SEC daily files. Type CMD into Program Search.
   - 2.1.4. Type: C:\Users\Jacquelyn\Wget> wget.exe -N -o download_edgar_index.log -r --accept="form.*" ftp://ftp.sec.gov/edgar/daily-index/. This pulls the daily index files from EDGAR and places them on your computer in exactly the same folder and subfolder structure that appears on the FTP server. It does this by recursively (-r command) storing them. A log of what is downloaded will appear in a text (log) file called download_edgar_index.log as specified above. For further information on how to use Wget.exe, see https://www.gnu.org/software/wget/manual/wget.html.
   - 2.1.5. Rerun the command in 2.1.4 until no new files are downloaded. Not every daily index file will download the first time. Run Wget.exe a few times until no new downloads appear in the download_edgar_index.log. Wget.exe will not download files that are already stored on your computer.

   2.2. Unzip Daily Index Files using WinRAR.
   - 2.2.1. The daily index files should now be stored on your computer. Ideally, each index file is an .idx file. If some files are formatted as a .gz or .tar file, you can easily extract these files into the .idx format using WinRAR.

   2.3. Run Python code called "10K_Collect.py."

[21] This code will be available at Jacquelyn Gillette's MIT website, and this procedure has been built following a similar process for NSAR forms as described in Clifford, Fulkerson, Jordan, and Waldman (2011) and by Robert Parham at http://www.kn.owled.ge/.

2.3.1. Run this code to output a text file listing each 10K to be downloaded and its location on the SEC FTP server. Copy this text file and store it in the same folder as Wget.exe (called 10k.txt below).

2.4. Download 10Ks from EDGAR using Wget.exe and the text file created in 2.3 above.

2.4.1. Run Wget.exe to download each 10K text file. Using the command prompt, type: C:\Users\Jacquelyn\Wget> wget -N --retr-symlinks -o download_10k.log -i 10k.txt –P E:\Gillette_10K_files ftp://ftp.sec.gov/edgar/daily-index/.

2.5. Run Python code called "10K_Parse.py." This parses each individual 10K file, extracts the company information, cleans the HTML/XBRL formatting, and provides a spelling error score for each 10K using Microsoft Word (see Appendix D).

Appendix D: Calculation of Spelling Errors

The following details the steps to calculate the number of spelling errors in a given 10K.[22]

1. **Alter settings in Microsoft Word.**
   1.1. In Word, select Options, select Proofing: Uncheck the box that says "Flag repeated words".

2. **Alter the custom dictionary in Microsoft Word.**
   2.1. Download the master word file of over 84,000 words from Bill McDonald's website : http://www3.nd.edu/~mcdonald/Word_Lists.html.
   2.2. Save the master list.
   2.3. Create a new column in the master list that turns the list of words into lower case (use "=lower()" in Excel).
   2.4. Save the master list as a text file.
   2.5. Cut and paste the master list of lowercase words in the text file into the CUSTOM.dic in Microsoft Word.
      2.5.1. In Windows XP, for example: C:\Documents and Settings\Zimmerman.UR\Application Data\Microsoft\UProof
      2.5.2. In Windows 7 or 8, for example: C:\Users\%username%\AppData\Roaming\Microsoft\

3. **Change Proofing Settings in Microsoft Word.**
   3.1. *Uncheck* Check spelling as you type
   3.2. *Uncheck* Use contextual spelling
   3.3. *Uncheck* Mark grammar errors as you type
   3.4. *Uncheck* Check grammar with spelling

4. **Download Pywin32, the VBA script that Python uses to run Word behind the scenes.**
   4.1. Download Pywin32 from sourceforge: http://sourceforge.net/projects/pywin32/files/pywin32/Build%20219/
   4.2. I recommend downloading this file: pywin32-219.win32-py2.7.exe. This release is compatible with Python 2.7 and 32-bit Python. It is important to download the release of Pywin32 that is compatible with the operating system on your computer and the version of Python that you have downloaded or it will not run the .exe file.
   4.3. Run the .exe file.
   4.4. Go to your Python scripts files. Pywin32 will download approximately 4 folders of data into your Python library, but some files will need to be rearranged.
      4.4.1. Go to C:\Python27\Lib\site-packages.
      4.4.2. From the Pywin32_Systems32, copy all 3 modules into Win32com -> Client folder.
      4.4.3. From the win32 folder, copy all files into Win32com -> Client folder.
   4.5. Run the python script using win32.client module.

[22] This code will be available at Jacquelyn Gillette's MIT website.

4.6. Note that if you receive an error message that win32api is not found in the client folder, then it is necessary to completely close IDLE and restart Python. Try performing the steps outlined in 4.4 above, and then try again after restarting Python.

5. **Run Python code called "10K_Parse.py." This parses each individual 10K file, extracts the company information, cleans the HTML/XBRL formatting, and provides a spelling error score for each 10K using Microsoft Word.**

6. **Modify the Python code and re-perform the analysis.**
   6.1. Alter the code such that the loop that deals with "\n" in the middle of words is turned on (if it was off the first time).
   6.2. Change the name of the output file to "_v2", for example.

7. **For each 10K, spelling errors is calculated as the minimum number of spelling errors after running the Python code twice as described in 5 and 6 above.**

8. **Notes:**
   8.1. Only one python script can run on a given computer at one time. The reason is that it uses Microsoft Word, and Word cannot enable more than one screen to be open at a time. To improve processing speed, run the code on several computers at one time, only one script per computer.
   8.2. Reset the computer's "Power Options" to never sleep and never go to screen saver so that the code can run uninterrupted until it is finished.
   8.3. Approximate run time for the "10K_Parse.py" python code is 4-8 weeks. This estimate assumes that the python code is run simultaneously on 3 computers on all 10Ks available in EDGAR. The steps above require that "10K_Parse.py" is modified and run twice on each 10K such that the total run time for calculating the number of spelling errors in 10Ks is 8-16 weeks.

References

Aier, J., J. Comprix, M. Gunlock, and D. Lee. 2005. The financial expertise of CFOs and accounting restatements. Accounting Horizons 19 (3): 123–135.

Armstrong, C., Jagolinzer, A., Larcker, D., 2010. Chief executive officer equity incentives and accounting irregularities. Journal of Accounting Research 48, 225–271.

Arif, S., Marshall, N., Schroeder, J., Yohn, T., 2016. The influence of mandatory disclosure regulation on voluntary disclosure: Financial reporting timeliness and the rise of concurrent earnings announcements. Working paper, Indiana University.

Armstrong, C., Jagolinzer, A., Larcker, D., 2010. Chief executive officer equity incentives and accounting irregularities. Journal of Accounting Research 48, 225–271.

Ashbaugh-Skaife, H., Collins, D., Kinney, W., 2007. The discovery and reporting of internal control deficiencies prior to SOX-mandated audits. Journal of Accounting and Economics 44, 166–192.

Baber, W., Kang, S., Liang, L., Zhu, Z., 2009. Shareholder rights, corporate governance and accounting restatement. Working paper, Georgetown University.

Bergstresser, D., Philippon, T., 2006. CEO incentives and earnings management. Journal of Financial Economics 80, 511–529.

Boland, C., Bronson, S., and Hogan, C., 2015. Accelerated Filing Deadlines, Internal Controls, and Financial Statement Quality: The Case of Originating Misstatements. Accounting Horizons 29 (3), 551–575.

Burns, N., Kedia, S., 2006. The impact of performance-based compensation on misreporting. Journal of Financial Economics 79, 35–67.

Carcello, J.V., Neal, T.L., Palmrose,Z.V., Scholz,S., 2011. CEO involvement in selecting board members, audit committee effectiveness, and restatements. Contemporary Accounting Research 28 (2), 396–430.

Chava, S., Purnanandam, A., 2010. CEOs versus CFOs: Incentives and corporate policies. Journal of Financial Economics 97, 263-278.

Chava, S., Roberts, M., 2008. How does financing impact investment? The role of debt covenants. The Journal of Finance 63, 2085-2121.

Chin, C.L., Chi, H.Y., 2009. Reducing restatements with increased industry expertise. Contemporary Accounting Research 26 (3), 729–765.

Choudhary, P., Merkley, K., Schipper, K., 2016. Qualitative characteristics of financial reporting errors deemed immaterial by managers. Working paper, Georgetown University.

Clifford, C., Fulkerson, J., Jordan, B., Waldman, S., 2011. Do investors care about risk? Evidence from mutual fund flows. Working paper, University of Kentucky.

Coles, J. L., Daniel, N. D., Naveen, L., 2006. Managerial incentives and risk-taking. Journal of financial Economics 79, 431-468.

Core, J., Guay, W., 2002. Estimating the value of employee stock option portfolios and their sensitivities to price and volatility. Journal of Accounting Research 40, 613-630.

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.,Ge, W., Larson, C., and Sloan, R., 2011. "Predicting Material Accounting Misstatements," Contemporary Accounting Research Vol. 28 No. 1 (Spring 2011) pp. 17–82.

Denis, D. J., Hanouna, P., Sarin, A., 2006. Is there a dark side to incentive compensation? Journal of Corporate Finance 12, 467-488.

Desai, H., Hogan, C., Wilkins, M., 2006. The reputational penalty for aggressive accounting: Earnings restatements and management turnover. The Accounting Review 81, 83-112.

Doyle, J., Ge, W., McVay, S., 2007b. Determinants of weaknesses in internal control over financial reporting. Journal of Accounting and Economics 44, 193–223.

Efendi, J., Srivastava, A., Swanson, E., 2007. Why do corporate managers misstate financial statements? The role of option compensation and other factors. Journal of Financial Economics 85, 667–708.

Erickson, M., Hanlon, M., Maydew, E., 2006. Is there a link between executive equity incentives and accounting fraud? Journal of Accounting Research 44, 113–143.

Fama, E. and French, K., 2004. New lists: fundamentals and survival rates. Journal of Financial Economics, 73 (2), 229-269.

Francis, J., LaFond, R., Olsson, P., Schipper, K., 2005. The market pricing of accruals quality. Journal of Accounting and Economics 39, 295–327.

Gao, F., 2015. To comply or not to comply: Understanding the discretion in reporting public float and SEC regulations. Working paper, University of Illinois at Chicago.

Harris, J., Bromiley, P., 2007. Incentives to cheat: The influence of executive compensation and firm performance on financial misrepresentation. Organization Science 18, 350-367.

Hayn, C., 1995. The Information Content of Losses. Journal of Accounting and Economics 20 (2), 125–153.

Hennes, K.M., Leone, A.J., Miller, B.P., 2008. The importance of distinguishing errors from irregularities in restatement research: The case of restatements and CEO/CFO turnover. The Accounting Review 83 (6), 1487–1519.

Ittner, C. D., Lambert, R. A., Larcker, D. F., 2003. The structure and performance consequences of equity grants to employees of new economy firms. Journal of Accounting and Economics 34, 89-127.

Irvine, P. and Pontiff, J., 2009. Idiosyncratic return volatility, cash flows and product market competition. Review of Financial Studies, 22 (3), 1149–1177.

Jayaraman, S., Milbourn, T. T., 2015. CEO equity incentives and financial misreporting: The role of auditor expertise. The Accounting Review 90, 321-350.

Lambert, T., Jones, K., Brazel, J., Showalter, D. S., 2016. Audit time pressure and earnings quality: An examination of accelerated filings. Working paper, Lehigh University.

Larcker, D.F., Zakolyukina, A., 2012. Detecting deceptive discussions in conference calls. Journal of Accounting Research 50, 495-540.

Kedia, S., Philippon, T., 2009. The economics of fraudulent accounting. Review of Financial Studies 22, 2169-2199.

Kinney, W., McDaniel, L., 1989. Characteristics of firms correcting previously reported quarterly earnings. Journal of Accounting and Economics 11, 71–93.

Kinney Jr., W.R., Palmrose, Z.V., Scholz, S., 2004. Auditor independence, non-audit services, and restatements: was the U.S., government right? Journal of Accounting Research 42 (3), 561–588.

Lee, D., 2008. Randomized experiments from non-random selection in U.S. house elections. Journal of Econometrics 142, 675-697.

Lee, D., Lemieuxa, T., 2010. Regression discontinuity designs in economics. Journal of Economic Literature 48, 281-355.

Lobo, J., and Zhao, Y., 2013. Relation between audit effort and financial report misstatements: evidence from quarterly and annual restatements. The Accounting Review, 88 (4), 2013, 1385-1412.

McNichols, M., Stubben, S., 2008. Does earnings management affect firms' investment decisions? The Accounting Review 83, 1571-1603.

Owens, E., Wu, J. and Zimmerman, J., 2015. Business model shocks and abnormal accrual models. Working Paper, University of Rochester.

Richardson, S., Tuna, I., Wu, M., 2002. Predicting earnings management: The case of earnings restatements. Working paper, University of Pennsylvania.

Roberts, M., Whited, T., 2013. Endogeneity in empirical corporate finance. Handbook of the Economics of Finance 2, 493-572.

Plumlee, M., and T. L. Yohn. 2010. An analysis of the underlying causes attributed to restatements. The Accounting Horizons 24 (1): 41–64.

Povel, P., Singh, R., Winton, A., 2007. Booms, busts, and fraud. Review of Financial Studies 20, 1219-1254.

Rajan, R. G., Zingales, L., 1998. Which capitalism? Lessons from the east Asian crisis. Journal of Applied Corporate Finance, 11(3), 40-48.

Skinner, D. J., Sloan, R. G., 2002. Earnings surprises, growth expectations, and stock returns or don't let an earnings torpedo sink your portfolio. Review of accounting studies 7, 289-312.

Srinivasan, S., 2005. Consequences of financial reporting failure for outside directors: Evidence from accounting restatements and audit committee members. Journal of Accounting Research 43 (2), 291–334.

Stock, J., Wright, J., Yogo, M., 2002. A survey of weak instruments and weak identification in generalized method of moments. Journal of Business and Economic Statistics 20, 518-529.

The Economist, 2002. Thumped. July 13, 2002, 11-12.

Wang, T., Winton, A., Yu, X., 2010. Corporate fraud and business conditions: Evidence from IPOs. The Journal of Finance 65, 2255-2292.

Table 1: Sample Selection

This table reports the sample selection process beginning from the set of 10Ks available on the SEC EDGAR website. The table provides the number of firm-year observations remaining in the sample after each data requirement.

Sample Selection Procedure	Firm-Years
10Ks Downloaded from EDGAR (August 31, 1993 - May 31, 2014)	135,120
10Ks Successfully Analyzed Spelling Errors using Python	134,300
10Ks with Filing Status in Audit Analytics	117,356
10Ks with Compustat and Compustat Segments Data	94,701
10Ks Filed in Years 1998-2014 to Match Audit Analytics Restatement Coverage	81,631
Retain First 10K Filed in Fiscal Year	62,793
10Ks with CRSP Data, SIC Data, and 10Ks that are not: ABS, REIT, Shell, Blank Check, Non-Operational, Fund, or Trust Companies	41,284
10Ks with Filing Timeliness between -15 days and 24 days	38,185
Final Sample	**38,185**

Figure 1: Mean restatement likelihood by dichotomous classification: Late versus Timely

The x-axis denotes firms that filed their 10Ks on or before the filing deadline (Timely) versus those that filed after the deadline (Late). The y-axis plots the average proportion of restatements for each group.

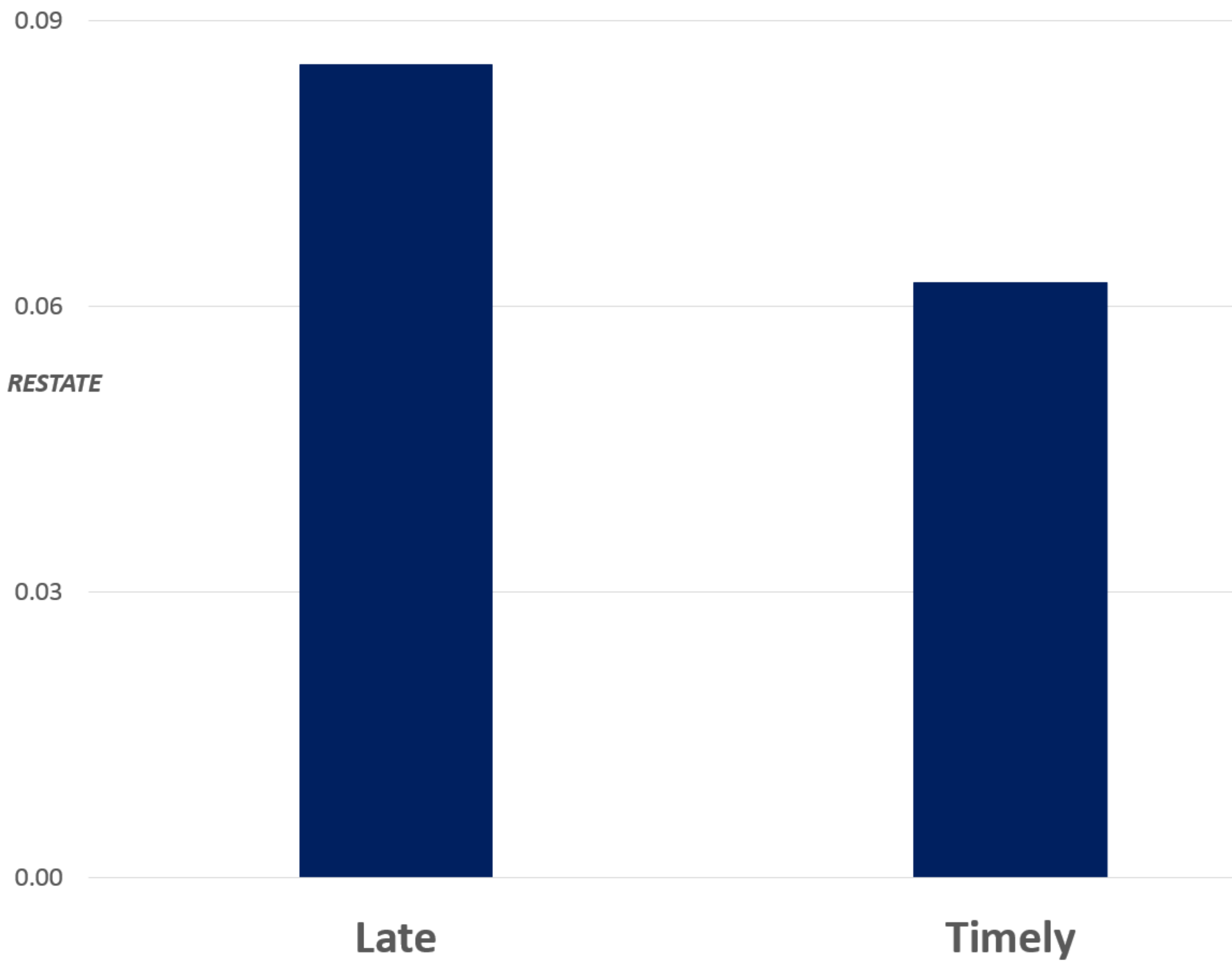


Table 2: Descriptive statistics

The sample consists of 38,190 firm-year observations for 6,642 unique firms for the period 1998 to 2014. *RESTATE* is an indicator variable equal to 1 if the firm restates their financial statements in the subsequent year, and 0 otherwise. Filing timeliness (*FILE_TIME*) is measured as the difference between the firm's filing deadline date and the date it filed its 10K. Spelling Errors (*SPELL_ERR*) is defined as the number of spelling errors in the 10K (times 10^3). *ACC_RES* is the composite measure of accounting resources estimated by combining *FILE_TIME* and *SPELL_ERR* using principal components. *WINDOW* denotes the 10K filing window defined as the difference between the filing deadline date and the date of the fiscal year end. *ROA* is defined as Income before extraordinary items scaled by lagged total assets. *RET* denotes excess stock returns defined as annual firm returns minus size-adjusted returns. *NCAP* is an indicator equal to one when the firm issued debt or equity greater than 5% of the firm's total assets, and 0 otherwise. *LEV* denotes book leverage and is defined as total debt scaled by total assets. *MTB* is defined as the market-to-book ratio, defined as the ratio of market value of equity plus book value of debt scaled by book value of assets. *SGR* is the annual growth in sales over the previous year. *CMPLX* denotes firm complexity and is measured as the (log of) number of business segments. *RETVOL* denotes stock return volatility, defined as the annual standard deviation of daily returns. *SIZE* denotes firm size and is computed as (the log of) total sales. *CEODELTA* (*CEOVEGA*) denotes the log of the increase in CEO wealth for a 1% increase in the company's stock price (stock price volatility). Analogous definitions extend to the CFO with *CFODELTA* and *CFOVEGA*. *FILE_TIME* is truncated between -15 and 24 days (the 5% tails). All control variables are winsorized at the 1% tails.

Variable	**Obs.**	**Mean**	**Median**	**S.D.**	**Min**	**Max**
RESTATE	38,185	0.065	0.000	0.247	0.000	1.000
FILE_TIME	38,185	4.391	2.000	7.183	-15.000	24.000
SPELL_ERR	38,185	14.767	8.000	21.277	0.000	136.000
ACC_RES	38,185	0.000	-0.008	1.002	-5.938	2.421
WINDOW	38,185	77.607	75.000	11.673	60.000	90.000
ROA	38,185	-0.020	0.034	0.221	-1.106	0.396
RET	38,185	0.012	-0.068	0.605	-1.051	2.947
NCAP	38,185	0.338	0.000	0.473	0.000	1.000
LEV	38,185	0.210	0.160	0.216	0.000	0.961
MTB	38,185	2.034	1.492	1.630	0.555	10.464
SGR	38,185	0.169	0.078	0.525	-0.697	3.553
CMPLX	38,185	1.545	1.099	0.741	0.000	3.045
RETVOL	38,185	0.148	0.122	0.096	0.034	0.565
SIZE	38,185	5.731	5.784	2.143	-0.322	10.604
AUD_FEES	32,621	0.707	0.542	0.604	0.053	2.907
IND_EXP	34,565	0.515	1.000	0.500	0.000	1.000
CEO_DELTA	11,663	5.318	5.317	1.527	0.000	13.153
CEO_VEGA	11,830	3.657	3.912	1.822	0.000	9.329
CFO_DELTA	11,530	3.574	3.609	1.336	0.000	11.235
CFO_VEGA	11,786	2.536	2.639	1.460	0.000	7.574

Table 3: Correlation matrix

This panel presents spearman correlations. * ($^{+}$) denotes statistical significance at the 1% (5%) significance level.

	RESTATE	*FILE_TIME*	*SPELL_ERR*	*ACC_RES*	*WINDOW*	*ROA*	*RET*	*NCAP*	*LEV*	*MTB*	*SGR*	*CMPLX*	*RETVOL*	*SIZE*	*CEO_DELTA*	*CEO_VEGA*	*CFO_DELTA*	*CFO_VEGA*
RESTATE	1.0																	
FILE_TIME	-0.045*	1.0																
SPELL_ERR	0.006	-0.018	1.0															
ACC_RES	-0.036*	0.767*	-0.553*	1.0														
WINDOW	0.023^{+}	0.078*	-0.214*	0.193*	1.0													
ROA	-0.048*	0.049*	-0.051*	0.054*	-0.132*	1.0												
RET	-0.010	0.069*	-0.007	0.054*	-0.011	0.246*	1.0											
NCAP	0.012	-0.016	0.054*	-0.038*	-0.058*	-0.079*	-0.016	1.0										
LEV	0.033*	0.006	0.080*	-0.037*	-0.052*	-0.248*	-0.048*	0.469*	1.0									
MTB	-0.050*	0.015	-0.011	0.001	-0.061*	0.583*	0.317*	-0.152*	-0.280*	1.0								
SGR	0.009	-0.066*	0.017	-0.064*	0.004	0.355*	0.202*	0.044*	-0.061*	0.287*	1.0							
CMPLX	0.018	0.040*	0.040*	0.020^{+}	-0.063*	-0.042*	0.027*	0.082*	0.153*	-0.172*	-0.068*	1.0						
RETVOL	0.048*	-0.060*	-0.037*	-0.019^{+}	0.357*	-0.228*	-0.049*	-0.046*	-0.081*	-0.128*	-0.001	-0.163*	1.0					
SIZE	-0.012	0.131*	0.080*	0.052*	-0.362*	0.092*	0.043*	0.145*	0.315*	-0.117*	-0.049*	0.330*	-0.386*	1.0				
CEO_DELTA	-0.019^{+}	0.037*	0.041*	-0.007	-0.156*	0.334*	0.223*	0.000	-0.000	0.382*	0.203*	0.058*	-0.257*	0.393*	1.0			
CEO_VEGA	-0.007	0.088*	0.066*	0.021^{+}	-0.136*	0.140*	0.062*	0.015	0.088*	0.174*	0.018	0.115*	-0.231*	0.446*	0.590*	1.0		
CFO_DELTA	-0.027*	0.071*	0.058*	0.011	-0.165*	0.303*	0.228*	0.024*	0.055*	0.368*	0.193*	0.063*	-0.273*	0.412*	0.647*	0.587*	1.0	
CFO_VEGA	-0.011	0.078*	0.045*	0.028*	-0.079*	0.167*	0.060*	0.000	0.053*	0.209*	0.059*	0.083*	-0.210*	0.383*	0.513*	0.812*	0.762*	1.0

Table 4: Restatement prediction model and the role of accounting resources

The dependent variable is *RESTATE* that denotes a restatement. *TIMELY* is an indicator that takes the value of 1 if the firm files its 10K on or before the filing deadline. *FILE_TIME* denotes filing timeliness. *SPELL_ERR* denotes the number of spelling errors (times 10^3). *ACC_RES* is the combined measure of accounting resources. *ROA* is income before extraordinary items. *RET* denotes annual (excess) return. *NCAP* is an indicator equal to one when the firm issued debt or equity, and 0 otherwise. *LEV* denotes book leverage. *MTB* is the market-to-book ratio. *SGR* is the growth in sales over the previous year. *CMPLX* denotes the (log of) number of segments. *RETVOL* denotes stock return volatility. *SIZE* denotes firm size defined as the log of total assets. Robust standard errors clustered by firm are presented in parentheses under the coefficients. In addition, Models (4) to (7) include year and industry fixed effects (defined at the 2-digit SIC code level). (***), (**), (*) denote statistical significance at the 1%, 5% and 10% levels, respectively. The incentive variables are *ROA, RET, NCAP, LEV, MTB, SGR,* and *SIZE.*

	Pred.	(1)	(2)	(3)	(4)	(5)	(6)
TIMELY	–	**-0.339**					
		[0.058]***					
FILE_TIME	**–**		**-0.018**				
			[0.003]***				
SPELL_ERR	**+**			**-0.001**			
				[0.001]			
ACC_RES	**–**				**-0.073**		**-0.111**
					[0.020]***		**[0.023]*****
ROA	–					-0.295	-0.257
						[0.130]**	[0.130]**
RET	–					0.037	0.046
						[0.036]	[0.036]
NCAP	+					0.096	0.096
						[0.049]*	[0.049]*
LEV	+					0.536	0.511
						[0.123]***	[0.122]***
MTB	+					-0.024	-0.026
						[0.017]	[0.017]
SGR	+					0.098	0.093
						[0.040]**	[0.040]**
CMPLX	?					0.260	0.259
						[0.034]***	[0.034]***
RETVOL	?					0.850	0.776
						[0.272]***	[0.273]***
SIZE	?					-0.001	0.002
						[0.015]	[0.015]
Partial F test of incentive variables						60.71	55.37
p value						<0.01	<0.01
Marginal effect at:							
TIMELY=0		0.085					
TIMELY=1		0.062					
5 days late			0.076				
5 days early			0.064				
Bottom decile				0.066	0.070		0.072
Top decile				0.064	0.060		0.057
Year and ind effects		No	No	No	No	Yes	Yes
Pseudo R^2		0.002	0.002	0.001	0.001	0.025	0.027
Obs.		38,185	38,185	38,185	38,185	38,185	38,185

Table 5: Ruling out audit verification effects

The dependent variable is *RESTATE* that denotes a restatement. *ACC_RES* is the combined measure of accounting resources. *AUD_FEES* denotes the log of audit fees (scaled to millions). *UNEXP_FEES* is unexplained audit fees defined as the residual of a regression of log audit fees on firm size (defined as the log of total assets). Robust standard errors clustered by firm are presented in parentheses under the coefficients. In addition, Models (1) to (3) include year and industry fixed effects (defined at the 2-digit SIC code level). (***), (**), (*) denote statistical significance at the 1%, 5% and 10% levels, respectively.

	Pred.	(1)	(2)
ACC_RES	–	-0.111	-0.113
		[0.025]***	[0.024]***
AUD_FEES	?	0.099	
		[0.076]	
UNEXP_FEES	?		0.090
			[0.079]
Marginal effect at:			
Bottom decile		0.074	0.074
Top decile		0.058	0.058
Controls		Yes	Yes
Year effects		Yes	Yes
Industry effects		Yes	Yes
Pseudo R^2		0.026	0.026
Obs.		32,621	32,621

Table 6: Addressing endogeneity of accounting resources

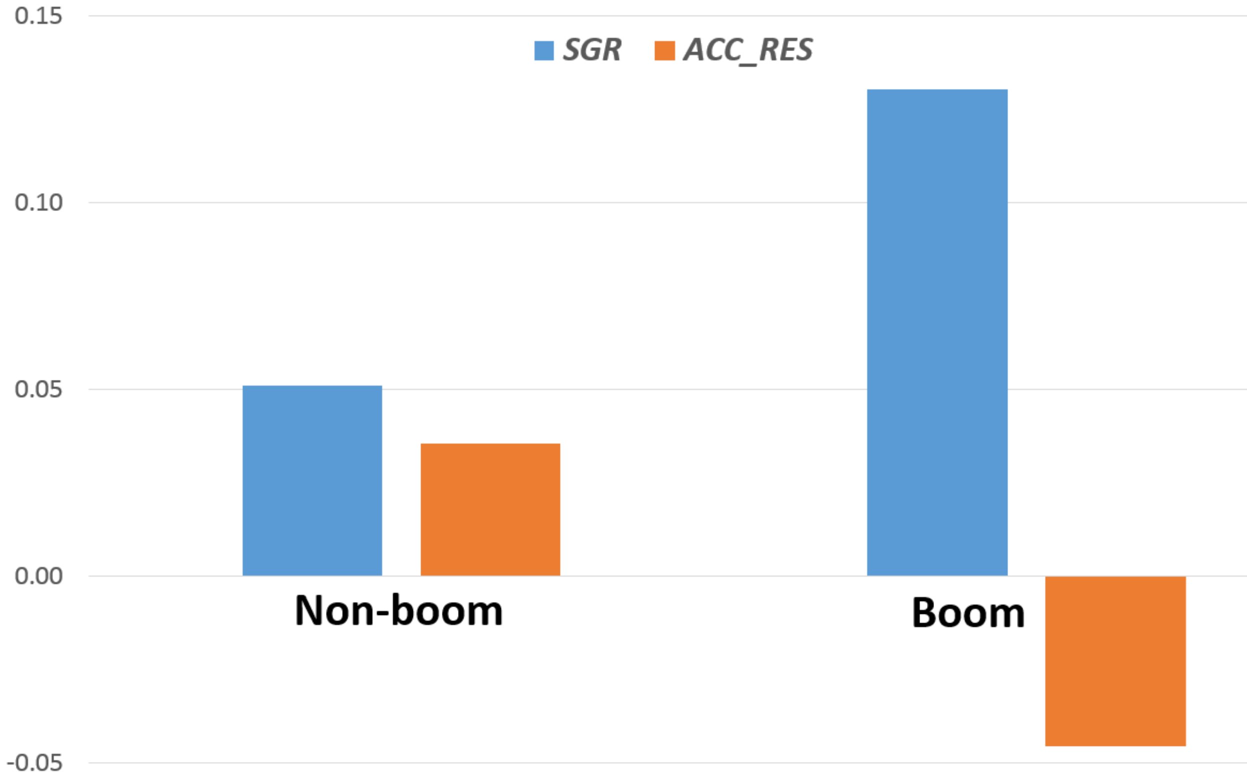
The dependent variable in Models (2), (4), and (6) is *RESTATE* that denotes whether the firm restated its earnings. Models (1), (3), and (5) use a 2SLS design where the first-stage dependent variable is *ACC_RES* and the instrument is the filing window (*WINDOW*). *ACC_RES_PRED* denotes the predicted value of *ACC_RES* estimated from the first stage. Models (3)-(6) use a matched-sample design based on changes in firms' filing deadlines. The sample is restricted to firms with changes in their filing deadlines and control firms with similar changes in MVE but without changes in their filing deadlines. Models (5) and (6) also include an additional control, *MVE_CHNG_Q2*, which denotes the change in MVE during the second fiscal quarter (i.e., the quarter that determines the firm's filing status and filing deadline). This controls for any residual differences between the affected and unaffected firms due to imperfect matching. All of the specifications include unexplained audit fees as a control. All regressions include year and industry fixed effects (defined at the 2-digit SIC code level). Robust standard errors clustered by firm are presented in parentheses under the coefficients. (***), (**), (*) denote statistical significance at the 1%, 5% and 10% levels, respectively.

Designs		Entire sample		Matched sample			
Dep. variable		*ACC_RES* (1st stage)	*RESTATE* (2nd stage)	*ACC_RES* (1st stage)	*RESTATE* (2nd stage)	*ACC_RES* (1st stage)	*RESTATE* (2nd stage)
	Pred.	(1)	(2)	(3)	(4)	(5)	(6)
WINDOW	+	**0.029**		**0.031**		**0.031**	
		[0.001]***		**[0.002]*****		**[0.002]*****	
ACC_RES_PRED	–		**-0.222**		**-0.293**		**-0.294**
			[0.128]*		**[0.158]***		**[0.158]***
MVE_CHNG_Q2						-0.003	0.014
						[0.011]	[0.048]
p. value of partial *F*-test		<0.001		<0.001		<0.001	
Marginal effect at:							
Bottom decile			0.083		0.083		0.083
Top decile			0.052		0.044		0.044
Audit-related controls		Yes	Yes	Yes	Yes	Yes	Yes
Firm-level controls		Yes	Yes	Yes	Yes	Yes	Yes
Year effects		Yes	Yes	Yes	Yes	Yes	Yes
Industry effects		Yes	Yes	Yes	Yes	Yes	Yes
Adj./Pseudo R^2		0.187	0.027	0.208	0.034	0.208	0.034
Obs.		32,621	32,621	20,686	20,884	20,884	20,884

Table 7: Filing timeliness and restatements across the business cycle

Panel A: Graphical evidence

The horizontal axis depicts Non-boom (Boom) periods are defined as years when sales growth (*SGR*) for all firms is below (above) the sample-wide median. The vertical axis plots sales growth and accounting resources that correspond to these periods.



Panel B: Multivariate evidence

The dependent variable in Model (1) is *RESTATE* that denotes whether the firm restated its earnings in year *t+1*. The composite measure of accounting resources (*ACC_RES*) is split into two based on boom versus bust periods (*ACC_RES_BOOM* and *ACC_RES_NON_BOOM*) respectively. Similar definitions extend to *ACC_RES_PRED_BOOM* and *ACC_RES_PRED_NON_BOOM* that decompose the predicted value of accounting resources (*ACC_RES_PRED*). All control variables are as defined in Table 4, and are included in the specification but not presented. All regressions include year and industry fixed effects (defined at the 2-digit SIC code level). Robust standard errors clustered by firm are presented in parentheses under the coefficients. (***), (**), (*) denote statistical significance at the 1%, 5% and 10% levels, respectively.

Dep. variable		*RESTATE*		
		Single-stage	2SLS (Entire sample)	2SLS (Matched sample)
	Pred.	(1)	(2)	(3)
ACC_RES_BOOM	–	**-0.166**		
		[0.034]***		
ACC_RES_NON_BOOM	–	**-0.075**		
		[0.028]***		
ACC_RES_PRED_BOOM	–		**-0.334**	**-0.307**
			[0.151]**	**[0.173]***
ACC_RES_PRED_NON_BOOM	–		**-0.164**	**-0.280**
			[0.132]	**[0.167]***
p value of (1) = (2)		0.028	0.117	0.831
Marginal effect at:				
Boom:				
Bottom decile of *ACC_RES*		0.076	0.091	0.084
Top decile of *ACC_RES*		0.053	0.045	0.043
Non-boom:				
Bottom decile of *ACC_RES*		0.070	0.078	0.083
Top decile of *ACC_RES*		0.060	0.055	0.046
Controls		Yes	Yes	Yes
Year effects		Yes	Yes	Yes
Industry effects		Yes	Yes	Yes
Pseudo R^2		0.028	0.026	0.034
Obs.		38,185	32,621	19,077

Table 8: Role of accounting resources by type of restatement (errors versus irregularities)

The dependent variable in all models is *RESTATE* that denotes whether the firm restated its earnings in year *t+1*. *ACC_RES (ACC_RES_PRED)* represents accounting resources (predicted value of accounting resources). This panel splits restatements into three groups based on the (tercile of the) market reaction to the announcement of the earnings announcement. The "Low" market reaction group represents irregularities while the "Medium/High" group represents errors. All regressions include year and industry fixed effects (defined at the 2-digit SIC code level). Robust standard errors clustered by firm are presented in parentheses under the coefficients. (***), (**), (*) denote statistical significance at the 1%, 5% and 10% levels, respectively. The incentive variables are *ROA*, *RET*, *NCAP*, *LEV*, *MTB*, *SGR*, and *SIZE*

		Low		**Med/High**	
Median *CAR*		-0.049		0.027	
Dep. variable		*RESTATE*		*RESTATE*	
		(Single stage)	(2SLS: 2nd stage)	(Single stage)	(2SLS: 2nd stage)
		(1)	(2)	(3)	(4)
	Pred.				
ACC_RES		**-0.177** **[0.039]*****		**-0.071** **[0.029]****	
ACC_RES_PRED	**?**		**-0.166** **[0.203]**		**-0.242** **[0.145]***
Partial *F* test of incentive variables: Without *ACC_RES* With *ACC_RES*		 14.26** 11.32		 37.26*** 36.93***	
Marginal effect at: Bottom decile of *ACC_RES* Top decile of *ACC_RES*		 0.025 0.017	 0.025 0.017	 0.046 0.040	 0.057 0.033
Control variables		Yes	Yes	Yes	Yes
Year effects		Yes	Yes	Yes	Yes
Industry effects		Yes	Yes	Yes	Yes
Pseudo R^2		0.024	0.022	0.031	0.026
Obs.		32,164	32,422	32,438	32,438

Table 9: Exploring the role of managerial incentives

The sample is restricted to firms with compensation data. The dependent variable is *RESTATE*. Model (1) presents the single-stage and Models (2) and (3) present 2SLS. *CEODELTA* (*CEOVEGA*) denotes sensitivity of CEO wealth to the stock price (volatility). Similar definitions extend to the CFO. All models include year and industry fixed effects and robust standard errors clustered by firm. (***), (**), (*) denote statistical significance at the 1%, 5% and 10% levels.

Designs		Single-stage		2SLS	
Dep. variable		*RESTATE*		*ACC_RES*	*RESTATE*
	Pred.	(1)	(2)	(3)	(4)
ACC_RES	**–**		**-0.138**		
			[0.044]***		
WINDOW	**+**			**0.041**	
				[0.003]***	
ACC_RES_PRED	**–**				**-0.620**
					[0.210]***
UNEXP_FEES		0.345	0.318	-0.168	0.255
		[0.145]**	[0.144]**	[0.049]***	[0.150]*
ROA	–/+/–	-0.503	-0.466	0.464	-0.224
		[0.322]	[0.325]	[0.140]***	[0.333]
RET	–/+/–	-0.008	0.006	0.061	0.056
		[0.104]	[0.104]	[0.022]***	[0.106]
NCAP	+/+/+	-0.055	-0.060	-0.023	-0.076
		[0.096]	[0.097]	[0.024]	[0.097]
LEV	+/?/+	0.853	0.836	-0.116	0.766
		[0.266]***	[0.267]***	[0.090]	[0.271]***
MTB	+/+/+	-0.111	-0.113	0.008	-0.125
		[0.045]**	[0.045]**	[0.015]	[0.046]***
SGR	+/+/+	0.327	0.308	-0.187	0.225
		[0.112]***	[0.115]***	[0.045]***	[0.119]*
CMPLX	?/+/?	0.095	0.099	0.026	0.113
		[0.068]	[0.068]	[0.026]	[0.068]*
RETVOL	?/+/?	0.354	0.263	-0.988	-0.071
		[0.752]	[0.753]	[0.206]***	[0.760]
SIZE	?/?/?	-0.111	-0.102	0.127	-0.075
		[0.044]**	[0.044]**	[0.020]***	[0.047]
CEODELTA	+/?/+	0.063	0.056	-0.034	0.037
		[0.043]	[0.043]	[0.015]**	[0.044]
CEOVEGA	+/?/+	-0.008	-0.005	0.022	-0.001
		[0.042]	[0.042]	[0.012]*	[0.042]
CFODELTA	+/?/+	-0.005	-0.003	0.021	-0.009
		[0.063]	[0.063]	[0.020]	[0.063]
CFOVEGA	+/?/+	-0.019	-0.021	-0.020	-0.018
		[0.064]	[0.064]	[0.019]	[0.064]
p value of partial *F*-test				<0.001	
Marginal effect at: Bottom (top) decile			0.069 (0.053)		0.102 (0.032)
Year effects		Yes	Yes	Yes	Yes
Industry effects		Yes	Yes	Yes	Yes
Adj./Pseudo R^2		0.039	0.041	0.210	0.041
Obs.		10,115	10,115	10,245	10,115