Auditor-Client Compatibility and Audit Quality*

Mert Erinc
BI Norwegian Business School
mert.erinc@bi.no

Tzachi Zach
The Ohio State University
zach.7@osu.edu

October 2024

Abstract

We develop a new auditor-client fit metric based on topical compatibility between auditors and their clients by combining the results of PCAOB inspections with clients' disclosures of their critical accounting policies. We show that auditor fit is negatively related to several traditional audit quality proxies, including restatements, abnormal accruals, and the likelihood of an auditor missing a material weakness in internal controls. Moreover, we report that our new proxy performs better than traditional measures of auditor-client compatibility, such as over twenty versions of industry specialization. We document that auditor fit is positively associated with higher levels of real earnings management, consistent with stronger auditor oversight imposing higher costs on accrual earnings management. Our proxy demonstrates the importance of the results of PCAOB's inspections, especially when combined with mandatory disclosures about companies' critical accounting policies.

Keywords: Audit Quality, PCAOB Inspections, Critical Accounting Policies.

* We thank participants at the 2023 European Accounting Association Congress, the 2023 International Symposium on Audit Research, the 2023 AAA Annual Meeting, EARNet 2023, HARC 2024, the brown bag seminar at BI Norwegian Business School, and the seminar at Maastricht University. We also appreciate comments from Bethany Brumley (discussant), Gopal Krishnan (discussant), Francine McKenna, Mikhail Pevzner, Heidi Vander Bauwhede (discussant), Yufan Dong (discussant), Ann Vanstraelen, Iver Wiertz, Marleen Willekens, and Ally Zimmerman.

1. Introduction

We develop a new auditor-client compatibility measure by combining the deficiencies of the nine largest U.S. auditors reported by the PCAOB with their clients' most critical accounting areas as disclosed in clients' 10'Ks. The resulting *auditor fit* metric captures the degree to which an auditor can provide a high-quality audit to its specific clients. Our new measure is based on topical compatibility between the auditor and the client. It aims to quantify the extent to which area-specific weaknesses (strengths) at the auditor compromise (enhance) the quality of an audit engagement based on its specific challenges.

The quality of financial statements and corporate earnings is central to the information environment in financial markets. Dechow, Ge, and Schrand (2010) state that "Higher quality earnings provide more information about the features of a firm's financial performance that are relevant to a specific decision made by a specific decision-maker." The literature on financial statement quality has spanned many areas, including developing accounting and auditing standards, managers' informative and opportunistic accounting choices, and auditors' production functions and expertise. Understanding the determinants of financial reporting quality is crucial for expanding the pool of capital in financial markets. Our goal in this study is to understand better how improved matching between an auditor and its client contributes to the quality of financial statements through one of its important determinants – audit quality, which plays a crucial role in shaping financial reports.

Past studies have spent considerable effort to understand the sources and determinants of audit quality. Francis (2023) emphasizes the importance, yet elusive nature, of audit quality. DeFond and Zhang (2014) define audit quality as "greater assurance that the financial statements faithfully reflect the firm's underlying economics, conditioned on its financial reporting system and innate characteristics." The literature has identified several underlying constructs that are expected to determine the level of quality an auditor provides. Such determinants include the competence of audit staff, the effectiveness of testing procedures, and the audit firm's ability to develop effective procedures and training (Francis, 2011). The literature has also offered ways to empirically capture audit quality with proxies such as restatements and abnormal accruals (Aobdia, 2019).

Auditor-client compatibility plays a central role in determining audit quality and fees. The primary variable the literature has used for auditor-client compatibility has been auditor industry specialization, beginning with Solomon et al. (1999). Our goal in this study is to evaluate auditor-client compatibility based on the characteristics of the client's financial reporting environment and the elements most crucial to it, as well as the auditors' strengths and weaknesses with respect to those same elements.

To develop our auditor-client compatibility variables, we use two mandatory disclosures. First, we utilize the disclosures of critical accounting policies (CAPs) to identify the most prominent areas for each SEC registrant. In developing these disclosures, the SEC argued that their goal is to "enhance investors' understanding of the application of companies' critical accounting policies…" (SEC 2002). Indeed, these disclosures provide investors with guidelines about which areas of the financial statements are most critical to understand and are most sensitive to measurement issues through disclosures about: "what gave rise to the initial adoption; the impact of the adoption; the accounting principle adopted and method of applying it; and the choices it had among accounting principles." (SEC 2002). For example, the 2015 annual report of PepsiCo Inc. lists the following CAPs: revenue recognition, impairment of goodwill, income tax, and pension plans. Presumably, those areas require the most robust auditor oversight.

Second, we collect the PCAOB inspection reports of all audit firms that are annually inspected.[1] These annual inspection reports identify areas of weakness at the auditor based on inspecting samples of client engagements. For example, the 2016 inspection report of PepsiCo's auditor, KPMG, summarizes the reviews of 51 of KPMG's 2015 engagements and finds deficiencies in the following areas, where the number of deficiencies is in parentheses: allowance for doubtful accounts (13), business combinations (7), cash & cash equivalents (8), impairment of goodwill (11), income taxes (3), inventory valuation (5), investments (6), PP&E (2), revenue recognition (9), and mortgage valuation (1).

The intuition for our metric is straightforward. Suppose the PCAOB has identified areas in which the auditing process of an audit firm is deficient. In that case, the audits of clients most affected by these areas, as identified by the clients' disclosures, are expected to be weaker. A weak audit process is a major contributor to low audit quality and is likely to result in lower-quality financial statements. Continuing with the PepsiCo-KPMG example, out of the four PepsiCo CAPs, KPMG was found to be weak in three: impairment of goodwill, revenue recognition, and income taxes. Our auditor fit measure assesses that this auditor-client pair has a weak fit in 2015, based on the areas of weakness and the preponderance of engagements in which a particular weakness was found.

While the intuition for our metric suggests that such topical compatibility will be informative to capital markets, the information gleaned from PCAOB inspection reports may not be useful. For example, Lennox and Pittman (2010) find that audit clients do not view these inspection reports as valuable tools for assessing audit quality. Some observers argued that the

[1] These include the Big Four firms as well as the next five largest auditors: PwC, EY, Deloitte, KPMG, Grant Thornton, BDO, RSM, Marcum, Crowe.

inspection reports only provided information about "extreme" deficiencies, and their information content could be improved further (Brown, 2020). Indeed, in response to concerns that the inspection reports are not sufficiently informative, the PCAOB introduced in 2020 a new format for the report that provides more details about the inspections to address such concerns (Brown, 2020). However, Brown (2020) continued to argue that the reports are still not as informative as he would like them to be. If the information from PCAOB inspections is not indicative of the overall quality of the audit process in an audit firm, then our auditor fit measure might not effectively identify engagements with low-quality audits. It is also possible that the CAP disclosures are strategic and incomplete. If, for example, an SEC registrant omits a critical policy from its disclosure, then our auditor fit measure may not effectively explain audit quality.

Our findings can be summarized as follows. First, we document that higher levels of auditor fit are associated with higher audit quality. For example, our auditor fit proxy is negatively associated with restatement frequency, abnormal accruals, and the likelihood of an auditor missing a material weakness in internal controls. These results hold in models containing industry and year-fixed effects, and using a one-step estimation procedure for abnormal accruals as advanced by Chen et al. (2018). In addition, we find that auditor fit is associated with stronger auditor oversight as proxied by the substitution away from accrual earnings management to real earnings management. Second, we uncover evidence that the reasons for restatements are related to areas in which clients and auditors lack fit, that is, areas of import to clients and in which auditors were found deficient.

Third, our auditor-client compatibility measure is more strongly associated with audit quality compared to other auditor-client compatibility measures, such as over twenty industry specialization variables and a client similarity score (Brown and Knechel, 2016). Finally, we show that investors seem to appreciate stronger compatibility between an auditor and its client. We document that the earnings response coefficients (ERCs) of clients with greater auditor compatibility are significantly larger. In these tests, both our proxy for auditor fit as well as the industry specialization variables are associated with higher ERCs. Interestingly, client similarity scores are associated with lower ERCs.

We perform several analyses to demonstrate the robustness of our fit measure. It is possible that the PCAOB inspections' results are only confined to the inspected sample and do not extend beyond it to the audit firm. In contrast, we provide evidence that the deficiencies detected by the PCAOB are predictive not only of future restatements by clients of the audit firms but also of the reasons for the restatements. That is, clients' restatements are more likely to be for reasons related to their auditors' deficient areas. This provides some confidence that the inspections' results indicate more systematic issues at the audit firm at large. Also, because the fit measure depends on the PCAOB's inspections' agenda, areas that the PCAOB neglects and does not inspect might bias our metric upward. We adjust our fit measure in robustness analyses to account for this possible neglect. We show that such adjustments do not impact our overall conclusions.

Our study contributes to the literature in several ways. First, we demonstrate that auditor-client topical compatibility is an important contributor to higher audit quality. As such, we advance the literature that studies the determinants of audit quality. Moreover, we document that our auditor fit measure performs better than industry specialization proxies used in the literature over the last two decades. This finding underscores the importance of understanding the topical alignment between an auditor and their clients and suggests that industry specialization is not able to capture this aspect of the auditor-client relationship. This finding is also consistent with Minutti-Meza (2013), who questions the ability of industry specialization to capture audit quality.

Second, our new auditor-client compatibility variable differs from and is incremental to other recently emerging compatibility proxies. For example, Brown and Knechel (2016) developed a measure based on the similarities of a client's narrative financial disclosures (e.g., MD&A and footnote sections of the 10-K) to other clients of the same auditor. This similarity score is an advanced way to capture auditor familiarity with and specialization in groups of companies that are similar to each other. On the other hand, our measure captures a different aspect of the audit production function. It is based on the topical compatibility (or lack thereof) between the auditor and its client. Indeed, the results are consistent with our new measure capturing complementary aspects of the audit process compared to the similarity score.

Third, we also contribute to our understanding of the usefulness of two prominent disclosures: the PCAOB inspection reports and corporate disclosures of companies' critical accounting policies. Some commentaries and researchers have shown and argued that the PCAOB inspection reports are not informative (e.g., Lennox and Pittman, 2010; Brown, 2020). One argument was that the use of deficiency rates is an invalid method for measuring the audit quality of U.S. auditors because, for example, the sample is not representative of the population, changes in engagement selection criteria are not considered, and differences in inspection team competency are not evaluated (e.g., Peecher and Solomon 2014; Palmrose 2013). Evidence in Acito et al. (2018), among others, suggests that the PCAOB inspections are informative about aspects of the auditor-client relationship. Our study is consistent with these results and shows that the *combination* of PCAOB inspection reports and CAP disclosures does provide information that can be used to assess auditor fit and audit quality.

Fourth, our study adds to our knowledge about the role of CAP disclosures in capital markets. As such, we complement the findings in Szerwo (2023). Szerwo (2023) finds that some

CAP disclosures are informative with respect to future restatements and that investors respond to restatements announcements more negatively if those restatements pertain to CAPs. Our study suggests that CAP disclosures are informative in assessing the overall fit between an auditor and its clients. Of course, our results do not speak to the degree of informativeness of these reports, and it is certainly possible that they could be more informative.

Our study proceeds as follows. In section 2, we overview the three related literatures to the study: audit quality and auditor-client compatibility, PCAOB inspection reports, and disclosures of critical accounting policies. Section 3 describes our data collection process and the methodology we use to develop our auditor fit metric. In section 4, we report the results, and we conclude in section 5.

2. Literature Review

2.1 Audit quality and auditor-client compatibility

According to Francis (2023) "*Nothing is arguably more important in auditing practice, regulation, and research than the concept of audit quality. Yet, despite its centrality to the understanding of auditing, defining it has proved elusive, and there is little consensus on the concept or the measurement of audit quality*." DeFond and Zhang (2014) provide a comprehensive summary of the literature on audit quality, its determinants, and consequences. Most attempts to capture audit quality involve the assessment of financial reporting outputs, such as restatements, and properties of financial statements, such as discretionary accruals. As DeFond and Zhang (2014) argue, these approaches suffer from the confluence of audit quality, the clients' innate characteristics, and the financial reporting system. An effective audit quality measure aims to "disentangle the effect of audit quality" from other clients' features unrelated to the audit process itself. To achieve that, the literature has turned to input-based measures of audit quality. These include auditor identity (e.g., Big N), specialization, and contract features between the auditor and the client (e.g., level of audit fees). Researchers have recently attempted to evaluate audit quality using information on individual partners in international settings (e.g., Aobdia et al., 2015). With the advent of Form AP in 2016, there have been some similar attempts with U.S. data (e.g., Burke et al., 2019; Cunningham et al., 2019).

To our paper, the most related area of this literature involves studies that combine auditor and client characteristics to evaluate auditor-client compatibility. Since Solomon et al. (1999), the literature has commonly adopted industry specialization to proxy for auditor's knowledge and expertise in the client's industry and audit quality in general. Despite its popularity as a proxy, there is no unanimity regarding how it should be measured. Studies have employed various empirical definitions of industry specialization. Generally, the proxy is used as an indicator that takes the value of one if the auditor's market share in the client's industry meets certain criteria.[2] The proxy's construction is affected by choices in three different layers. First, the market share can be based on audit fees (Francis et al., 2005; Reichelt and Wang, 2010; Guo et al., 2022), the client's total assets (Behn et al., 2008; Minutti-Meza, 2013) or client's total sales (Lim and Tan, 2008; Cao et al., 2012). Second, the market share can be accumulated and compared at national level (Payne, 2008; Bae et al., 2017), city/office level (Cassell et al., 2019), partner level (Dekeyser et al., 2023), auditor team (Cahan, Che, Knechel and Svanstrom, 2022), or multiple levels at the same time (Reichelt and Wang, 2010). Third, the condition for the auditor to qualify as an industry specialist can be based on having the highest market share in the industry (Lim and Tan, 2008; Bae et al., 2017), exceeding a cutoff share such as 30% (Bills et al., 2015; Guo et al. 2022), and/or

[2] In a few studies, industry specialization is measured as the market share itself, i.e. a continuous variable (e.g., Behn et al., 2008).

exceeding a cutoff share difference with the follower auditor's market share (Reichelt and Wang, 2010; Cao et al., 2012; Minutti-Meza, 2013).

Extant evidence for the association between industry specialization and audit quality is mixed. Earlier studies predominantly provide evidence of a positive association, using discretionary accruals and earnings response coefficients (ERC) as audit quality proxies (Krishnan, 2003; Balsam et al., 2003; Kwon, Lim, and Tan, 2007; Gul, Fung, and Jaggi, 2009), and experimental methods (Low, 2004; Hammersley, 2006). Dunn and Mayhew (2004) find that this positive association exists in unregulated industries, based on analysts' disclosure quality rankings. Payne (2008) finds that clients of specialist auditors are associated with less benchmark-beating behavior. Behn et al. (2008) document that earnings predictability, proxied by higher forecast accuracy and lower forecast dispersion, is higher for clients of an industry specialist only if the auditor is not a Big 5 auditor. Reichelt and Wang (2010) find that industry specialization is associated with lower discretionary accruals, higher ERCs, and a higher likelihood of issuing a going-concern opinion only if the auditor is an industry specialist at both national and city levels.

Minutti-Meza (2013) points to methodological issues with using within-industry market shares to proxy for specialization, as he argues that client characteristics that are correlated with choosing a high-market share auditor are also associated with commonly-used audit quality proxies. In other words, clients who prefer and can afford an industry-leading auditor are more likely to have better disclosure quality. Using a matched sample besides the full sample, and employing the same audit quality proxies as Reichelt and Wang (2010), he documents that there is no significant association between them and the auditor's industry specialization in the matched samples. More recently, Dekeyser et al. (2023) documented using a new "industry range" construct (i.e., the opposite of industry specialization) that audit partners with experience in a wider range of industries are more likely to require audit adjustments from their clients, indicating higher audit quality.

Other attempts to capture auditor-client mismatch involved determining whether a client is properly matched with a Big Four audit firm or not, based on the coefficients of a prediction model estimated for the entire Compustat universe (Shu, 2000; Landsman, 2009; Schroeder and Hogan, 2013).

In our paper, we capture auditor-client compatibility based on more granular and fundamental characteristics of both the auditor and the client. Our approach is similar in spirit, although different in its details and goals, to two studies. First, Brown and Knechel (2016) develop a similarity score of auditor-client compatibility based on the closeness of the textual disclosures of one client to all other clients of its auditor that belong to the same industry. They find that clients with low similarity scores with other companies in the same industry-year-auditor cluster are more likely to switch auditors. Their new auditor's clients have higher similarity scores. They also find that similarity scores are negatively associated with discretionary accruals, hinting at a possible positive relation between these scores and audit quality. However, based on their results regarding going concern opinions, the ability of the similarity score to proxy for audit quality is inconclusive.

Second, Acito et al. (2018) evaluate how an auditor's performance in PCAOB inspections and its exposure to a particular client are related to audit fees and auditor turnover. They find that engagements with high exposure, based on both auditor and client characteristics, are positively associated with audit fees and, for smaller clients, negatively associated with auditor turnover. In a sense, the building blocks of their analyses are similar to ours, in that a mismatch between a client and an auditor implies higher audit risk. We differ from Acito et al. (2018) because our tests are geared towards evaluating whether auditor fit is an important determinant of audit quality whereas Acito et al. (2018) are interested in the business relationship between the auditor and the client (audit fees and viability of the engagement). We further depart from Acito et al. (2018) by how we identify clients' exposure to certain accounting areas. Unlike Acito et al. (2018), who use the Folsom et al. (2017) dictionary for U.S. GAAP areas, we rely on the clients' direct disclosures of their critical accounting policies.

2.2 PCAOB inspection reports

We use the inspection results of the PCAOB as one of the inputs to our auditor fit metric. As such, we also contribute to the literature that examines the informativeness of the PCAOB reports, particularly the inspection results section. The literature on the PCAOB reports started with an early examination by Lennox and Pittman (2010), who provided the initial evidence on the matter. They document that the deficiencies disclosed in PCAOB inspection reports fail to predict subsequent changes in audit firms' market shares and thus are uninformative to clients. DeFond (2010) is reluctant to adopt this interpretation, however. He argues that the inspection reports do not have to be informative to be effective, and that the increased scrutiny provided by the PCAOB, in and of itself, can improve audit quality, irrespective of the content of the inspection reports. DeFond (2010) also argues that the reason why the inspections are not informative may be because the PCAOB targets the riskiest audit issues among the riskiest clients. The clients might not consider this a representative sample that can speak to the auditor's performance in the whole client base.

Later studies provide some evidence that the inspection reports are informative. For example, Abbott et al. (2013) show that GAAP-deficient triennially inspected (small) auditors are more likely to be dismissed in favor of non-deficient other small auditors. Relatedly, Gunny and

Zhang (2013) show that inspection deficiencies are indicative of low audit quality, as captured by abnormal accruals and future restatements. However, their results only applied to smaller auditors, and were not evident in auditors who are inspected annually. In another example, Aobdia (2018) shows (with proprietary data) that inspections are informative for auditors themselves, as he finds audit firm effort increases in inspected and non-inspected engagements of offices or partners that receive a Part I Finding. On the other hand, audit effort decreases in engagements that were inspected but did not receive a Part I Finding. The study also shows that clients whose audits received a Part I finding are 18% more likely to switch auditors. Finally, Abbott et al. (2024) document that audit firms respond to revenue-related deficiencies, leading to improvement in their clients' revenue quality. This implies that the inspections provide useful information that can enhance the quality of audits.

2.3 Critical Accounting Policy Disclosures

The second input to our auditor fit measure is companies' disclosures of critical accounting policies. These disclosures started in 2003. The accounting literature on CAP disclosures and their relatives – critical accounting estimates, is relatively thin, and there is not much evidence of their informativeness to capital market participants. Levine and Smith (2011) provide the first evidence on these disclosures. They show that "On average, firms disclose six to seven policies as critical, with the most commonly cited five policies (revenue recognition, taxes, contingencies, marketable securities, and impairments) covering about half of the total disclosures. Glendening (2017) finds, using a sample of S&P 500 firms, that the predictive value of earnings with respect to future cash flows is negatively associated with the presence of a critical accounting estimate disclosure. Glendening, Mauldin, and Shaw (2019) find that quantitative CAEs are negatively associated with management's incentives to misreport (proxied by portfolio

Vega) and positively associated with audit committee accounting expertise and audit offices with multiple quantitative CAE clients. Szerwo (2023) demonstrates that some, but not all, CAP areas are informative regarding future restatements. Particularly, he finds that revenue, derivatives, accruals and short-term liabilities, and capitalization of expenditures, are the areas reported by companies that seem to be correlated with future restatements. Moreover, he shows that investors' response to a restatement is more negative if the restatement occurs in an area previously-reported by the company as CAP. Our study shows that combining the CAP disclosures with other relevant information can provide information about audit quality to capital market participants.

3. Data

3.1 Data collection and classification

To construct our sample, we collect all the CAP disclosures available in Calcbench. This marks the beginning of our sample period as 2009, because Calcbench's parsing relies on the interactive (XBRL-encoded) 10-K documents. We extract all the CAP disclosures through 2019 using Calcbench, which provides access to the parsed SEC documents.

Critical Accounting Policies (CAP) Disclosures. Our first goal is to build an extraction algorithm to identify the CAP section in 10-Ks because Calcbench does not tag it separately. To achieve that, we begin with the 10K's of the 30 Dow Jones companies. After manually confirming a success rate of above 97% in identifying the CAP section in this training sample, we employ the algorithm over all companies in the Calcbench universe.[3] We test the success of our algorithm over a random set of companies from our entire sample and find that the algorithm is successful in

[3] There are several common titles that companies use to identify their CAP discussions, including "Critical Accounting Estimates" and "Critical Accounting Policies and Estimates." Calcbench's API lets us separately extract the MD&A section of each 10-K, allowing a better and easier identification of the CAP sub-section. This makes us confident that our algorithm does not extract the Note to Financial Statements with the similar title: "Summary of Significant Accounting Policies".

identifying the CAP section in over 99% of company-years. Table 1 summarizes our sample selection procedure. Our initial extraction starts with 8,622 companies and 55,720 company-year observations.

Next, we identify and extract the sub-headers within the CAP disclosures which include the exact critical accounting areas. Using the pool of extracted CAP sub-headers, we identify 24 distinct accounting areas that appear as CAPs. Accounting for the variety of words that companies use to refer to the same areas, we generate a dictionary of areas, and adopt it to code each company-year as either disclosing or not disclosing the respective area as CAP.[4] For each CAP area, we assign a dummy variable $CAP_{i,t,a}$ where a is the accounting area code. This step eliminates 855 companies and 7,194 company-years that do not include any one of the 24 common and distinct critical accounting areas. Next, we eliminate 1,678 company-years where the auditor is not among the nine annually audited big auditors. After several data requirement conditions, we end up with a sample of 28,435 company-years encompassing 4,354 companies. The number of observations used in the analyses depends on the adopted compatibility measure. The regressions that include *SPEC_OFFICE* exclude city-industry-year clusters with fewer than two observations following Minutti-Meza (2013), leaving 24,363 company-years. Similarly, the analyses that include *SIM_COMB* miss observations due to small auditor-industry-year groups, following Brown and Knechel (2016), leaving 13,326 company-years.

Table 2 reports the occurrence rates of each CAP area conditional on the company's industry based on the Fama-French-12 classification scheme. Each cell includes the fraction of companies in a particular industry that report a certain accounting area as critical. For each CAP

[4] More details about our CAP identification process are available upon request.

area, we highlight in shaded cells the two industries in which the area is most common. Similarly, for each industry, we denote the two most common CAP areas in bold font. We observe that *Revenue Recognition* (Area #22), *Income Taxes* (Area #11), and *Impairment of Goodwill and Other Long-Lived Assets* (Area #10) are the most popular CAP areas, regardless of industry. On average, these areas appear as CAP in 59%, 58%, and 57%, respectively, of all 10-Ks during our sample period. These numbers are similar to the reported frequencies in Levine and Smith (2011).

Other areas seem to be common to some industries but not to others. For example, 51% of companies in the Money industry report CAPs related to *Allowance for Credit Losses* (Area #1), as expected, given the importance of credit in this industry. In the same vein, 52% of companies in the Money industry report Investments (Area #14) as important. Similarly, 53% of companies in the Shops industry (i.e., wholesalers and retailers) emphasize inventory valuation in their CAP disclosures. Pensions are popular in Utilities, with 56% of companies in that sector reporting them as critical. Finally, stock-based compensation is a critical accounting policy in the Health industry, with 67% of the companies reporting it as such. The shaded cells help to illustrate which industries are affected the most by a given CAP area, even if the area itself is not popular in general. For instance, R&D Costs (Area #21) do not appear frequently in CAP disclosures (only 13%). However, they are relatively more prevalent in the Health (37%) and in the Business Equipment (18%), covering Technology and software companies.

Figure 1 illustrates the time-series variation of occurrence rates of the ten most popular CAP areas over the sample period. We observe that CAPs, such as revenue recognition, income taxes, and goodwill impairments, are quite common, hovering around 60% throughout most of the sample period, with a slight decline in recent years. Most other CAP areas exhibit frequencies of around 20% to 30%. The steepest declines in frequencies over our sample period occur in allowance for credit losses (from 40% to 25%) and inventory valuation (from 30% to 20%).

PCAOB Inspection Reports. Next, we gather information from the PCAOB's website on the results of 144 inspections published between 2004 – 2020 of all audit firms that were subject to annual inspections by the PCAOB as of 2020. To identify the areas of deficiencies reported by the PCAOB, we use the same dictionary of area-relevant keywords that we developed for client CAPs.[5] After identifying the deficient areas, we count the number of deficiencies per area in each inspection report. For every inspection report of audit firm A, at time t, we count the number of engagements that were found deficient in area a, and label the count as $Num\_def_{A,t,a}$.[6] For example, if revenue recognition is identified as a deficiency area in 6 of the 10 engagements reviewed for Deloitte in 2009, then $Num\_def_{A,t,a}$ for that observation is counted as 6.

Table 3 reports the annual number of engagements that were inspected by the PCAOB at each audit firm in our sample. Naturally, the Big Four are subject to more engagement reviews that are inspected each year. The average number of engagements for a Big Four per year during our sample period is 54, while this average is about 19 for the next five auditors in our sample.

Table 4 provides summary statistics for the number of deficiencies found in the inspection reports for each area across all nine audit firms throughout the 2004 to 2020 period. Because we use the same dictionaries to identify CAP areas and deficiency areas, the number of the deficient areas reported in Table 4 corresponds to the number of CAP areas reported in Table 2. The most prevalent deficiency areas are revenue recognition (Area #22) with 669 deficiencies, allowance

[5] We identify 3 additional auditor-specific areas that do not appear in clients' CAP disclosures: Assessment of going-concern, Fraud Risk, Internal Control over Financial Reporting (ICFR). We exclude them.
[6] We denote in t the fiscal year for which an inspection is performed. Typically, this inspection is performed during the following fiscal year, t+1.

for credit losses and doubtful accounts (Area #1) with 518 deficiencies, and fair value of assets and liabilities (Area #8) with 339 deficiencies. When divided by the number of inspections, these deficiencies translate to 4.6, 3.6, and 2.3 deficiencies per inspection respectively for each area. Figure 2 plots the time trend, over our sample period, of the number of deficiencies, separately for each audit firm in our sample.

The number of deficiencies could vary by the quality of the auditor, but also by the number of engagements that the PCAOB reviewed for each auditor. To control for the number of engagements, and to limit the impact of a few engagements that could potentially have undue influence on our measure, we compute for each auditor-year-area a weakness score, $Weakness_{A,t,a}$ as follows:

$$Weakness_{A,t,a} = \frac{Num\_def_{A,t,a}}{Eng\_Rev_{A,t}} \tag{1}$$

where $Num\_def_{A,t,a}$ is the number of deficiencies auditor A received for area a across all engagements that were subject to PCAOB inspections for fiscal year t. This number is scaled by $Eng\_Rev_{A,t}$, which is the total number of engagements of audit firm A that were reviewed by the PCAOB for fiscal year t.

Figure 3 illustrates the variation in each auditor's weakness score over time for the ten most common deficiency areas. One prominent pattern is the high deficiency rates of KPMG in the area of "allowance for credit losses" from 2011 to 2017. This pattern served as the backdrop for the audit inspection scandal of 2017 (McKenna et al., 2022a, 2022b).

3.2 Measuring Auditor-Client Compatibility

3.2.1 Our FIT variables

We now describe *FIT*, the new measure of the fit between an auditor and its clients. *FIT* is based on the critical accounting policies for each client, and whether the auditor is deemed deficient by the PCAOB in those areas. The idea is to identify pairs of auditors and clients in which audit quality might be weaker or stronger based on signals that originate at the client and at the PCAOB. The intuition is as follows. If a client has two critical accounting areas, say revenue recognition and asset impairments, and if the auditor is deemed deficient by the PCAOB in both areas, then the fit between the auditor and the client is weak from an audit quality perspective.

To operationalize *FIT*, we need to match a client's 10-K for fiscal year *t* with its auditor's PCAOB inspection report for the same fiscal year. The inspections underlying this report are typically conducted in fiscal year *t+1* and the report is issued later in that year or even in the year that follows.[7] Since our aim is to assess the fit between the auditor and the client during the fiscal year in which the audits were conducted, we match a client's CAP disclosures in fiscal year *t* to the auditor deficiencies for that same fiscal year as they are revealed in PCAOB reports that are issued later.

We build our *FIT* measure based on the 24 areas that we identified in CAP disclosures, as described earlier. We calculate it as follows:

$$FIT_{i,A,t} = \frac{\sum_{a=1}^{24}(CAP_{i,t,a} \times Weakness_{A,t,a})}{\sum_{a=1}^{24} CAP_{i,t,a}} \times -1 \qquad (2)$$

[7] E.g. KPMG's 2017 Inspection Report: "The inspection procedures included reviews of portions of the Firm's work on 52 issuer audits, which generally related to issuer year ends in 2016."

where $the\ CAP_{i,t,a}$ is an indicator for whether client i included area a in its CAP disclosures in year t, and $Weakness_{A,t,a}$ is as defined above. The intuition of this measure is as follows. For each auditor-client pair, we compute the average, across all CAPs that are relevant for client *i* in fiscal year *t*, of the weakness ratios of the client's auditor, *A*, during that same fiscal year. In other words, we sum up the weakness scores of the auditor in the areas its client reported as CAP, and scale it by the total number of CAPs reported. Because this raw average is decreasing in the quality of the fit, we multiply it by (-1). As a result, higher values of *FIT* correspond to a better fit between the client and the auditor.

Figure 4 illustrates the trends in *FIT* for each of the nine auditors. For each audit firm and fiscal year, the graph reports the simple average of *FIT* across all of the auditor's clients. A noteworthy observation is the continuous increase in *FIT* at Deloitte. Part of this increase can be explained by a general improvement in deficiencies at Deloitte, as illustrated by Figure 2. PwC's *FIT* is generally stable over time, with a slight jump in 2019, coinciding with a sudden drop in PwC's deficiencies. EY's *FIT* has decreased in the early years, reaching a low in 2011 that coincided with a peak in its deficiencies (see Figure 2), but has been rising since. KPMG's *FIT* pattern dips in 2015, shortly before the PCAOB-KPMG scandal, and has been rising since. BDO experienced a significant decline in its fit until 2013, and since then, it has gradually increased. We observe such an increase at GT as well. *FIT* has been generally stable for RSM, and for Crowe, although Crowe's has been more volatile.

Our fit measure relies on the notion that the PCAOB inspections' results are not confined to the inspected sample, but extend beyond it to the audit firm at large. That is, the inspections uncover some systematic issues at the auditor that apply to their entire client portfolio.[8] To evaluate this assumption, we investigate whether the areas found to be deficient for auditor a in the PCAOB's inspection samples are related to the reasons of restatements of the clients in auditor a's portfolio. Such analysis evaluates whether the findings of the PCAOB inspections have the potential to extend to the auditor at large. We estimate the following logistic model, separately for each of the 24 critical accounting areas:[9]

$$\begin{aligned} RESTATE_{i,t,a} &= \beta_0 + \beta_1 DEF_{A,t,a} + Auditor\text{-}Characteristcis_{A,t} \\ &\quad + Firm\text{-}Controls_{i,t} \ \&\ FE \end{aligned} \tag{4}$$

, where (1) $RESTATE_{i,t,a}$ is equal to one if the reason for the restatement of company i in period t is deemed related to area a. Our assessment of the restatement reason is based on textual analysis of the announcement of a restatement against the keywords we used to classify CAPs, as previously described; (2) $DEF_{A,t,a}$ is equal to 1 if auditor A had a deficiency related to area a in year t. Of the 24 CAP areas, we could only estimate 17 models effectively due to insufficient restatements in 7 areas. Of those 17 models (unreported – results available upon request), we find that the coefficient on $DEF_{A,t,a}$ is positive and significantly different from 0 at the 10% level or better in 7 models. To assess the collective statistical evidence of a relation between the reason for restatement and the deficiency area, we utilize a binomial test to evaluate the likelihood of obtaining seven significant coefficients out of 17 with a base likelihood of 10%. The p-value of this test is extremely small, meaning that the likelihood of seven significant coefficients by chance

[8] Abbott et al. (2024) provide some evidence consistent with this assumption. They show that audit firms respond to PCAOB deficiencies in the area of revenue recognition, and as a result client's revenue quality increases following the publication of inspection results.

[9] We thank Iver Wiertz for proposing this test.

is extremely low. Therefore, we believe that the deficiency areas found in PCAOB inspections are revealing and extend beyond the inspected sample alone.

3.2.2 Other compatibility variables

Together with our proposed new metric for auditor-client compatibility discussed in the previous section, we also evaluate the relation between audit quality and other compatibility metrics used in the literature. The most common approach to proxy for auditor-client compatibility is based on auditors' industry specialization. It is natural to presume that specialization provides auditors with specific skills originating from a wide array of clients in the same industry (Dopuch and Simunic, 1982). These skills theoretically transfer to other engagements and have the potential to scale. Defond and Zhang (2014) discuss the merits of using this approach. They argue that "industry specialization …. provides a measure of quality variation within Big N auditors…(which) allows researchers to address questions that pertain to within Big N quality differences." However, because the specialization metrics are typically measured as indicator variables, they "fail to capture relatively subtle variations in audit quality." Furthermore, there is some debate about the proper way to measure specialization (Neal and Riley, 2004; Minutti-Meza, 2013; Guo et al., 2022), and the level at which specialization should be evaluated (e.g., national, office/city, partner).

In our paper, we employ Guo et al.'s (2022) office-level specialization measure as the main industry specialization proxy. We define an auditor as a specialist if its annual market share, in terms of aggregate audit fees within a Metropolitan Statistical Area (MSA) and a two-digit SIC code, is more than 30 percent. We chose to report our main results using the Guo et al. (2022) metric because it is the most recent one used in the literature to date, utilizing the more specific office/city classification. However, we ran our results with over 20 different industry specialization measures used in the literature over the last 20 years, from nine studies. These included the following: (i) Guo et al.'s (2022) national-level measure, (ii) Minutti-Meza's (2013) national-level and city-level industry leader measures, (iii) Cao et al. (2012)'s measure that utilizes both an absolute and a relative market share threshold to define industry leadership, (iv) Reichel and Wang (2010)'s ten measures of different combination of leadership at the city and national levels, (v) Behn et al. (2008)'s single measure of national leadership using clients' total assets as the basis of determination, (vi) Lim and Tan (2008)'s single measure that uses a cutoff of 30% of audit fees to define industry leadership, (vii) Francis, Reichel and Wang (2005)'s two measures of leadership based on audit fees at the national and city, and one measures based on the combined leadership at both national and city levels, (viii) Dunn and Mayhew (2004)'s two measures based on national market share and dominance defined based on a 20% cutoff, (ix) and Balsam et al. (2003)'s two measures based on national dominance and market share.

Brown and Knechel (2016) introduce an alternative way to measure auditor-client compatibility that is closer in spirit to our approach. Under their approach, auditor-client compatibility is measured by the similarity of client *i* to its auditor's other clients in the same industry group. They judge similarity by the language in three of client *i*'s 10-K sections (Business Description, Management Discussion & Analysis, and Footnotes) compared to the language other clients use in their three 10-K sections. In most of our analysis in this paper, we use the similarity measure that is based on the combined score of these three sections. Like Brown and Knechel (2016), we label this variable *SIM_COMB*.[10] Appendix A provides more detail about the computation of the similarity scores.

[10] As an alternative to *SIM_COMB*, we also considered reference groups that are based on the Text-based Network Industry Classifications (TNIC) metrics developed in Hoberg and Philips (2010, 2016). Our conclusions do not

3.3 Measuring Audit Quality

We use several proxies to measure audit quality. We are motivated by Aobdia (2019), who evaluates audit quality proxies based on their association with the results of PCAOB inspections at the engagement level. Our first proxy is restatements (*RESTATE*), which equals one if company *i*'s financial statements of fiscal year *t* have been restated. We use the Big R definition as Thompson (2023), which is based on the filing of 8-K (Item 4.02) along with the restatement. Our second proxy is abnormal discretionary accruals. Traditionally, abnormal accruals have been measured as residuals from a first-stage model such as the Modified Jones. Abnormal accruals were then used in second-stage models as dependent variables to test various hypotheses (e.g., their relation with auditor identity). Chen et al. (2018) study the use of estimated regressors in second-stage models and find that a two-step procedure frequently yields biased coefficients and incorrect inferences, including in cases when abnormal accruals are the variable of study. As a result, in our paper, we adopt the suggestion of Chen et al. (2018) (p. 782) and employ a single-step procedure to evaluate the relations between abnormal accruals and our auditor-client compatibility variables. A single-step approach involves including as independent variables both the variables that explain "normal" accruals (the traditional variables used in a modified Jones model) as well as the variables of interest (in our case the auditor-client compatibility variables) and other controls, including industry and year fixed effects.[11]

---

qualitatively change. In these untabulated analyses (available upon request) we find that *SIM_COMB* is more likely to be associated with audit quality proxies compared to metrics based on TNIC.

[11]We also tested (i) the accrual quality proxy based on the approach advocated in Dechow and Dichev (2002), using a single-step estimation approach, and (ii) the "small profit" measure advanced in Aobdia (2019). The conclusions are generally similar to those we report in the study, with the results from the Dechow and Dichev (2002) measure tracking closely with those we report on abnormal accruals. All results from these robustness analyses are available upon request.

Our third proxy for audit quality is the non-reporting of material weaknesses in internal controls (*MISSED_MW*), as advanced in Rice and Weber (2012). They find that missed material weaknesses are partly related to the quality of the audit performed. We equate *MISSED_MW* to one in cases where restatements occurred with no warning of material weaknesses in internal controls to precede them, and zero otherwise. Pittman et al. (2023) and Beardsley et al. (2021) also use this proxy as an audit quality surrogate.

3.4 Research Design

In this section, we outline the approach for evaluating the efficacy of our auditor-client compatibility metric. Figure 5 depicts our approach. First, the ability of our auditor-client compatibility measure to capture audit quality is represented by Line 1 in Figure 5. As a first step, when developing a measure such as ours, we would like to evaluate its correlation with the underlying theoretical concept that it may explain, in this case, audit quality. Thus, we seek to achieve some *convergent validity* with the unobservable audit quality (Line 2 in Figure 5). To that end, we examine our measure's association with a variety of empirical proxies that have been used in the literature to proxy for audit quality and that we discussed in section 3.3 (Line 5-7 in Figure 5). The stronger our auditor-client compatibility proxy is correlated with the audit quality proxies, the closer we get to convergent validity. Furthermore, one can compare the performance of our metric with other auditor-client compatibility metrics along this dimension.

One challenge we face is that the empirical proxies we use for audit quality are not perfect. They are, in essence, financial statement quality metrics. That is, they are not only affected by audit quality (Line 4 in Figure 5), but also by the accounting quality of the client (Line 3 in Figure 5). Since our interest only lies in the relation between audit quality and auditor-client compatibility, we also need to evaluate our measure's *divergent validity*. That is, to what extent does our proxy converge to audit quality and diverge from accounting quality? This will depend on the degree to which each empirical proxy we use for audit quality differentially captures audit and accounting quality (Line 8-10 in Figure 5). This construct is unobservable but can be speculated on. In addition, to try and further isolate the impact of accounting quality, our research design will aim to control for the impact of a client's accounting quality on these financial statement quality metrics.

3.5 Summary Statistics

Table 5 provides some summary statistics for the variables that we use in our subsequent tests, over all company-year observations in our sample. Starting with the auditor-client compatibility metric, we find that the average level of *FIT* is -0.06 with some reasonable variation reflected in a standard deviation that equals 0.04. The average similarity score (*SIM_COMB*) we report is 8.79, and is close to that reported in Brown and Knechel (2016). We report statistics for two variants of industry specialization, at the office and national levels, following Guo et al. (2022). Based on the specialists variables, *SPEC_OFFICE* and *SPEC_NATIONAL*, we find that in 45% (20%) of auditor-client pairs, auditors are considered specialists at the office (national) level. These numbers are comparable to those reported by Guo et al. (2022).

About 1.4 percent of company-years experience a BIG R restatement (*RESTATE*). The average level of total accruals (*TACC*) is -0.19 of total assets, while the median is -0.07, which is comparable to the values reported in recent studies (e.g., Brown and Knechel, 2016 and Nallareddy et al., 2020). In about 1% of observations, auditors missed a material weakness in internal controls.

The average number of CAPs (*NUM_CAPS*) reported annually by our sample companies is 4.07 with an interquartile range of 2 to 6. The average auditor tenure (*TENURE*) in our sample is about 12 years, and the average level of audit fees (*AUDIT_FEES*) is around $2.0 million. Sixty-three percent of clients in our sample are audited by a Big Four (*BIG4*) firm. Auditor changes in our sample are rare, occurring in 4% of the observations (*AUDITOR_CHANGE*). The average client in our sample has a market capitalization of $4.3 billion and a BM ratio of 0.38. About 53% of sample companies have foreign operations.

In Table 6, we report the univariate correlations between some of the main variables in our analyses. First, we focus on three of the auditor compatibility metrics – our own *FIT* score, and two others that have been discussed in the literature. We find that *FIT* is negatively correlated with the similarity score from Brown and Knechel (2016) – *SIM_COMB* (-0.14) and positively correlated (0.11) with the auditor specialty variable (*SPEC_OFFICE*). These correlations are important because they show that our auditor-client compatibility measure does not merely capture similarities in clients' accounting policies within the auditor's portfolio or auditor specialization. Given that the three metrics are purported to capture some dimension of auditor-client compatibility, they do not appear to be overlapping. Although the correlations are low (and sometimes negative), we do not believe that they reflect any drawback to any one of the three metrics. That is, they do not suggest that these metrics cannot each achieve convergent validity with auditor-client compatibility. Because there is no consensus agreement in the literature that any one of these metrics captures auditor-client compatibility perfectly, then any univariate correlations between them are merely descriptive, and do not reflect on the efficacy of any one of the metrics to capture auditor-client compatibility.

We note some of the other interesting correlations. With respect to audit quality metrics (*RESTATE* and *MISSED_MW*) we find negative, but small, univariate correlations with *FIT*, positive, but small, correlations with *SIM_COMB*, and no significant correlations with *SPEC_OFFICE*.

4. Results

4.1 Preliminaries – Determinants of FIT

We begin our analysis by familiarizing ourselves with the new variable we constructed, *FIT*, whose purpose is to capture the quality of the matching between the auditor and its clients. Recall that *FIT* is based on the performance of the auditor in PCAOB inspections in the areas that are most critical to the specific client, based on the client's CAP disclosures. We explore several specifications of the levels and changes in *FIT*. We include a variety of client-based explanatory variables, such as *SIZE* and *LEVERAGE*, and four variables that are related to the client-auditor engagement: (1) *AUDITOR_CHANGE*, an indicator to flag the company-year in which a new auditor was hired; (2) *SPEC_OFFICE*, an indicator of whether the auditor specializes in the client's industry at the office level, described in section 3.2.2. and based on Guo et al. (2022); (3) *TENURE*, the log of the length in years the auditor has been with the client; and (4) *AUDIT_FEES*. For more details on variable measurement, please see Appendix A. We report the estimation results in Table 7. For the explanatory variables noted with a *, we use levels in the "levels" specifications (Columns 1 and 2) and changes in the "changes" specifications (Column 3).

With respect to the two other auditor-client compatibility metrics used in the literature, *SIM_COMB* and *SPEC_OFFICE*, we find no evidence that they are correlated with *FIT* or with changes in *FIT* (with the exception of a weakly significant and negative coefficient on

SIM_COMB in column 1), *s*omewhat consistent with the univariate correlations documented in Table 6. Collectively, the evidence from Table 7 concerning *FIT* suggests that it might be capturing aspects of audit-client compatibility that are different from those tracked by auditor specialization and similarities of clients in an auditor's portfolio.

With respect to auditor characteristics, the *BIG4* and *AUDITOR_CHANGES* dummies, as well as *TENURE* are not associated with *FIT*. We find a consistent and positive association between an internal control weakness (*ICW*) and *FIT* in both "levels" specifications. In terms of general client-level characteristics, we find some evidence that *FIT* is positively associated with *SIZE*, negatively associated with the *LOSS* dummy, and is lower in companies with high levels of working capital (*ARINV*).

Finally, we examine the associations of *FIT* with its building blocks, the number of CAPs of clients (*NUM_CAPS*) and the number of auditor deficiencies (*NUM_DEF*). We find that *FIT* is lower when the number of auditor deficiencies is higher (although this is almost by construction). We also find strong evidence that an increase in auditor deficiencies leads to lower *FIT*, as evidenced by the negative coefficient in the change specification. This makes sense and validates what our measure attempts to capture. With more deficient areas attributed to an auditor, they will be less likely to be compatible with their clients.

We find a different correlation between *FIT* and *NUM_CAPS*. *FIT* is positively associated with *NUM_CAPS*. Similar results emerge when examining the changes specification. Because in computing *FIT* we attempt to control for the number of CAPs, by dividing the weakness scores by *NUM_CAPS* (see equation (2)), this positive association makes sense. To the extent, however, that *NUM_CAPS* serves as a proxy for accounting quality, we might need to control for it in the next analyses to make sure that we are able to achieve divergent validity as explained in section 3.4.

4.2 How does FIT relate to existing measures of Audit Quality?

The goal of an auditor-client matching variable, like *FIT*, is to evaluate the likelihood that an auditor can perform a high-quality audit for that specific client. If the auditor is found deficient in the areas that are most critical for the client, it is reasonable to expect that the fit between the client and the auditor is weak. In this section, we examine the ability of our auditor-client compatibility variable, *FIT*, to explain variations in audit quality across time and companies (Lines 5 to 7 in Figure 5). We also would like to evaluate how *FIT* stacks against other popular compatibility measures used in the literature and discussed in section 3.2.2. To accomplish this, we estimate the following model:

$$AQ_{i,t} = \beta_0 + \beta_1 COMPAT_{i,t} + Auditor\text{-}Characteristcis_{A,t} + Firm\text{-}Controls_{i,t} \& \ FE \qquad (5)$$

where *AQ* is proxied by three traditional audit quality measures; as discussed in section 3.3, they are: (1) *RESTATE* – an indicator for the existence of a Big R restatement in company *i*'s financial statements of fiscal year *t*, (2) Modified Jones abnormal accruals estimated using a single-step procedure following Chen at al. (2018), and (3) *MISSED_MW* – an indicator variable that equals 1 if a restatement was *not* preceded by an issuance by the auditor of a warning about material weaknesses in internal controls.

We use three variables for *COMPAT*, as described in section 3.2: (1) *FIT*, (2) *SIM_COMB* developed in Brown and Knechel (2016), and (3) *SPEC_OFFICE* following Guo et al. (2022). Our goal is to evaluate whether *FIT* is related to audit quality and how similar or different it is from other auditor-client compatibility proxies. We also include, but do not report, all the variations of *SIM_COMB*, and over twenty other auditor specialization variables (see section 3.2.2). The results with these alternative variables are similar.

To control for other auditor characteristics, we include the following variables in our regressions: a Big 4 auditor indicator (*BIG4*), the number of years the audit firm has been with the client (*TENURE*), audit fees charged by the auditor (*AUDIT_FEES*), an indicator for whether the auditor is different from previous fiscal year's auditor (*AUDITOR_CHANGE*), and an indicator for whether the auditor has identified internal control weaknesses over the company's financial reporting (*ICW* – in Panels A and B only). Finally, we use a variety of company-level control variables, including variables that are aimed to capture the accounting complexity and quality of the client (e.g., *FOREIGN*, *SEGMENTS*, and *NUM_CAPS*) to control for "non-audit" quality issues as depicted in Line 3 of Figure 5. Variable definitions are available in Appendix A.

Table 8 reports the estimation results of Equation (5), where each panel employs a different audit quality metric. Beginning with Panel A, in which the dependent variable proxying for audit quality is *RESTATE*, we find in column (1) that the coefficient on our auditor-client compatibility variable, *FIT*, is negative and statistically significant. This evidence is consistent with *FIT* capturing aspects of audit quality that reduce the likelihood of future restatements. In columns (2) and (3) we evaluate the ability of the other two common compatibility measures, *SIM_COMB* and *SPEC_OFFICE*, to explain future restatements. We find no statistical association with restatements for these two variables.

Finally, in column (4) we examine the ability of *FIT* to explain restatements when jointly competing with the other two variables. We find that *FIT* retains its negative association with restatements, even after including *SIM_COMB* and *SPEC_OFFICE* in the same model. The lack of association between *SPEC_OFFICE* and restatements is consistent with prior studies that have failed to find such an association (e.g., Cao et al., 2012; Bills et al., 2015). In unreported results (available upon request) we continue to find no association between *SPEC_OFFICE* and restatements using over twenty different proxies from nine studies (details are described in section 3.2.2), while *FIT* remains consistently negative and significant. This leads us to conclude that *FIT* better aligns with the elements of audit quality that are captured by restatements, compared to industry specialization. As for control variables, we find consistent evidence that the existence of an internal control weakness (*ICW*) is strongly associated with restatements, similar to results documented in Beardsley et al. (2021). We also find some evidence that lengthier auditor tenures are associated with lower rates of restatements, consistent with findings in Stanley et al. (2007).

We now turn to examining the second audit quality proxy. In Panel B, we use a single-step model of Total Accruals (*TACC*) to evaluate its association with the auditor-client compatibility variables. Recall that a single-step estimation procedure is the preferred way to study abnormal accruals, which is essentially what we evaluate in Panel B, per the analyses conducted by Chen et al. (2018). The results we report in Panel B omit the "modified Jones" accrual-determinant variables that we include in the regression as well as the industry- and year-fixed effects and their interactions with the accrual determinants.

In column (1), we find a negative and significant association with *FIT*, demonstrating that clients with more compatible auditors tend to report more conservatively with income-increasing accruals tending to be lower. In column (2), *SIM_COMB* has no association with abnormal accruals. In contrast, in column (3), the coefficient on *SPEC_OFFICE* is significant and *positive*, contrary to what one would expect. That is, clients of specialized auditors tend to report higher income-increasing accruals. In column (4), when all three auditor-client measures are included, we find strong evidence that *FIT* is negatively and significantly associated with abnormal accruals, while *SIM_COMB*'s and *SPEC_OFFICE*'s associations with abnormal accruals are non-existent.

The lack of association of *SPEC_OFFICE* is also observed when it is proxied by over twenty other versions used in prior studies (results available upon request).

Finally, in Panel C, we use *MISSED_MW* as a dependent variable. The results in column (1) provide strong evidence of a negative association between the likelihood of missing a material weakness in internal control and *FIT*. This relationship is not evident in Columns (2) and (3), where we use *SIM_COMB* and *SPEC_OFFICE* as measures of auditor-client compatibility. This lack of association for *SPEC_OFFICE* holds for over twenty other proxies for industry specialization (results available upon request). When we include *FIT*, together with *SIM_COMB* and *SPEC_OFFICE*, in column (4), we continue to find a negative and significant coefficient, although the statistical significance is weakened. Notably, *SPEC_OFFICE* is mildly positively associated with *MISSED_MW*, which is at odds with how one might expect a compatibility variable to behave in this model.

Overall, the results in Table 8 suggest that *FIT* captures elements of audit quality related to all three audit quality variables we examine. *FIT* appears to be the strongest among the three auditor-client compatibility variables when they are included in the same model. Furthermore, *SPEC_OFFICE* cannot explain the three audit quality metrics that we examine, either independently or jointly. In fact, in one model, its sign is flipped relative to expectations. In this regard, it is important to emphasize again that auditor's industry specialization's inability to explain our three audit quality measures is evident when using over twenty other variants of specialization, which we list in section 3.2.2.

4.2.1 Restatements Reasons

While the results in Panel A of Table 8 indicate that clients with higher-*FIT* auditors are less likely to experience future restatements, the results do not speak to the *reasons* for restatements. It is natural to ask whether restatements are more likely to occur for reasons linked to the specific areas in which the client's auditor is deficient. Because *FIT* is aggregated across all client areas and auditor's deficiencies, we are unable to answer this question with *FIT* alone.

In the next analysis we probe one layer deeper into the reasons for restatements and shed light on the relation between the reasons for the restatements and auditor-client (in)compatibility in a specific CAP area. To do so, we estimate a set of logistic models as follows:

$$RESTATE_{i,t,a} = \beta_0 + \beta_1 AREA\_MISFIT_{A,t,a,i} + Auditor\text{-}Characteristcis_{A,t} + Firm\text{-}Controls_{i,t} \;\&\; FE \tag{6}$$

We classify the reason for each restatement based on either (i) our own textual analysis of the restatement announcement against the keywords we used to define the CAP areas or (ii) using the restatement reason provided by Audit Analytics. We define $RESTATE_{i,t,a}$ to equal one if the reason for the restatement of company *i* in period *t* is deemed related to area *a*. For each area *a*, we construct the explanatory variable, $AREA\_MISFIT_{A,t,a,i}$, to be equal to one when the auditor of client *i* is deficient in the clients' critical accounting area *a*. We then estimate equation (6) separately for each CAP area (down the lines) and for each definition of restatement reason (across the columns). We report the coefficients of *AREA_MISFIT* in Panel D of Table 8.

Based on a sufficient number of restatements, we can effectively estimate 18 of the 24 models when we define the reasons for restatements based on our own algorithm (Column 1), and 10 of the 24 models when restatements reasons are defined by Audit Analytics (Column 2). In column (1), the coefficient on *AREA_MISFIT* is positive and significant in 9 out of the 18 models.

To evaluate the overall relation between the reason for restatements and the lack of compatibility between the auditor and the client, we use a binomial approach with a base probability of 10%. Nine significant coefficients out of 18 models under a null of no relation is rejected at conventional levels. In column (2) the coefficient on *AREA_MISFIT* is positive and significant in 6 out of the 10 models, which similarly reject the null of no relation at conventional levels. We conclude that not only are restatements more likely when clients and auditors are incompatible, as in Panel A, there is also evidence that the underlying reason for the restatement is related to the lack of compatibility between a client and an auditor in a specific area.

4.2.2 Sensitivity analysis – PCAOB Neglect

Our *FIT* measure has the potential to be affected by the choices that the PCAOB makes with respect to the areas of inspection and focus. Areas that the PCAOB neglects and does not inspect might not be reflected well in *FIT*. Specifically, companies with critical accounting policies in areas not inspected by the PCAOB are inclined to have artificially higher fit scores because the absence of an inspection is erroneously interpreted as suggestive of no deficiencies. This issue factors into the *FIT* ratio in equation (2) as follows. Suppose a neglected area is truly deficient for auditor a. In that case, the numerator in equation (2) should have been higher, causing the ratio overall to be higher and *FIT* to be lower (after the multiplication by -1). Since this is not observed, the actual *FIT* score might be overstated.

To address this issue, we first identified areas that the PCAOB neglected. We deemed an accounting area to be neglected if the PCAOB did not find any deficiency in this area in a particular year across all inspections and audit firms. We then recompute *FIT* to account for neglect. As discussed earlier, the numerator of *FIT* could lead to an overstatement of the *FIT* ratio. While we cannot adjust the numerator directly, because we do not observe which neglected area is actually deficient, we recompute *FIT* by excluding the neglected areas from the denominator.[12] We label the resulting adjusted *FIT* measure as *FIT_NO_NEGL,* and we report estimation results in Panel E of Table 8. These estimation results correspond to column (4) of Panels A – C of Table 8. The results are similar to those with unadjusted *FIT*. Thus, our overall conclusion is that the relation between audit quality and *FIT* is not materially affected by the coverage strategy of the PCAOB.

4.3 The role of auditor-client compatibility in oversight – Real vs. Accrual earnings management

Our previous analysis of the relationship between *FIT* and the three audit quality variables relies on the assumption that these variables reasonably capture audit quality (the assumption underlying Lines 8-10 in Figure 5). While the evidence in prior literature suggests that this is indeed the case, to varying degrees, we opt for a different approach to proxy for audit quality in this section. We also believe that using this proxy minimizes the impact of accounting quality on the proxy, furthering the achievement of divergent validity.

The literature on audit and earnings management recognizes the substitution that might exist between accrual-based and real actions that managers take to accomplish their financial reporting goals (Cohen et al., 2008; Zang, 2012). It has been argued that auditors have different capabilities in how they oversee the two types of earnings management approaches. While auditors are capable of restraining managers' accrual-based actions, they are far more limited in deterring managers from engaging in real earnings management. As a result, when audit quality is higher,

[12] In a few cases, where all CAP areas of a firm are neglected by the PCAOB, the denominator becomes zero and therefore the ratio is undefined. We drop these firms from the analysis, giving rise to slightly smaller samples in Panel E of Table 8 compared to the original analyses.

managers will experience a higher degree of external constraints on accrual-based actions, and they might resort to real earnings management actions (Chi et al., 2011). Therefore, the substitution between accrual and real earnings management could serve as an alternate proxy for audit quality.

Our next set of tests explores this logic. If the compatibility between the auditor and the client is high, we expect a higher audit quality, and a higher likelihood of substitution away from accrual-based earnings management to real earnings management. We examine the association of the three compatibility measures we used in Table 8 with several variables that are meant to proxy for real earnings management (*REM*). We estimate the following model:

$$REM_{i,t} = \beta_0 + \beta_1 COMPAT_{i,t} + Auditor\text{-}Characteristcis_{A,t} + Firm\text{-}Controls_{i,t} \,\&\, FE \qquad (8),$$

where *REM* is represented by these two variables: (1) *DISEXP* – discretionary expenditures, and (2) *PROD* – production costs. In a similar approach taken with respect to accrual measures in Table 8, we estimate the abnormal portion of *DISEXP* and *PROD* in a single step, within Equation (8) as prescribed by Chen et al. (2018). Therefore, we include in this model, but do not report, the determinants of *DISEXP* and *PROD* as "first-stage regressors." The models from which we adopted the first-stage regressors are advocated in Kothari et al. (2016) and Huang et al. (2020). We also multiply *DISEXP* by (-1) to make its sign consistent with *PROD.*

Table 9 reports the results. In Panel A, where the REM measure is *DISEXP*, we find a positive and significant coefficient on *FIT*. Because *DISEXP* is multiplied by (-1), the positive and significant coefficient suggests that for client-auditor pairs with higher compatibility, there are more attempts to increase reported earnings by reducing discretionary expenditures. In columns (2) and (3), when we use *SIM_COMB* and *SPEC_OFFICE* as the compatibility measures, we find no evidence of increased *DISEXP*, as both coefficients are not significant. When including *FIT* together with *SIM_COMB* and *SPEC_OFFICE*, in column (4), the strong and positive coefficient on *FIT* remains. This reinforces the inferences from column (1), suggesting, again, that with more stringent auditor oversight over accruals (i.e., higher *FIT*), companies tend to reduce discretionary expenditures. Interestingly, we also find a consistent negative and significant coefficient on *ICW* across all models. A finding of an internal control weakness by the auditor might suggest heightened risk exists for accrual-based earnings management during that year, as internal controls over financial reporting have been compromised. This elevated risk of accrual-based earnings management is consistent with lower levels of REM (or higher discretionary expenditures), as reflected in the negative coefficient on *ICW*.

Moving to Panel B, where the REM measure is *PROD*, we find a positive and significant coefficient in column (1) on *FIT*. This is consistent with an increase in production costs, reflecting possible attempts to manage earnings through overproduction, in the presence of strong auditor oversight over accrual-based earnings management. A positive and mildly significant coefficient is present for *SIM_COMB* in column (2), but no evidence of an association between *PROD* and *SPEC_OFFICE* is seen in column (3). When combining the auditor compatibility measures in column (4), we find consistent evidence supporting the existence of higher level of production for *FIT*. Interestingly, we find in the same specifications that *SIM_COMB* is positively associated with higher levels of production costs, reflecting higher level of auditor scrutiny over financial reporting. Similar to other specifications, we find no evidence that *SPEC_OFFICE* is associated with abnormal levels of production.

Collectively, Table 9 provides consistent evidence of higher levels of REM in companies whose auditor is more compatible with their critical accounting policies. This suggests that such auditors provide strong oversight over accrual-based earnings management methods, leading managers to substitute into REM approaches. This result is also consistent with the findings in Cohen et al. (2021) based on auditor adjustment data from China and those in Chi et al. (2011) who rely on traditional metrics of audit quality such as audit fees and industry expertise.

4.4 *Does the market take account of auditor-client compatibility?*

In the previous sections, we established that our new auditor-client compatibility measure captures audit quality attributes, based on their associations with metrics previously used to assess audit quality. We also showed that *FIT*'s ability to track audit quality is generally superior to other compatibility metrics. We now turn to evaluate how the market perceives the earnings of companies, as a function of their compatibility with their auditors. Do investors respond more strongly to earnings that have been audited by auditors that are more compatible with the client?

To do so, we evaluate earnings response coefficients (ERCs), adopting the approach taken by Baber et al. (2014), and Gipper et al. (2020), among others. To evaluate the relation between stock market response to earnings and auditor-client compatibility, we estimate the following regression model:

$$\begin{aligned} CAR_{i,t} = \beta_0 + \beta_1 UE_{i,t} + \beta_2 UE_{i,t} x TOP\_COMPAT_{i,t} + \beta_3 TOP\_COMPAT_{i,t} \\ + \beta_m Controls + FE \end{aligned} \quad (7),$$

where *CAR* represents the three-day abnormal returns (adjusted using Fama-French three factors) around an earnings announcement, *UE* is unexpected earnings measured as the actual earnings minus the median of all analysts' forecasts computed in the statistical period (in IBES) closest to the earnings announcement. Our main variables of interest are those represented by *TOP_COMPAT*. Each year, we rank all auditor-client pairs based on their *FIT,* and *SIM_COMB. TOP_FIT, and TOP_SIM_COMB* are indicator variables that equal to one if the auditor-client pair was in the top half of the ranked distribution as known by the market at the time of the earnings announcement.[13] These variables should capture whether a pairing between a client and its auditor, based on information that is available to investors as of the earnings announcement date, is stronger than the median pairing each year. The third variable, *SPEC_OFFICE*, which is based on industry specialization, is defined as in the previous sections.

We report the estimation results of Equation (7) in Table 10. Our focus in this table is on the interaction term, *UE*TOP_COMPAT*, which represents the portion of ERCs that is attributed to the stronger compatibility between the auditor and the client. A positive coefficient means that the market responds more strongly to the announcements of earnings of companies whose auditor is more compatible. For each compatibility variable, we report the results from a simple specification, without control variables, and one that includes a series of basic firm controls. In all specifications, we use a set of fixed effects that include industry, year *x* quarter, and auditor.

We first note the expected positive and significant coefficient on *UE* in all specifications. As for the compatibility interaction terms with *UE*, in columns (1) and (2), we find that the incremental stock price response to unexpected earnings issued by companies with high *FIT,* as reflected in the coefficients on *UE x TOP_FIT,* is higher than other companies, with a significance level of 5% in column (2). This result is consistent with the market perceiving these companies' earnings to be of higher quality.

In contrast, in columns (3) and (4), we find evidence that the market responds less strongly to earnings announcements of companies audited by an auditor whose clients are similar, as proxied by *SIM_COMB*. The coefficients on *UE x TOP_SIM_COMB* are negative and significant.

[13] For compatibility measures computed using PCAOB inspection reports and/or firm's 10-K we verify that the inspection reports and or the 10'K's we use for a specific earnings announcement have been issued prior to the earnings announcement date.

This means that investors do not view the earnings of firms audited by auditors with similar clients to be of higher quality. It is possible that similarity through financial statement narrative does not capture the qualities that investors expect to see in earnings. Finally, it appears that investors respond strongly to the earnings of firms audited by specialists, as reflected in the positive and significant coefficients on *UE x SPEC_OFFICE* in columns (5) and (6).

When we include all three compatibility measures together in one model, in column (7), we find that the coefficient on *UE* SPEC_OFFICE* remains positive and significant, with the coefficient on *UE x TOP_FIT* remaining positive but losing some of its statistical strength. The coefficient on *UE x TOP_SIM_COMB* remains negative but also loses some of its statistical strength.

In summary, the results in Table 10 provide strong evidence that the market affords more credibility to financial statements audited by industry specialists. The evidence shows that the market also lends some credence to the earnings of companies audited by more compatible auditors in terms of topical alignment. Coupled with our results from Table 8-9, however, it remains an open question of whether the market response is reflective of a firm's audit quality because, based on those results, *FIT* is a much stronger predictor of audit quality metrics, such as restatements, than industry specialization.

5. Conclusion

We utilize two important disclosures – the PCAOB inspection reports of an auditor and the critical accounting policies of this auditor's clients – to develop a new measure of auditor-client fit that is based on the topical compatibility between an auditor and its individual clients. If, for example, an auditor is found deficient in auditing of, say, the allowance for loan losses, then the measure will tend to deem this auditor a weak fit for clients whose allowance for loan losses is considered a critical accounting policy.

The auditor fit measure we develop is associated with existing proxies of audit quality, such as restatements, abnormal accruals, and the likelihood of an auditor missing a material weakness in internal controls. These associations demonstrate the new measure's convergence validity, that is, it captures some useful attributes of the auditor-client relationship that influence the quality of the audit. We also show that our auditor-client fit performs better than a comprehensive set of over twenty industry specialization variables, which have been the most widely used proxies for auditor-client compatibility in the literature to date. Furthermore, our tests show that the market response to earnings is stronger for companies whose auditor is of stronger compatibility with their important areas of accounting. The stock price reaction tests also reveal that the market responds more strongly to the earnings of clients audited by industry specialists.

The success of our measure in explaining audit quality is important for several reasons. It advances our understanding of the determinants of audit quality by showing that topical compatibility is an important attribute in the matching between auditors and clients. This topical compatibility goes beyond the specialization concept that has been the literature's gold standard for capturing auditor-client compatibility. Second, our findings suggest that both the PCAOB inspection reports and the disclosures of critical accounting policies provide useful information *when combined,* to the marketplace, including investors, regulators, and SEC registrants. However, our metric is not without limitations. For example, our metric relies on the proper and accurate disclosures by companies of their Critical Accounting Policies. Thus, it will be less effective when companies strategically report these CAPs to mislead and obfuscate.

Although beyond the scope of our current research, our measure has the potential to explain other aspects of the auditor-client relationship. For example, do clients take account of the fit between their own accounting exposures and their auditor's expertise? Is there evidence that the decision to hire an auditor is related to fit? A second aspect is audit fees. Are audit fees related to the degree of fit between an auditor and a potential client?

References

Abbott, L. J., Boland, C. M., McCarthy, S. M., & Swenson, L. A. (2024). The Association between PCAOB Revenue-Deficient Audit Engagements and Revenue Quality. *The Accounting Review*, 1-27.

Abbott, L. J., Gunny, K. A., & Zhang, T. C. (2013). When the PCAOB talks, who listens? Evidence from stakeholder reaction to GAAP-deficient PCAOB inspection reports of small auditors. *Auditing: A Journal of Practice & Theory, 32*(2), 1-31.

Acito, A. A., Hogan, C. E., & Mergenthaler, R. D. (2018). The effects of PCAOB inspections on auditor-client relationships. *The Accounting Review, 93*(2), 1-35.

Aobdia, D. (2018). The impact of the PCAOB individual engagement inspection process—Preliminary evidence. *The Accounting Review, 93*(4), 53-80.

Aobdia, D. (2019). Do practitioner assessments agree with academic proxies for audit quality? Evidence from PCAOB and internal inspections. *Journal of Accounting and Economics, 67*(1), 144-174.

Aobdia, D. (2020). The economic consequences of audit firms' quality control system deficiencies. *Management Science, 66*(7), 2883-2905.

Aobdia, D., Lin, C. J., & Petacchi, R. (2015). Capital market consequences of audit partner quality. *The Accounting Review, 90*(6), 2143-2176

Baber, W. R., Krishnan, J., & Zhang, Y. (2014). Investor perceptions of the earnings quality consequences of hiring an affiliated auditor. *Review of Accounting Studies*, 19, 69-102.

Bae, G. S., Choi, S. U., Dhaliwal, D. S., & Lamoreaux, P. T. (2017). Auditors and client investment efficiency. *The Accounting Review, 92*(2), 19-40.

Balsam, S., Krishnan, J., & Yang, J. S. (2003). Auditor industry specialization and earnings quality. *Auditing: A journal of practice & Theory, 22*(2), 71-97.

Barton, J., & Simko, P. J. (2002). The balance sheet as an earnings management constraint. *The Accounting Review*, 77(s-1), 1-27.

Balsam, S., Krishnan, J., & Yang, J. S. (2003). Auditor industry specialization and earnings quality. *Auditing: A Journal of Practice & Theory, 22*(2), 71-97.

Beardsley, E. L., Imdieke, A. J., & Omer, T. C. (2021). The distraction effect of non-audit services on audit quality. *Journal of Accounting and Economics, 71*(2-3), 101380.

Behn, B. K., Choi, J. H., & Kang, T. (2008). Audit quality and properties of analyst earnings forecasts. *The Accounting Review, 83*(2), 327-349.

Bills, K. L., Jeter, D. C., & Stein, S. E. (2015). Auditor industry specialization and evidence of cost efficiencies in homogenous industries. *The Accounting Review*, *90*(5), 1721-1754.

Brown, J.R., Jr. (2020). *Seeing Through the Regulatory Looking Glass: PCAOB Inspection Reports*. CFA Institute's Corporate Disclosure Policy Council and Capital Markets Policy Council, July 23. Available at: https://pcaobus.org/news-events/speeches/speech-detail/seeing-through-the-regulatory-looking-glass-pcaob-inspection-reports_724.

Brown, S. V., & Knechel, W. R. (2016). Auditor–client compatibility and audit firm selection. *Journal of Accounting Research*, *54*(3), 725-775.

Burke, J. J., Hoitash, R., & Hoitash, U. (2019). Audit partner identification and characteristics: Evidence from US Form AP filings. *Auditing: A Journal of Practice & Theory*, *38*(3), 71-94.

Cahan, S. F., Che, L., Knechel, W. R., & Svanström, T. (2022). Do Audit Teams Affect Audit Production and Quality? Evidence from Audit Teams' Industry Knowledge. *Contemporary Accounting Research* 39(4), 2657-2695.

Cao, Y., Myers, L. A., & Omer, T. C. (2012). Does company reputation matter for financial reporting quality? Evidence from restatements. *Contemporary Accounting Research*, *29*(3), 956-990.

Cassell, C., Hunt, E., Narayanamoorthy, G., & Rowe, S. P. (2019). A hidden risk of auditor industry specialization: evidence from the financial crisis. *Review of Accounting Studies*, *24*, 891-926.

Chen, W. E. I., Hribar, P., & Melessa, S. (2018). Incorrect inferences when using residuals as dependent variables. *Journal of Accounting Research*, *56*(3), 751-796.

Chi, W., Lisic, L. L., & Pevzner, M. (2011). Is enhanced audit quality associated with greater real earnings management?. *Accounting Horizons*, 25(2), 315-335.

Cohen, D. A., Dey, A., & Lys, T. Z. (2008). Real and accrual‐based earnings management in the pre‐and post‐Sarbanes‐Oxley periods. *The Accounting Review*, *83*(3), 757-787.

Cohen, D. A., Dai, Z., Han, X., Wu, L., & Zhou, K. (2021). Unintended Consequences from Downward Audit Adjustments: Evidence from China. Working Paper, Vanderbilt University. Available at SSRN 4016635.

Cunningham, L. M., Li, C., Stein, S. E., & Wright, N. S. (2019). What's in a name? Initial evidence of US audit partner identification using difference-in-differences analyses. *The Accounting Review*, *94*(5), 139-163.

Dechow, P. M., & Dichev, I. D. (2002). The quality of accruals and earnings: The role of accrual estimation errors. *The accounting review*, *77*(s-1), 35-59.

Dechow, P., Ge, W., & Schrand, C. (2010). Understanding earnings quality: A review of the proxies, their determinants and their consequences. *Journal of Accounting and Economics*, *50*(2-3), 344-401.

DeFond, M. L. (2010). How should the auditors be audited? Comparing the PCAOB inspections with the AICPA peer reviews. *Journal of Accounting and Economics, 49*(1-2), 104-108.

DeFond, M., & Zhang, J. (2014). A review of archival auditing research. *Journal of Accounting and Economics*, *58*(2-3), 275-326.

Dekeyser, S., He, X., Xiao, T., & Zuo, L. (2023). Auditor industry range and audit quality. *Journal of Accounting and Economics*, 101669.

Dopuch, N., & Simunic, D. (1982, June). Competition in auditing: An assessment. In *Fourth Symposium on auditing research* (Vol. 401, p. 405). Urbana, IL: University of Illinois.

Dunn, K. A., & Mayhew, B. W. (2004). Audit firm industry specialization and client disclosure quality. *Review of accounting studies, 9*, 35-58.

Folsom, D., Hribar, P., Mergenthaler, R. D., & Peterson, K. (2017). Principles-based standards and earnings attributes. *Management Science, 63*(8), 2592-2615.

Francis, J. R., Reichelt, K., & Wang, D. (2005). The pricing of national and city‐specific reputations for industry expertise in the US audit market. *The accounting review, 80*(1), 113-136.

Francis, J., LaFond, R., Olsson, P., & Schipper, K. (2005). The market pricing of accruals quality. Journal of accounting and economics, 39(2), 295-327.

Francis, J. R., & Yu, M. D. (2009). Big 4 office size and audit quality. *The accounting review, 84*(5), 1521-1552.

Francis, J. R. (2011). A framework for understanding and researching audit quality. *Auditing: A journal of practice & theory, 30*(2), 125-152.

Francis, J. R. (2023). What exactly do we mean by audit quality?. *Accounting in Europe,* 1-11.

Gipper, B., Leuz, C., & Maffett, M. (2020). Public oversight and reporting credibility: Evidence from the PCAOB audit inspection regime. *The Review of Financial Studies, 33*(10), 4532-4579.

Glendening, M. (2017). Critical accounting estimate disclosures and the predictive value of earnings. *Accounting Horizons, 31*(4), 1-12.

Glendening, M., Mauldin, E. G., & Shaw, K. W. (2019). Determinants and consequences of quantitative critical accounting estimate disclosures. *The Accounting Review, 94*(5), 189-218.

Gramling, A. A., Krishnan, J., & Zhang, Y. (2011). Are PCAOB-identified audit deficiencies associated with a change in reporting decisions of triennially inspected audit firms?. Auditing: *A Journal of Practice & Theory, 30*(3), 59-79.

Gul, F. A., Fung, S. Y. K., & Jaggi, B. (2009). Earnings quality: Some evidence on the role of auditor tenure and auditors' industry expertise. *Journal of Accounting and Economics*, 47(3), 265-287.

Gunny, K. A., & Zhang, T. C. (2013). PCAOB inspection reports and audit quality. *Journal of Accounting and Public Policy, 32*(2), 136-160.

Guo, Q., Koch, C., & Zhu, A. (2022). The value of auditor industry specialization: Evidence from a structural model. *The Accounting Review*, *97*(7), 193-222.

Hammersley, J. S. (2006). Pattern identification and industry - specialist auditors. *The Accounting Review, 81*(2), 309-336.

Hoberg, G., & Phillips, G. (2010). Product market synergies and competition in mergers and acquisitions: A text-based analysis. *The Review of Financial Studies*, *23*(10), 3773-3811.

Hoberg, G., & Phillips, G. (2016). Text-based network industries and endogenous product differentiation. *Journal of Political Economy*, 124(5), 1423-1465.

Huang, S., Roychowdhury, S., & Sletten, E. (2020). Does litigation deter or encourage real earnings management?. *The Accounting Review, 95*(3), 251-278.

Kothari, S. P., Leone, A. J., & Wasley, C. E. (2005). Performance matched discretionary accrual measures. *Journal of Accounting and Economics, 39*(1), 163-197.

Kothari, S. P., Mizik, N., & Roychowdhury, S. (2016). Managing for the moment: The role of earnings management via real activities versus accruals in SEO valuation. *The Accounting Review, 91*(2), 559-586.

Krishnan, G. V. (2003). Does Big 6 auditor industry expertise constrain earnings management?. *Accounting horizons*, 17, 1-16.

Kwon, S. Y., Lim, C. Y., & Tan, P. M. S. (2007). Legal systems and earnings quality: The role of auditor industry specialization. *Auditing: A Journal of Practice & Theory, 26*(2), 25-55.

Landsman, W. R., Nelson, K. K., & Rountree, B. R. (2009). Auditor switches in the pre-and post-Enron eras: Risk or realignment?. *The Accounting Review, 84*(2), 531-558.

Lennox, C., & Pittman, J. (2010). Auditing the auditors: Evidence on the recent reforms to the external monitoring of audit firms. *Journal of Accounting and Economics, 49*(1-2), 84-103.

Levine, C. B., & Smith, M. J. (2011). Critical accounting policy disclosures. *Journal of Accounting, Auditing & Finance, 26*(1), 39-76.

Lim, C. Y., & Tan, H. T. (2008). Non‐audit service fees and audit quality: The impact of auditor specialization. *Journal of accounting research, 46*(1), 199-246.

Low, K. Y. (2004). The effects of industry specialization on audit risk assessments and audit‐planning decisions. *The accounting review, 79*(1), 201-219.

McKenna, F., Pevzner, M., Sheneman, A., & Zach, T. (2022a). Corruption in the Auditor Inspection Process: The Case of KPMG and the PCAOB. *Issues in Accounting Education*.

McKenna, F., Pevzner, M., Sheneman, A., & Zach, T. (2022b). Deconstructing the PCAOB: Using organizational economics to assess the state of a regulator. Available at SSRN 4227295.

Minutti-Meza, M. (2013). Does auditor industry specialization improve audit quality?. *Journal of Accounting Research, 51*(4), 779-817.

Nallareddy, S., Sethuraman, M., & Venkatachalam, M. (2020). Changes in accrual properties and operating environment: Implications for cash flow predictability. *Journal of Accounting and Economics, 69*(2-3), 101313.

Neal, T. L., & Riley Jr, R. R. (2004). Auditor industry specialist research design. *Auditing: A journal of practice & Theory, 23*(2), 169-177.

Palmrose, Z. V. (2013). PCAOB audit regulation a decade after SOX: Where it stands and what the future holds. *Accounting Horizons*, 27(4), 775-798.

Payne, J. L. (2008). The influence of audit firm specialization on analysts' forecast errors. *Auditing: A Journal of Practice & Theory, 27*(2), 109-136.

Peecher, M. & I. Solomon. 2014. The PCAOB's 'Audit Failure' Rate Is Highly Suspect. *CFO.com*. Available at: https://www.cfo.com/auditing/2014/02/pcaobs-audit-quality-highly-suspect/

Pittman, J., Stein, S. E., & Valentine, D. F. (2023). The importance of audit partners' risk tolerance to audit quality. *Contemporary Accounting Research*, *40*(4), 2512-2546.

Reichelt, K. J., & Wang, D. (2010). National and office‐specific measures of auditor industry expertise and effects on audit quality. *Journal of Accounting Research, 48*(3), 647-686.

Rice, S. C., & Weber, D. P. (2012). How effective is internal control reporting under SOX 404? Determinants of the (non-) disclosure of existing material weaknesses. *Journal of Accounting Research*, *50*(3), 811-843.

Schroeder, J. H., & Hogan, C. E. (2013). The impact of PCAOB AS5 and the economic recession on client portfolio characteristics of the Big 4 audit firms. *Auditing: A Journal of Practice & Theory, 32*(4), 95-127.

Securities and Exchange Commission (2002). Proposed Rule: Disclosure in Management's Discussion and Analysis about the Application of Critical Accounting Policies. (https://www.sec.gov/rules/proposed/33-8098.htm)

Shu, S. Z. (2000). Auditor resignations: Clientele effects and legal liability. *Journal of Accounting and Economics, 29*(2), 173-205.

Solomon, I., Shields, M. D., & Whittington, O. R. (1999). What do industry-specialist auditors know?. *Journal of accounting research, 37*(1), 191-208.

Stanley, J. D., & DeZoort, F. T. (2007). Audit firm tenure and financial restatements: An analysis of industry specialization and fee effects. *Journal of Accounting and Public Policy, 26*(2), 131-159.

Szerwo, B. (2023). MD&A disclosure of critical accounting policies and financial reporting risk: evidence from restatements. *Journal of Accounting, Auditing & Finance, 38*(1), 104-129.

Thompson, R. A. (2023). Reporting misstatements as revisions: An evaluation of managers' use of materiality discretion. *Contemporary Accounting Research, 40*(4), 2745-2784.

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

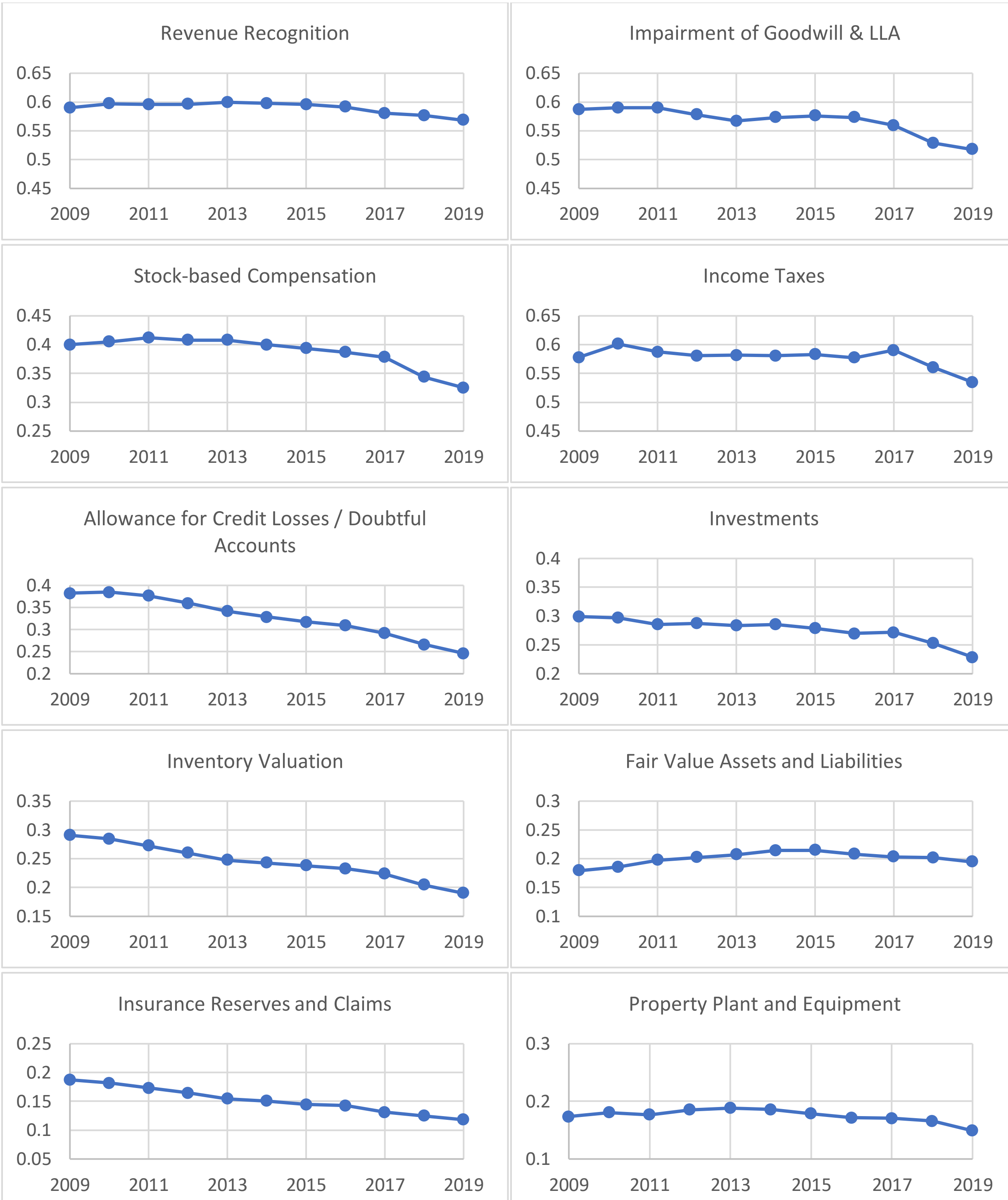
APPENDIX A: Variable Definitions

FIT	Our measure of auditor-client fit, calculated based on the client's reported Critical Accounting Policies (CAPs) and its auditor's PCAOB inspection deficiencies. For each company-year in our sample, we extract the CAP sections from their 10-Ks. First, using these disclosures we construct a list and dictionary of the most popular 24 accounting areas reported by companies as critical. Then we code each 10-K as to whether they report a given area as CAP. Next, we extract the deficient areas of Big Four auditors from their annual PCAOB inspections and count the number of engagements they were found deficient for each area, using the same dictionary. We calculate a $Weakness$ score for each auditor-year as the number of deficient engagements in the respective area, divided by the total number of engagements reviewed for the inspection. We finally combine the client's CAPs and the auditor's weaknesses with the following formula. It scales the measure by the total number of CAPs reported and multiplies it with -1 to make it decrease in the extent of overlapping areas, and increase in the lack of overlaps. $FIT_{i,A,t} = \frac{\sum_{a=1}^{24}(CAP_{i,t,a} \times Weakness_{A,t,a})}{\sum_{a=1}^{24} CAP_{i,t,a}} \times -1$
FIT_NO_NEGL	A version of the *FIT* measure, adjusted for areas possibly neglected by the PCAOB. The formula used for calculating *FIT* applies to *FIT_NO_NEGL* as well. However, we exclude the neglected CAP areas from the original set of 24 areas. We consider an area neglected if in a given year, the PCAOB does not issue any deficiencies for it across all nine auditors in our sample. Accordingly, the summations in both parts of the ratio are done across a number of CAPs that changes every year. Since the absence of deficiencies in an area does not change the numerator, this adjustment is equivalent to excluding the areas from the denominator, thereby increasing the ratio, and decreasing the *FIT* measure correcting it for a possible overstatement.
AREA_MISFIT	For each of the 24 areas, *AREA_MISFIT* is a dummy variable taking the value of one if (1) the area is among the CAPs of the client, and (2) the area is among the deficient areas of the auditor.
NUM_CAPS	The total number of Critical Accounting Policies (CAPs) reported by the client firm in its 10-K.
SIM_COMB	Composite similarity score calculated following Brown and Knechel (2016) to proxy auditor-client compatibility, based on the similarity of a client's auditor-reviewed narrative disclosures to that of other clients of its auditor in the same industry and year. To calculate this score, we first identify the reference groups (i.e. same GICS industry-year-auditor groups) for each observation in our sample. Then we extract their business descriptions, MD&As and notes to the financial statement from their 10-Ks, and calculate their pairwise similarity scores with the same disclosure of each of the other companies in same reference group using cosine similarity algorithm. We take the average of the pairwise scores for each reference group, resulting in three scores per company-year ranging between 0 and 1, one for each disclosure type: *RAWSIM_BUS*, *RAWSIM_MDA* and *RAWSIM_NOTES*. To account for the positive correlation of length of the documents with the similarity scores, we follow Brown and Tucker (2011) and regress the raw similarity scores on the first three powers of the document length measured in number of words. The residuals of the regression (namely *SIM_BUS*, *SIM_MDA* and *SIM_NOTES*) represent the variation in the raw similarity scores that cannot be explained by document length. In addition, Brown and Knechel (2016) develop the composite score, *SIM_COMB*, by converting each score to quintile rankings, and summing them across the three scores. For purposes of brevity, in our reported analyses we only include this composite score that ranges between 3 and 15 where higher levels imply better compatibility. However, we test that our results are not affected when the three component scores are used separately.

SPEC_OFFICE	Office industry specialist indicator, taking the value of 1 if the audit firm has a market share of more than 30 percent in terms of audit fees within the two-digit SIC industry at the local level (MSA), and 0 otherwise, following Guo et al. (2022).
SPEC_NATIONAL	National industry specialist indicator, taking the value of 1 if the audit firm has a market share of more than 30 percent in terms of audit fees within the two-digit SIC industry at the national level, and 0 otherwise, following Guo et al. (2022).
MISSED_MW	Missed material weaknesses, denoted as an indicator for observations where the client did not disclose a material weakness in a given year, and eventually restated that year's financial statement.
AUDITOR_CHANGE	An indicator for whether the company's auditor for the fiscal year is different from its auditor in the previous year.
TENURE	Natural logarithm of the number of years the auditor served the client.
RESTATE	An indicator for whether the fiscal year's financial statement was later restated with an 8-k Item 4.02 filing (commonly referred to as Big R restatements).
AUDIT_FEES	Natural logarithm of audit fees.
TACC	Total accruals calculated as the income before extraordinary items minus operating cash flow from continuing operations, scaled by total assets.
WCA	Working capital accruals of the client, computed as net income before extraordinary items, plus depreciation and amortization, minus cash flow from operations.
PROD	Level of production calculated as the sum of COGS and change in inventory during the year scaled by lagged total assets, following Huang et al. (2020),
DISEXP	Discretionary expenses calculated as the sum of advertising expenses, R&D expenses and SG&A expenses, all scaled by lagged total assets, multiplied by -1, following Huang et al. (2020) and Kothari et al. (2016).
SIZE	Natural logarithm of client company's assets.
ROA	Return on assets of the client, calculated as income before extraordinary items divided by average total assets.
LEVERAGE	Computed as total liabilities divided by total assets.
LOSS	An indicator for company-years in which the client firm reported a net loss.
BM	Book to Market ratio, calculated as total shareholder's equity divided by the absolute value of price per share times common shares outstanding.
OCF	Operating Cash Flows divided by lagged total assets
ARINV	Ratio of non-cash current assets, calculated as account receivables plus inventories divided by lagged total assets
FOREIGN	An indicator for companies that have operations in foreign countries.
FYE_DEC	Indicator for client companies that have a fiscal year-end in December, accounting for the busy season of auditors.
SALESVOL	Sales volatility calculated as the standard deviation of sales revenue using a rolling window and requiring three years of data, as per Francis et al. (2009).
CASHVOL	Cash volatility calculated as the standard deviation of Cash Flow from Operations, using a rolling window and requiring three years of data, as per Francis et al. (2009).
ICW	An indicator for whether the auditor has identified any internal control weaknesses over client's financial reporting.
LIQUIDITY	Computed as current assets divided by current liabilities.
NOA	Net Operating Assets relative to sales as per Barton et al (2002).
ANALYST	Number of analysts following the company.
ZSCORE	Altman Z-Score computed as 1.2*(working capital / total assets) + 1.4*(retained earnings / total assets) + 3.3*(earnings before interest and tax / total assets) + 0.6*(market value of equity / total liabilities) + 1.0*(sales / total assets).
BIG4	Indicator for company-years where the auditor is a Big4 auditor.

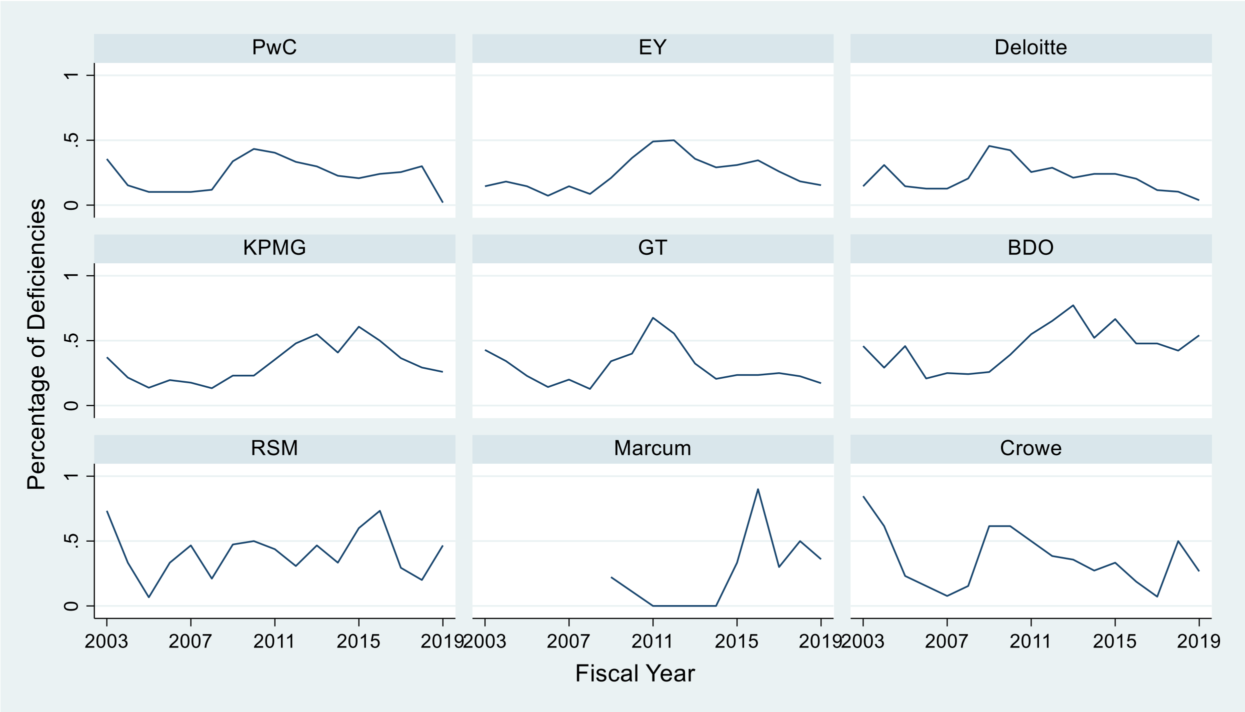
SEGMENTS	Number of business segments the client operates in.
CAR	A company's 3-day cumulative abnormal stock returns around the annual earnings announcement date, using the Fama-French three-factor model. Factors were obtained from Ken French's data library.
UE	I/B/E/S unexpected earnings, calculated as the actual quarterly EPS minus the mean of analysts' most recent forecasts of quarterly EPS prior to the earnings announcement of the client.
BETA	The market beta coefficient from regressing excess daily returns for a company on excess market returns over one fiscal quarter.
TOP_FIT	An indicator for client-auditor pairs that had above median *FIT* in the fiscal year prior to the earnings announcement. For each pair, the *FIT* score that we compare to the median score is based on the auditor's most recent inspection report publicly available before the date of earnings announcement.
TOP_SIM_COMB	*TOP_SIM_COMB*, calculated as an indicator for client-auditor pairs that had above median *TOP_SIM_COMB* in the fiscal year prior to the earnings announcement.

Figure 1: Frequency Trend of 10 Selected CAP Areas over the Sample Period



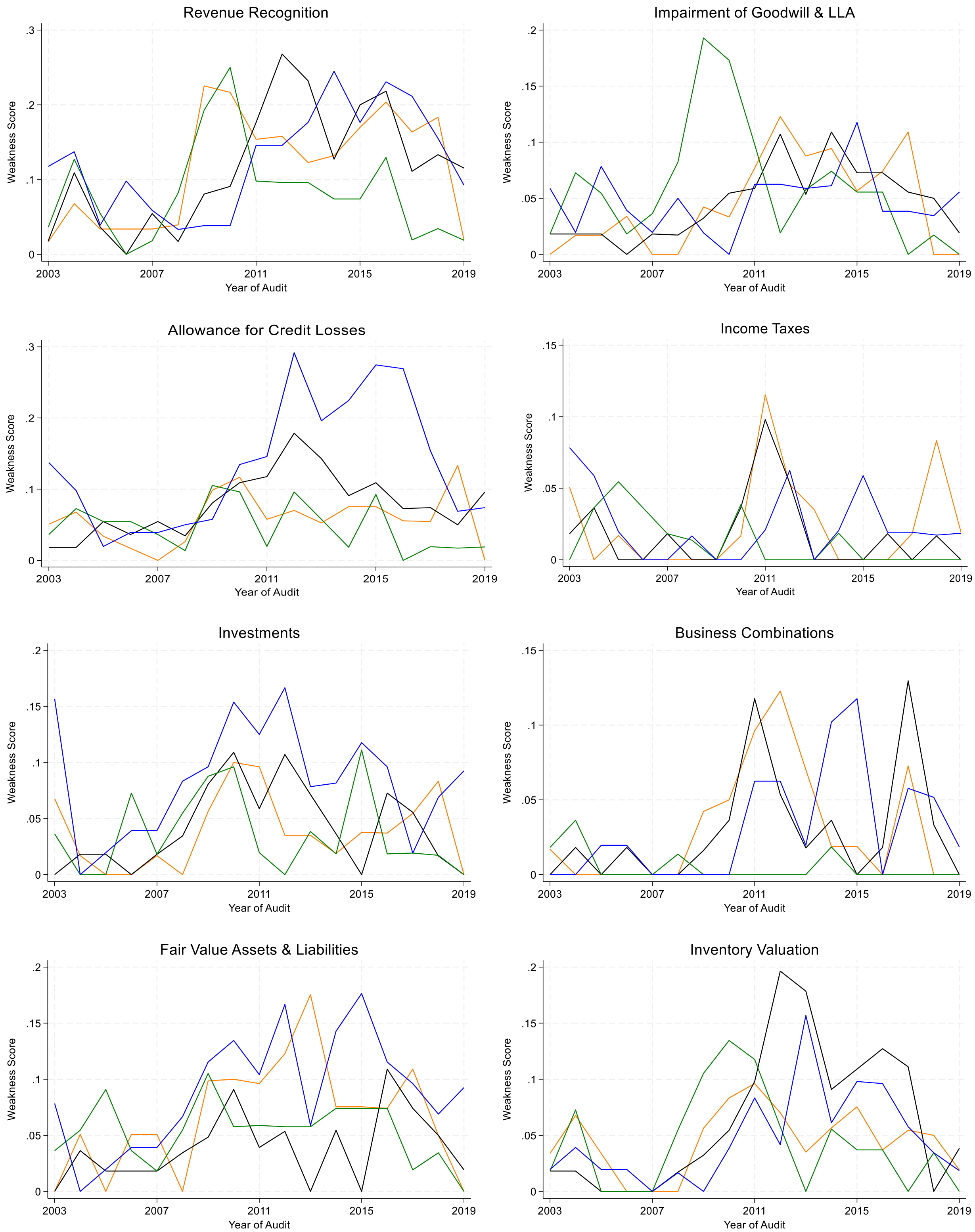
The graphs plot the occurrence frequency trends of the ten most popular CAP areas over the sample period. We identify whether the CAP section of a 10-K includes the respective area by using a dictionary of areas based on the CAP sub-headers we extracted. The dictionary helps us to identify the mentions of a CAP area even if the company does not use area sub-headers in its CAP section, and to account for the usage of different words relating to the same CAP area (e.g., inventory and merchandise).

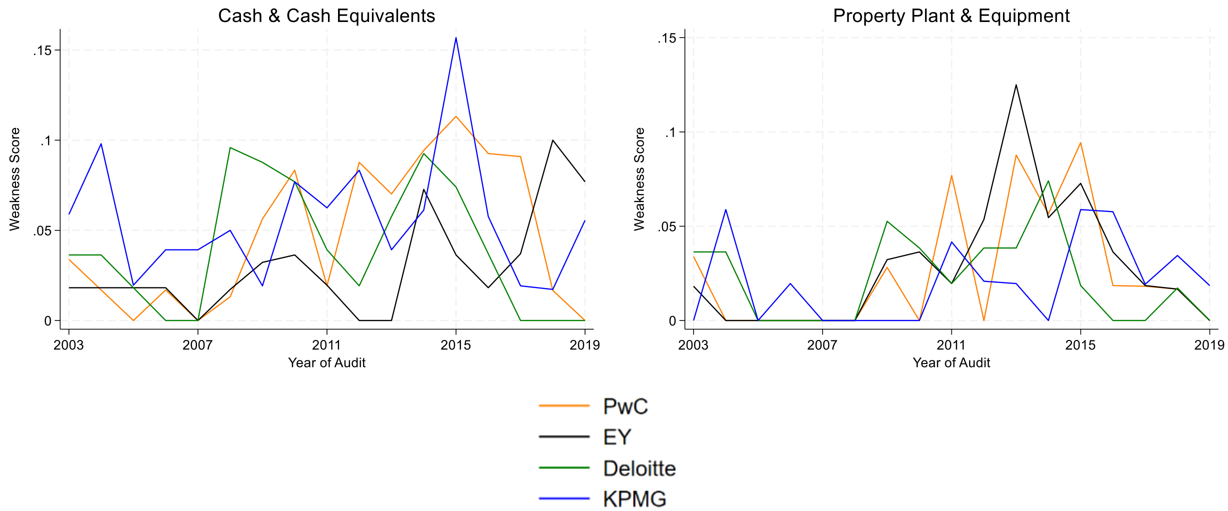
Figure 2: Number of Deficiencies over Time



The graphs illustrate the trend of auditors' percentage of engagements inspected by the PCAOB with at least one deficiency, from the beginning of the inspection regime (2003) until 2019.

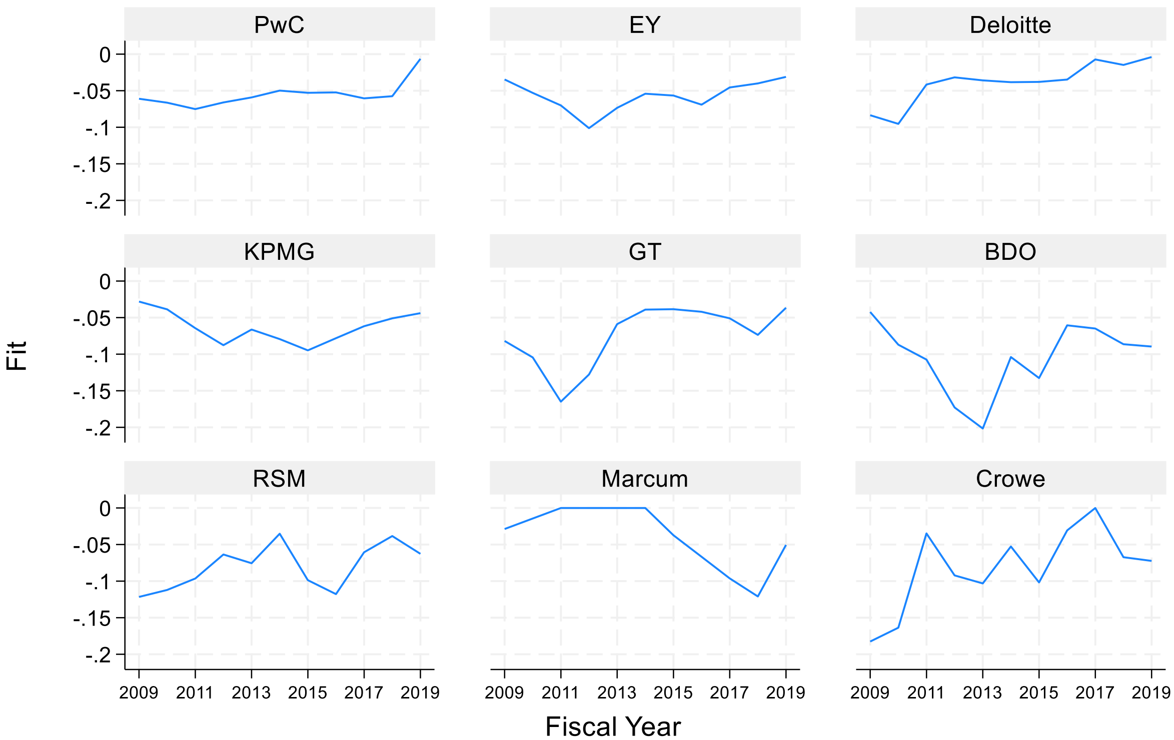
Figure 3: Ten Most Frequent Weakness Areas by Audit Firm





Each line in the tables represents the weakness score of an audit firm in the respective area. The weakness scores are calculated as the number of deficient engagements in the respective area, divided by the total number of engagements reviewed for the inspection. The years in the horizontal axis are the fiscal years of audit engagement where deficiencies occur. Weakess scores of non-Big4 auditors are excluded for visibility reasons.

Figure 4: FIT Trends



Each graph represents the variation of *FIT* over the sample period for each auditor. *FIT* is calculated as follows and is an inverse measure where 0 indicates the best possible fit and the fit degrades with more negative values.

$$FIT_{i,A,t} = \frac{\sum_{a=1}^{24}(CAP_{i,t,a} \times Weakness_{A,t,a})}{\sum_{a=1}^{24} CAP_{i,t,a}} \times -1$$

Figure 5: Research Design Framework

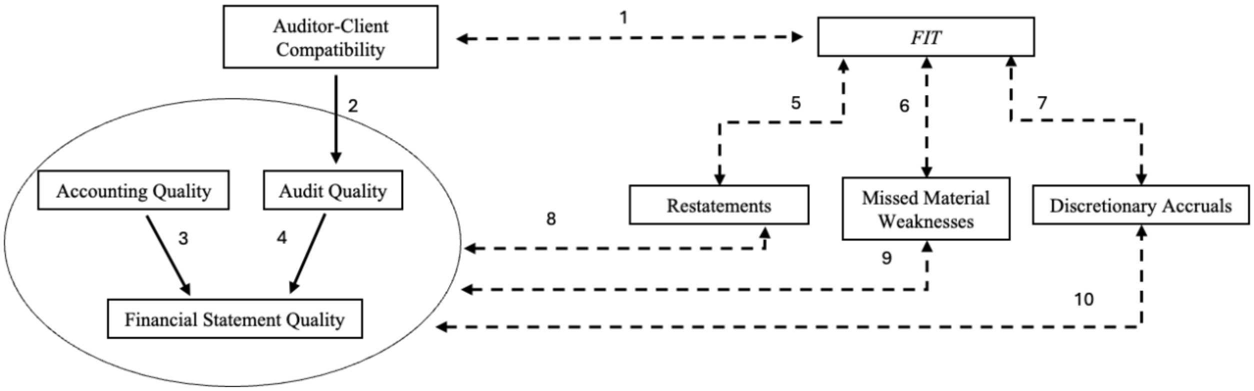


Table 1: Sample Selection

Elimination Steps	**Company-Years**	**Distinct Companies**
All CAP disclosures from 2009 to 2019 available in Calcbench	55,720	8,622
less CAP disclosures with none of the 24 determined areas	48,526	7,767
less clients of small (not annually inspected) auditors	46,848	7,075
less missing Compustat and Audit Analytics data needed for control variables	39,127	6,570
less missing data required for audit quality proxies	28,435	4,354
Observations available for SPEC_OFFICE:		
less city-industry-fiscal year combinations with less than 2 observations	24,363	3,732
Observations available for FIT:		
less missing data on PCAOB inspections deficiencies	21,694	3,568
Observations available for SIM_COMB:		
less clients of audit firms with small within-industry portfolios	13,326	2,372

Table 2: CAP Areas' Occurrence Rates- Industry Breakdown

Area	Description	NoDur	Durbl	Manuf	Enrgy	Chems	BusEq	Telcm	Utils	Shops	Hlth	Money	Other	All
1	Allowance for Credits	0.33	0.41	0.32	0.11	0.25	0.35	0.37	0.12	0.26	0.22	**0.51**	0.28	0.33
2	Asset Retirement Obligations	0.01	0.01	0.02	0.36	0.07	0.00	0.04	0.20	0.02	0.00	0.00	0.05	0.04
3	Business Combinations	0.08	0.07	0.10	0.13	0.12	0.19	0.10	0.07	0.10	0.08	0.11	0.12	0.12
4	Cash & Cash Equivalents	0.07	0.04	0.05	0.08	0.09	0.06	0.06	0.09	0.06	0.06	0.09	0.07	0.07
5	Discontinued Operations	0.01	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.02	0.02	0.01
6	Discount Rates	0.02	0.03	0.06	0.02	0.03	0.01	0.03	0.12	0.03	0.01	0.01	0.02	0.02
7	Environmental Costs	0.04	0.06	0.10	0.12	0.23	0.01	0.00	0.14	0.04	0.01	0.01	0.04	0.04
8	Fair Value Assets & Liab	0.15	0.14	0.13	0.20	0.15	0.18	0.20	0.23	0.13	0.20	0.32	0.17	0.20
9	FX Gains & Losses	0.08	0.04	0.06	0.05	0.08	0.05	0.05	0.02	0.03	0.04	0.01	0.04	0.04
10	Impairment of Goodwill & LLA	**0.74**	**0.65**	**0.72**	0.44	**0.68**	0.68	**0.77**	0.49	**0.74**	0.42	0.39	**0.62**	0.57
11	Income Taxes	**0.64**	**0.70**	**0.64**	**0.54**	**0.61**	**0.68**	0.65	0.48	**0.64**	0.48	0.51	0.57	0.58
12	Insurance Reserves & Claims	0.14	0.41	0.27	0.04	0.06	0.16	0.08	0.02	0.30	0.07	0.11	0.18	0.15
13	Inventory Valuation	0.47	0.51	0.50	0.08	0.34	0.34	0.07	0.01	0.53	0.29	0.02	0.10	0.24
14	Investments	0.18	0.13	0.15	**0.44**	0.21	0.22	0.21	0.33	0.12	0.19	**0.52**	0.24	0.28
15	Leases	0.04	0.06	0.04	0.07	0.03	0.04	0.09	0.06	0.13	0.04	0.11	0.08	0.07
16	Litigation & Other Cont.	0.12	0.13	0.18	0.16	0.31	0.18	0.24	0.26	0.15	0.11	0.09	0.17	0.15
17	Non-GAAP Financial Measures	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.00	0.01	0.00	0.03	0.01	0.01
18	Noncontrolling Interests	0.01	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.01	0.00	0.02	0.02	0.01
19	Pensions & Benefits	0.25	0.30	0.34	0.10	0.38	0.07	0.18	**0.56**	0.09	0.03	0.06	0.10	0.13
20	Property Plant & Equipment	0.17	0.08	0.13	0.26	0.24	0.09	0.24	0.19	0.16	0.07	0.27	0.27	0.18
21	R&D Costs	0.06	0.11	0.05	0.06	0.07	0.18	0.07	0.02	0.03	0.37	0.08	0.10	0.13
22	Revenue Recognition	0.59	0.63	0.62	0.43	0.48	**0.85**	**0.65**	**0.58**	0.50	**0.72**	0.32	**0.62**	0.59
23	Stock-based Compensation	0.35	0.33	0.31	0.34	0.31	0.58	0.30	0.10	0.34	**0.67**	0.17	0.36	0.39
24	Mortgage Valuation	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.11	0.01	0.02

We extract the sub-headers (i.e., the area titles) within the CAP disclosures. Using the pool of extracted CAP sub-headers, we identify 24 distinct accounting areas that appear as CAPs. Using the pool of area-relevant words under the CAP sub-headers as our reference dictionary for each area, we code each 10-K CAP section as to whether it reports the respective area as CAP or not. The figures in the table above are the fractions of 10-Ks that reports the respective area as CAP in a given industry. The last column reports the same fraction for all the 10-Ks collected. The bolded cells indicate the two most reported CAP areas for each industry. The shaded cells report the two industries that report the respective CAP area the most. Industry classifications are based on Fama-French 12 classification scheme.

Table 3: Number of Engagements Inspected by the PCAOB

Auditor	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	Total	%
PwC	71	60	52	57	57	53	53	54	55	60	52	**624**	**18.0%**
EY	62	55	51	56	56	55	55	55	54	60	52	**611**	**17.6%**
Deloitte	57	52	51	52	52	54	54	54	52	58	53	**589**	**16.9%**
KPMG	52	52	48	48	51	49	51	52	52	58	54	**567**	**16.3%**
GT	41	35	34	36	34	68	34	34	32	31	29	**408**	**11.7%**
BDO	31	23	20	23	22	23	24	23	23	26	24	**262**	**7.5%**
RSM	19	16	16	13	15	15	15	15	17	15	15	**171**	**4.9%**
Crowe	13	13	12	13	14	11	15	16	14	14	15	**150**	**4.3%**
Marcum	9	0	11	0	0	8	9	10	10	12	25	**94**	**2.7%**
Total	**355**	**306**	**295**	**298**	**301**	**336**	**310**	**313**	**309**	**334**	**319**	**3,476**	**100%**

The table reports the number of engagements (i.e. issuer audits) found deficient by the PCAOB in each engagement throughout our sample period. Note that the year is the auditor's engagement year and the counts are of audits, not deficiencies. There can be multiple deficiencies identified for a single audit. The percentages in the last column represent the proportion of deficient audits the respective audit firm has out of all the deficient audits in our sample.

Table 4: Inspection Deficiency Areas

Area	Description	N (Reports)	Num_Def	Mean	SD	Median
1	Allowance for Credit Losses / Doubtful Accounts	144	518	3.6	3.01	3
2	Asset Retirement Obligations	144	6	0.04	0.2	0
3	Business Combinations	144	153	1.06	1.6	0
4	Cash and Cash Equivalents	144	233	1.62	1.79	1
5	Discontinued Operations	144	18	0.12	0.33	0
6	Discount Rates	144	23	0.16	0.45	0
7	Environmental Costs	144	38	0.26	0.54	0
8	Fair Value Assets and Liabilities	144	339	2.35	2.28	2
9	Foreign Currency Gains and Losses	144	4	0.03	0.16	0
10	Impairment of Goodwill and Other Long-Lived Assets	144	284	1.97	2.08	1
11	Income Taxes	144	126	0.88	1.19	0
12	Insurance Reserves and Claims	144	55	0.38	0.63	0
13	Inventory Valuation	144	273	1.9	2.25	1
14	Investments	144	275	1.91	2.01	1
15	Leases	144	34	0.24	0.49	0
16	Litigation and Other Contingencies	144	80	0.56	0.8	0
17	Non-GAAP Financial Measures	144	0	0	0	0
18	Noncontrolling Interests	144	0	0	0	0
19	Pensions and Postretirement Benefit Obligations	144	19	0.13	0.43	0
20	Property Plant and Equipment	144	142	0.99	1.54	0
21	Research and Development Costs	144	52	0.36	0.78	0
22	Revenue Recognition	144	669	4.65	3.81	4
23	Stock-based Compensation	144	77	0.53	0.69	0
24	Valuation of Mortgage Servicing Rights	144	39	0.27	0.53	0

For every annual inspection report of audit firm A, we count the number of engagements that were found deficient in area a, and label the count as $Num\_def_{A,t,a}$ where t is the auditor's engagement year. Num_def reports the total number of deficiencies per area across all 144 inspection reports of all nine auditors from inspection year 2004 to inspection year 2020. $Mean$ is the average number of deficiencies per inspection report (Num_def divided by 144). *SD* and *Median* report the standard deviation and median deficiencies across inspection reports respectively.

Table 5: Descriptive Statistics

	N	Mean	SD	Min	p25	p50	p75	Max
Auditor-Client Compatibility Measures								
FIT	21,694	-0.06	0.04	-0.29	-0.08	-0.05	-0.03	0.00
SIM_COMB	13,326	8.79	3.43	3	6	8	11	15
SPEC_OFFICE	24,363	0.45	0.50	0	0	0	1	1
SPEC_NATIONAL	32,846	0.20	0.40	0	0	0	0	1
FIT_NO_NEGL	21,567	-0.06	0.04	-0.23	-0.08	-0.06	-0.03	0.00
Audit Quality Variables								
RESTATE	26,210	0.014	0.116	0	0	0	0	1
TACC	24,542	-0.19	0.70	-6.19	-0.14	-0.07	-0.03	0.45
MISSED_MW	26,210	0.011	0.106	0	0	0	0	1
REM Measures								
DISEXP	20,451	-0.41	0.46	-8.97	-0.53	-0.3	-0.14	0
PROD	21,295	0.68	0.7	0	0.22	0.47	0.87	4.74
Other Variables								
NUM_CAPS	26,210	4.07	2.44	0.00	2.00	4.00	6.00	15.00
AUDITOR_CHANGE	26,210	0.04	0.21	0.00	0.00	0.00	0.00	1.00
TENURE	26,210	11.71	13.65	0.00	4.00	8.00	14.00	86.00
AUDIT_FEES ('000)	26,210	2,048	3,525	6	309	905	2,152	28,400
SIZE (mil)	26,210	4,343	16,184	0	64	393	1,991	273,869
ROA	26,210	-0.27	1.15	-9.52	-0.15	0.01	0.06	0.33
LEVERAGE	26,210	0.35	0.79	0.00	0.02	0.19	0.38	6.75
LOSS	26,210	0.45	0.50	0.00	0.00	0.00	1.00	1.00
BM	26,210	0.38	1.08	-6.42	0.16	0.37	0.69	4.65
OCF	26,210	-0.07	0.62	-5.80	-0.03	0.07	0.13	0.51
ARINV	26,210	0.24	0.21	0.00	0.08	0.20	0.35	1.21
FOREIGN	26,210	0.53	0.50	0.00	0.00	1.00	1.00	1.00
FYE_DEC	26,210	0.72	0.45	0.00	0.00	1.00	1.00	1.00
LIQUIDITY	26,210	2.93	3.26	0.01	1.23	2.01	3.36	25.39
NOA	26,210	0.91	4.11	-13.66	0.22	0.52	1.04	30.75
ANALYST	26,210	6.16	7.42	0.00	0.00	4.00	9.00	48.00
ZSCORE	26,210	-1.04	35.50	-640.62	0.81	2.55	4.80	108.48
BIG4	26,210	0.63	0.48	0.00	0.00	1.00	1.00	1.00
SALESVOL	26,210	0.26	0.48	0.00	0.05	0.12	0.26	3.86
CASHVOL	26,210	0.20	0.72	0.00	0.02	0.05	0.11	6.79
ICW	26,210	0.04	0.20	0.00	0.00	0.00	0.00	1.00
SEGMENTS	26,210	2.31	1.82	1.00	1.00	1.00	3.00	15.00
ERC Variables								
CAR	84,029	-0.004	0.090	-2.030	-0.038	-0.001	0.035	2.342
UE	84,029	-0.007	0.168	-1.509	-0.001	0.001	0.003	0.785
BETA	84,029	0.011	0.005	-0.001	0.008	0.011	0.014	0.028
TOP_FIT	84,029	0.507	0.500	0.000	0.000	1.000	1.000	1.000
TOP_SIM_COMB	46,559	0.398	0.489	0.000	0.000	0.000	1.000	1.000

All variables are defined in Appendix A. The continuous variables are winsorized at 99% level. *SIZE*, *AUDIT_FEES*, *ANALYST* and *TENURE* variables are described in their original levels in this table, while they are used in their logged forms in the regressions.

Table 6: Correlations

	[1]	*[2]*	*[3]*	*[4]*	*[5]*	*[6]*	*[7]*	*[8]*	*[9]*	*[10]*	*[11]*	*[12]*	*[13]*	*[14]*
FIT [1]	1													
SIM_COMB [2]	-0.14***	1												
SPEC_OFFICE [3]	0.11***	-0.01	1											
NUM_CAPS [4]	0.09***	0.05***	0.02***	1										
AUDITOR_CHANGE [5]	-0.03***	-0.02*	-0.09***	0.02**	1									
TENURE [6]	0.11***	-0.01	0.19***	-0.03***	-0.17***	1								
RESTATE [7]	-0.03***	0.02***	-0.01	0	0.02***	-0.03***	1							
AUDIT_FEES [8]	0.14***	-0.04***	0.23***	0.05***	-0.08***	0.33***	-0.01*	1						
TACC [9]	0.01	0.01	0.07***	0.03***	-0.06***	0.08***	0	0.07***	1					
MISSED_MW [10]	-0.02***	0.02**	-0.01	0	0.02***	-0.03***	0.90***	-0.01**	0.01	1				
PROD [11]	-0.01	0.04***	0.04***	-0.05***	0.01	0	0	-0.06***	0.02**	0.01	1			
DISEXP [12]	0.08***	-0.01	0.11***	0.04***	-0.04***	0.12***	0	0.18***	0.34***	-0.01	-0.03***	1		
SIZE [13]	0.10***	-0.03***	0.15***	0	-0.05***	0.27***	-0.01*	0.74***	0.04***	-0.01	-0.07***	0.13***	1	
ROA [14]	0.03***	0.01*	0.06***	0.05***	-0.07***	0.11***	-0.01	0.10***	0.70***	-0.01	0	0.48***	0.07***	1
LEVERAGE [15]	0.07***	-0.10***	0.05***	-0.03***	0.01**	0.02***	0.02***	0.08***	-0.17***	0.01*	-0.04***	0.06***	0.05***	-0.17***

The table reports the pairwise correlations for selected variables. *, **, *** Denote statistical significance levels at 10 percent, 5 percent, and 1 percent, respectively. All variables are defined in Appendix A.

Table 7: Determinants of FIT

	FIT (1)	*FIT (t+1)* (2)	*ΔFIT* (3)
*SIM_COMB**	-0.000*	-0.000	-0.000
	(0.098)	(0.590)	(0.241)
*SPEC**	-0.000	0.001	0.001
	(0.729)	(0.339)	(0.544)
TENURE	0.000	-0.000	-0.001
	(0.980)	(0.926)	(0.266)
*AUDIT_FEES**	-0.001	-0.000	-0.002
	(0.460)	(0.873)	(0.260)
*ICW**	0.046***	0.032***	-0.002
	(0.000)	(0.000)	(0.218)
BIG4	-0.001	0.002	-0.001
	(0.599)	(0.408)	(0.102)
AUDITOR_CHANGE	-0.000	-0.005*	-0.005
	(0.842)	(0.071)	(0.166)
*SIZE**	0.002***	0.003***	-0.001
	(0.000)	(0.000)	(0.554)
*ROA**	-0.021*	-0.014	-0.019
	(0.069)	(0.319)	(0.165)
*LEVERAGE**	0.000	0.002	0.002
	(0.924)	(0.585)	(0.673)
LOSS	-0.003**	-0.003*	0.000
	(0.015)	(0.054)	(0.645)
*BM**	-0.002**	-0.003**	-0.001
	(0.015)	(0.011)	(0.236)
*OCF**	0.002	0.001	0.003
	(0.471)	(0.672)	(0.560)
*ARINV**	-0.013***	-0.008**	-0.005
	(0.000)	(0.024)	(0.249)
FOREIGN	-0.000	-0.000	-0.000
	(0.945)	(0.799)	(0.764)
FYE_DEC	0.001	0.002	0.000
	(0.279)	(0.225)	(0.922)
*LIQUIDITY**	0.001**	0.000	0.000
	(0.015)	(0.124)	(0.112)
*lag_NOA**	-0.000	0.000	0.000
	(0.837)	(0.573)	(0.692)
*ln_ANALYST**	-0.001	-0.001	0.001
	(0.280)	(0.102)	(0.514)
*ln_ZSCORE**	-0.001*	-0.001*	0.000
	(0.062)	(0.084)	(0.680)
*SALESVOL**	-0.004***	-0.005***	-0.000
	(0.004)	(0.002)	(0.782)
*CASHVOL**	0.004***	0.006***	-0.000
	(0.001)	(0.001)	(0.998)
*NUM_CAPS**	0.003***	0.002***	0.004***
	(0.000)	(0.000)	(0.000)
*NUM_DEF**	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)
Industry FE	Yes	Yes	Yes
Year FE	Yes	Yes	No
Observations	14,300	12,203	7,555
R-squared	0.447	0.317	0.274

The table examines the determinants of *FIT*. The sample period is from 2009 to 2019. The model in column (1) report OLS regression estimates with levels of variables, and estimate the potential determinants of the *FIT* measure in year *t*. Column (2) report the results of the same model, except a one year leaded *FIT* measure is used. The model in column (3) examines the same determinants with changes specifications, where *FIT* variable as well as the variables with * are specified in changes from year *t*-1 to year *t*. Note that *Client's Number of CAPs (NUM_CAPS*)* and *Auditor's Number of Deficiencies (NUM_DEF*)*– are determinants of our measure *FIT* by definition. Standard errors are corrected for heteroskedasticity and clustered at client-firm level. The corresponding p-values are reported in parentheses. *, **, *** denote two-tailed significance levels at 10 percent, 5 percent, and 1 percent, respectively. All variables are defined in Appendix A.

Table 8: FIT Measures' Association with Audit Quality

Panel A: Restatements

	RESTATE (1)	*RESTATE* (2)	*RESTATE* (3)	*RESTATE* (4)
FIT	-4.140*** (0.002)			-6.048*** (0.008)
SIM_COMB		0.002 (0.964)		0.027 (0.568)
SPEC_OFFICE			0.037 (0.858)	0.302 (0.301)
TENURE	-0.202* (0.061)	-0.407*** (0.005)	-0.160* (0.099)	-0.405** (0.020)
AUDIT_FEE	-0.187 (0.126)	-0.103 (0.575)	-0.122 (0.289)	-0.184 (0.428)
BIG4	0.019 (0.947)	-0.276 (0.521)	-0.471* (0.070)	-0.202 (0.699)
ICW	1.719*** (0.000)	2.084*** (0.000)	1.532*** (0.000)	1.906*** (0.000)
AUDITOR_CHANGE	-0.001 (0.997)	-0.129 (0.769)	-0.159 (0.523)	-0.044 (0.932)
SIZE	0.148 (0.114)	0.139 (0.336)	0.158* (0.050)	0.163 (0.332)
ROA	0.463 (0.135)	1.631 (0.119)	0.131 (0.440)	1.608 (0.186)
LEVERAGE	0.603*** (0.004)	1.023*** (0.004)	0.102 (0.505)	1.188** (0.018)
LOSS	0.110 (0.606)	0.146 (0.631)	0.353** (0.036)	0.278 (0.418)
BM	0.116 (0.131)	0.138 (0.184)	0.049 (0.416)	0.132 (0.325)
OCF	-0.842*** (0.004)	-1.224** (0.010)	-0.137 (0.272)	-1.042* (0.050)
ARINV	0.916* (0.099)	0.360 (0.592)	0.923*** (0.006)	0.233 (0.771)
FOREIGN	-0.291 (0.227)	-0.049 (0.884)	-0.200 (0.319)	-0.240 (0.552)
FYE_DEC	0.128 (0.570)	0.089 (0.765)	0.103 (0.586)	-0.031 (0.930)
LIQUIDITY	0.006 (0.874)	0.002 (0.983)	-0.032 (0.357)	0.009 (0.918)
lag_NOA	0.017 (0.466)	-0.016 (0.722)	0.006 (0.749)	-0.041 (0.394)
ln_ANALYST	0.052 (0.639)	0.084 (0.620)	0.000 (0.996)	0.134 (0.513)
ln_ZSCORE	-0.007 (0.252)	0.001 (0.954)	0.000 (0.153)	0.001 (0.909)
NUM_CAPS	-0.008 (0.845)	-0.054 (0.226)	0.018 (0.585)	-0.013 (0.828)
SALESVOL	0.342** (0.042)	0.412 (0.159)	0.117 (0.481)	0.389 (0.228)
CASHVOL	-1.312* (0.055)	-0.990 (0.260)	-0.081 (0.677)	-0.911 (0.350)
SEGMENTS	0.052 (0.359)	-0.002 (0.976)	0.038 (0.451)	0.041 (0.641)
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
N	21,694	13,326	24,363	10,059
Pseudo R-squared	0.0796	0.116	0.0574	0.116

Panel B: Discretionary Accruals (Modified Jones – Single Step)

	TACC (1)	***TACC*** (2)	***TACC*** (3)	***TACC*** (4)
FIT	-0.0682**			-0.1013**
	(0.019)			(0.016)
SIM_COMB		-0.0002		-0.0003
		(0.645)		(0.549)
SPEC_OFFICE			0.0151***	0.0015
			(0.008)	(0.632)
TENURE	0.0059***	0.0045***	0.0106***	0.0063***
	(0.000)	(0.004)	(0.002)	(0.001)
AUDIT_FEES	0.0124***	0.0043*	0.0433***	0.0055*
	(0.000)	(0.074)	(0.000)	(0.064)
BIG4	-0.0019	-0.0051	0.0159**	-0.0035
	(0.564)	(0.267)	(0.033)	(0.541)
ICW	-0.0199***	-0.0124**	-0.0167	-0.0109
	(0.000)	(0.032)	(0.177)	(0.124)
AUDITOR_CHANGE	0.0033	-0.0171**	0.0237*	-0.0105
	(0.575)	(0.021)	(0.071)	(0.240)
SIZE	-0.0174***	-0.0106***	-0.0605***	-0.0114***
	(0.000)	(0.000)	(0.000)	(0.000)
ROA	0.3550***	0.2159***	0.3836***	0.2232***
	(0.000)	(0.000)	(0.000)	(0.000)
LEVERAGE	0.0176***	0.0054	0.0392***	-0.0194***
	(0.000)	(0.291)	(0.000)	(0.002)
LOSS	-0.0579***	-0.0664***	-0.0512***	-0.0711***
	(0.000)	(0.000)	(0.000)	(0.000)
BM	-0.0004***	0.0001	-0.0009***	0.0001
	(0.001)	(0.457)	(0.000)	(0.507)
OCF	-0.2599***	-0.1992***	0.0369***	-0.2255***
	(0.000)	(0.000)	(0.000)	(0.000)
ARINV	-0.0031	0.0272***	-0.0487***	-0.0025
	(0.672)	(0.002)	(0.001)	(0.821)
FOREIGN	-0.0036	-0.0062**	-0.0093	-0.0066*
	(0.173)	(0.043)	(0.128)	(0.082)
FYE_DEC	-0.0039	-0.0027	-0.0001	-0.0036
	(0.119)	(0.363)	(0.989)	(0.337)
LIQUIDITY	-0.0009*	0.0016***	-0.0024***	0.0014**
	(0.077)	(0.006)	(0.005)	(0.036)
lag_NOA	0.0017***	0.0003	0.0054***	-0.0001
	(0.000)	(0.541)	(0.000)	(0.828)
ln_ANALYST	-0.0028**	-0.0060***	0.0145***	-0.0067***
	(0.048)	(0.000)	(0.000)	(0.001)
ln_ZSCORE	0.0004***	-0.0001	-0.0001***	-0.0003
	(0.006)	(0.709)	(0.001)	(0.159)
NUM_CAPS	-0.0005	-0.0002	-0.0008	0.0010
	(0.370)	(0.628)	(0.451)	(0.150)
SALESVOL	0.0040	-0.0043	-0.0242***	-0.0033
	(0.346)	(0.440)	(0.000)	(0.612)
CASHVOL	-0.0496***	-0.0651***	-0.0313***	-0.0496***
	(0.000)	(0.000)	(0.000)	(0.000)
SEGMENTS	0.0010	0.0014**	0.0029*	0.0018**
	(0.108)	(0.043)	(0.061)	(0.041)
First Stage Regressors	Yes	Yes	Yes	Yes
Industry x Year FE	Yes	Yes	Yes	Yes
FE x First Stage Regressors	Yes	Yes	Yes	Yes
N	20,001	12,352	22,636	9,415
R-squared	0.5743	0.5608	0.7428	0.5847

Panel C: Missed Material Weaknesses

	MISSED_MW (1)	***MISSED_MW*** (2)	***MISSED_MW*** (3)	***MISSED_MW*** (4)
FIT	-3.209** (0.024)			-4.418* (0.077)
SIM_COMB		-0.009 (0.859)		0.032 (0.537)
SPEC_OFFICE			0.104 (0.597)	0.562* (0.083)
TENURE	-0.192* (0.076)	-0.478*** (0.001)	-0.173* (0.084)	-0.426** (0.018)
AUDIT_FEE	-0.143 (0.341)	-0.074 (0.734)	-0.106 (0.430)	-0.030 (0.916)
BIG4	-0.037 (0.899)	-0.160 (0.741)	-0.568** (0.036)	-0.372 (0.533)
AUDITOR_CHANGE	0.249 (0.405)	0.099 (0.840)	-0.053 (0.848)	0.432 (0.435)
SIZE	0.084 (0.387)	0.015 (0.911)	0.144 (0.102)	0.020 (0.914)
ROA	0.382 (0.136)	1.266 (0.170)	0.130 (0.451)	1.454 (0.213)
LEVERAGE	0.509*** (0.010)	0.596** (0.028)	0.057 (0.710)	0.731 (0.100)
LOSS	0.101 (0.636)	-0.071 (0.820)	0.254 (0.142)	0.116 (0.731)
BM	0.115 (0.272)	-0.020 (0.876)	0.050 (0.465)	0.017 (0.907)
OCF	-0.920*** (0.002)	-1.000** (0.020)	-0.179 (0.151)	-0.855* (0.064)
ARINV	0.935* (0.079)	-0.249 (0.745)	0.642* (0.072)	-0.256 (0.777)
FOREIGN	0.014 (0.952)	0.350 (0.305)	-0.073 (0.740)	0.087 (0.831)
FYE_DEC	-0.236 (0.302)	-0.219 (0.501)	0.008 (0.966)	-0.104 (0.782)
LIQUIDITY	0.034 (0.303)	0.013 (0.852)	-0.034 (0.335)	-0.005 (0.950)
lag_NOA	0.027 (0.196)	0.001 (0.973)	0.004 (0.860)	-0.025 (0.513)
ln_ANALYST	-0.057 (0.610)	-0.146 (0.393)	-0.092 (0.369)	-0.073 (0.708)
ln_ZSCORE	-0.008 (0.135)	-0.003 (0.690)	0.000 (0.523)	-0.003 (0.723)
NUM_CAPS	-0.014 (0.724)	-0.040 (0.401)	0.019 (0.563)	0.014 (0.793)
SALESVOL	0.292 (0.139)	0.075 (0.850)	0.167 (0.355)	-0.150 (0.756)
CASHVOL	-1.536** (0.024)	-0.625 (0.319)	-0.108 (0.613)	-0.313 (0.524)
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
N	22,399	13,551	23,544	8,874
Pseudo R-squared	0.0740	0.0952	0.0465	0.103

Panel D: Reasons for Restatements

#	Area	Self-algorithm Classification (1) MISFIT Coef.	Audit Analytics Classification (2) MISFIT Coef.
1	Allowance For Credits	0.479***	1.149***
3	Business Combinations	1.069*	0.467**
4	Cash & Cash Equivalents	-0.206	
5	Discontinued Operations	2.674***	
8	Fair Value Assets & Liab	0.093	
10	Impairment Of Goodwill & Lla	1.029***	1.578***
11	Income Taxes	0.142*	0.178
12	Insurance Reserves & Claims	0.098	
13	Inventory Valuation	0.773***	0.635**
14	Investments	0.235*	-0.012
15	Leases	-0.202	
16	Litigation & Other Cont.	0.460	
19	Pension and Benefits	-0.714	1.400
20	Property Plant & Equipment	0.662*	0.452**
21	R&D Costs	0.103	-0.412
22	Revenue Recognition	0.053	0.515**
23	Stock-Based Compensation	0.394*	0.531
24	Mortgage Valuation	0.782	

Panel E: Robustness Analysis for Possibly Neglected Areas – Adjusted *FIT*

	RESTATE (1)	***TACC*** (2)	***MISSED_MW*** (3)
FIT_NO_NEGL	-6.000** (0.017)	-0.126*** (0.003)	-3.914* (0.084)
SIM_COMB	0.016 (0.733)	-0.001 (0.176)	-0.002 (0.972)
SPEC_OFFICE	0.287 (0.335)	-0.001 (0.820)	0.492* (0.089)
Controls	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
N	8,396	11,118	13,181
Pseudo / R-squared	0.155	0.512	0.104

The table reports the estimation results of the Equation 5 and examines the association of three different auditor-client compatibility metrics with the following audit quality proxies: Big-R restatements (Panel A), discretionary accruals (Panels B) and missed material weaknesses (Panel C). In Panels A, B and C, the first column of each panel includes our *FIT* measure. Their second and third columns have *SIM_COMB* - the composite similarity score by Brown and Knechel (2016), and *SPEC_OFFICE* - office-level industry specialization as the variable of interest. The fourth column combine the three auditor-client compatibility proxies together in a single model. The models in Panel A and C are estimated using logistic regressions. In Panel B, the coefficients are estimated following the recommendations of Chen et al. (2018) to avoid incorrect inferences. Accordingly, the estimation is done in a single step where the dependent variable are total accruals (*TACC*) while the regressors of the Modified-Jones model, as well as their interactions with industry-year fixed effects are included respectively. Panel D reports the coefficients of *AREA_MISFIT* in the logistic regression model in equation (6) estimated for each accounting area separately, where *AREA_MISFIT* is an indicator for company-years where the area is among the client's CAPs and also is among the auditor's deficient areas. Panel E reports our robustness test to account for areas that may have been neglected by the PCAOB. Models in Panel E estimate the same regressions in Panels A, B and C with a *FIT* measure adjusted for possibly neglected areas. The sample period is from 2009 to 2019. Standard errors are corrected for heteroskedasticity and clustered at client-firm level. The corresponding p-values are reported in parentheses. *, **, *** Denote two-tailed significance levels at 10 percent, 5 percent, and 1 percent, respectively. All variables are defined in Appendix A.

Table 9: FIT's Association with Real Earnings Management (REM)

Panel A: Discretionary Expenses – Single Step

	DISEXP (1)	*DISEXP* (2)	*DISEXP* (3)	*DISEXP* (4)
FIT	0.081*** (0.006)			0.187*** (0.000)
SIM_COMB		-0.001 (0.249)		-0.000 (0.642)
SPEC_OFFICE			0.004 (0.182)	-0.001 (0.819)
TENURE	0.000 (0.995)	0.000 (0.848)	0.003 (0.111)	0.000 (0.912)
BIG4	-0.005** (0.021)	0.001 (0.613)	-0.010*** (0.000)	0.002 (0.543)
ICW	-0.023*** (0.000)	-0.017*** (0.003)	-0.019*** (0.000)	-0.024*** (0.001)
AUDIT_FEES	0.005 (0.405)	0.002 (0.790)	0.004 (0.586)	-0.002 (0.839)
AUDITOR_CHANGE	-0.000 (0.994)	0.002 (0.828)	-0.004 (0.646)	0.001 (0.899)
SIZE	0.014*** (0.000)	0.010*** (0.000)	0.022*** (0.000)	0.010*** (0.000)
ROA	0.698*** (0.000)	0.598*** (0.000)	0.391*** (0.000)	0.543*** (0.000)
LEVERAGE	0.099*** (0.000)	0.099*** (0.000)	0.143*** (0.000)	0.114*** (0.000)
LOSS	-0.024*** (0.000)	-0.030*** (0.000)	-0.029*** (0.000)	-0.037*** (0.000)
BM	0.035*** (0.000)	0.035*** (0.000)	0.046*** (0.000)	0.048*** (0.000)
OCF	0.180*** (0.000)	0.198*** (0.000)	0.313*** (0.000)	0.211*** (0.000)
ARINV	0.108*** (0.000)	0.107*** (0.000)	0.120*** (0.000)	0.135*** (0.000)
FOREIGN	-0.013*** (0.000)	-0.014*** (0.000)	-0.020*** (0.000)	-0.013** (0.012)
FYE_DEC	-0.010*** (0.000)	-0.012*** (0.000)	-0.012*** (0.000)	-0.012*** (0.010)
LIQUIDITY	0.000 (0.666)	0.001 (0.382)	0.001** (0.011)	0.000 (0.921)
lag_NOA	-0.000 (0.349)	-0.000 (0.524)	0.000 (0.656)	-0.001 (0.379)
ln_ANALYST	-0.015*** (0.000)	-0.015*** (0.000)	-0.025*** (0.000)	-0.017*** (0.000)
ln_ZSCORE	-0.001 (0.451)	0.000 (0.913)	0.001 (0.537)	-0.001 (0.752)
NUM_CAPS	-0.001* (0.068)	-0.001 (0.186)	-0.002*** (0.001)	-0.002* (0.059)
SALESVOL	0.090*** (0.000)	0.096*** (0.000)	0.068*** (0.000)	0.089*** (0.000)
CASHSVOL	0.052*** (0.000)	0.040*** (0.000)	0.091*** (0.000)	0.039*** (0.000)
First Stage Regressors	Yes	Yes	Yes	Yes
Industry x Year FE	Yes	Yes	Yes	Yes
FE x First Stage Regressors	Yes	Yes	Yes	Yes
N	19,954	13,707	19,400	8,950
R-squared	0.824	0.819	0.806	0.829

Panel B: Production – Single Step

	PROD (1)	*PROD* (2)	*PROD* (3)	*PROD* (4)
FIT	0.074*** (0.005)			0.155*** (0.000)
SIM_COMB		0.001* (0.066)		0.001** (0.018)
SPEC_OFFICE			-0.002 (0.427)	-0.000 (0.891)
TENURE	0.001 (0.448)	0.003** (0.048)	0.005*** (0.002)	0.005** (0.012)
AUDIT_FEES	-0.014*** (0.000)	-0.011*** (0.000)	-0.011*** (0.000)	-0.010*** (0.002)
BIG4	0.000 (0.989)	-0.003 (0.545)	0.003 (0.348)	-0.005 (0.367)
ICW	0.003 (0.549)	-0.001 (0.819)	0.006 (0.345)	0.003 (0.637)
AUDITOR_CHANGE	-0.004 (0.436)	0.002 (0.791)	-0.003 (0.651)	0.004 (0.639)
SIZE	0.022*** (0.000)	0.016*** (0.000)	0.022*** (0.000)	0.017*** (0.000)
ROA	-0.240*** (0.000)	-0.268*** (0.000)	-0.290*** (0.000)	-0.253*** (0.000)
LEVERAGE	0.053*** (0.000)	0.040*** (0.000)	0.039*** (0.000)	0.039*** (0.000)
LOSS	-0.012*** (0.000)	-0.012*** (0.001)	-0.004 (0.191)	-0.016*** (0.001)
BM	0.021*** (0.000)	0.024*** (0.000)	0.027*** (0.000)	0.027*** (0.000)
OCF	-0.412*** (0.000)	-0.392*** (0.000)	-0.361*** (0.000)	-0.419*** (0.000)
ARINV	0.099*** (0.000)	0.058*** (0.000)	0.107*** (0.000)	0.084*** (0.000)
FOREIGN	-0.015*** (0.000)	-0.015*** (0.000)	-0.026*** (0.000)	-0.015*** (0.000)
FYE_DEC	-0.005** (0.038)	-0.005* (0.066)	-0.003 (0.205)	-0.002 (0.512)
LIQUIDITY	0.002*** (0.001)	0.002*** (0.001)	0.001** (0.013)	0.000 (0.644)
lag_NOA	-0.001*** (0.006)	-0.002*** (0.000)	-0.000 (0.318)	-0.002*** (0.001)
ln_ANALYST	-0.008*** (0.000)	-0.007*** (0.000)	-0.012*** (0.000)	-0.009*** (0.000)
ln_ZSCORE	0.003*** (0.000)	0.002** (0.030)	0.003*** (0.002)	0.002 (0.103)
NUM_CAPS	-0.001*** (0.007)	-0.001** (0.022)	-0.002*** (0.000)	-0.002*** (0.001)
SALESVOL	-0.047*** (0.000)	-0.064*** (0.000)	-0.057*** (0.000)	-0.065*** (0.000)
CASHSVOL	-0.031*** (0.000)	-0.004 (0.523)	-0.042*** (0.000)	0.015 (0.105)
First Stage Regressors	Yes	Yes	Yes	Yes
Industry x Year FE	Yes	Yes	Yes	Yes
FE x First Stage Regressors	Yes	Yes	Yes	Yes
N	21,034	14,486	22,454	12,601
R-squared	0.970	0.974	0.963	0.975

The table reports the estimation results of the Equation 8 and examines the association of four different auditor-client compatibility metrics with the traditional proxies of real earnings management: Abnormal Discretionary Expenses (note that the measure is multiplied by -1, higher levels indicating more earnings management) (Panel A), and Abnormal Production (Panel B). The first column of each panel includes our *FIT* measure. The second and third models of each panel have *SIM_COMB*, - the composite similarity score by Brown and Knechel (2016), and *SPEC_OFFICE* - office-level industry specialization as the variable of interest. The fourth column combines the auditor-client compatibility proxies together in a single model. The coefficients are estimated following the recommendations of Chen et al. (2018) to avoid incorrect inferences. Accordingly, the estimation is done in a single step where the regressors of the discretionary expenses model and the production model, as well as their interactions with industry-year fixed effects are included respectively. Note that the R-Squared values are affected by the autoregressive nature of the production model we adopted from Kothari et al. (2016). Standard errors are corrected for heteroskedasticity and clustered at client-firm level. The p-values are reported in parentheses. *, **, *** denote two-tailed significance levels at 10%, 5%, and 1%, respectively.

Table 10: Analysis of Earnings Response Coefficients for Auditor-Client Pairs

	Dependent Variable: CAR						
Compatibility Variable →	*FIT*		*SIM_COMB*		*SPEC_OFFICE*		*ALL*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
UE	0.039***	0.028***	0.044***	0.038***	0.031***	0.026***	0.021***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
UE x TOP_FIT	0.006*	0.008**					0.012*
	(0.093)	(0.040)					(0.067)
UE x TOP_SIM_COMB			-0.016***	-0.017***			-0.012*
			(0.001)	(0.002)			(0.068)
UE x SPEC_OFFICE					0.025***	0.019***	0.033***
					(0.000)	(0.000)	(0.000)
TOP_FIT	0.000	-0.001					-0.001
	(0.989)	(0.442)					(0.589)
TOP_SIM_COMB			0.001	0.000			0.001
			(0.387)	(0.811)			(0.298)
SPEC_OFFICE					-0.000	-0.001	-0.001
					(0.542)	(0.139)	(0.262)
SIZE		-0.001***		-0.001***		-0.001***	-0.001*
		(0.000)		(0.002)		(0.000)	(0.055)
MTB		0.000***		0.000***		0.000***	0.000***
		(0.000)		(0.000)		(0.000)	(0.000)
LEVERAGE		-0.003**		-0.004**		-0.002	-0.003
		(0.045)		(0.043)		(0.172)	(0.251)
BETA		0.385***		0.448***		0.414***	0.442***
		(0.000)		(0.000)		(0.000)	(0.000)
LOSS		-0.022***		-0.022***		-0.021***	-0.022***
		(0.000)		(0.000)		(0.000)	(0.000)
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year x Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Auditor FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	84,029	73,269	50,866	45,266	66,210	58,600	31,829
R-squared	0.010	0.019	0.007	0.018	0.010	0.019	0.021

The table reports the OLS estimation results of the Equation 7 and examines differential investor reactions to earnings announcements of companies with highly compatible auditors versus non-compatible auditors, using four different proxies for compatibility. The dependent variable Cumulative Abnormal Returns (CAR) are a company's 3 trading day return, centered on the quarterly earnings announcement date. Unexpected earnings (*UE*) is based on the difference between the I/B/E/S actual quarterly EPS and the analysts' mean quarterly EPS forecast. *TOP_FIT* and *TOP_SIM_COMB* are coded as indicators for the auditor-client pairs that had high (i.e. above median) *FIT* and *SIM_COMB* scores in the fiscal years prior to the earnings announcement, respectively. See Appendix A for detailed description of the variables. The corresponding p-values are reported in parentheses. *, **, *** denote two-tailed significance levels at 10 percent, 5 percent, and 1 percent, respectively.