Internet Appendix

I. Control Variables

Audit opinions on internal controls are likely to be persistent if companies fail to remediate their control weaknesses, or if the auditor’s approach to testing and evaluating internal controls is similar from one year to the next. Thus, we control for the prior year’s internal control opinion ($ICOP_{i,t-1}$), which equals one if the company received an adverse opinion in the previous year, and zero otherwise. We expect a positive coefficient on $ICOP_{i,t-1}$.

We control for material misstatements ($MISSTATE_{i,t}$) because they indicate a material internal control weakness over financial reporting, signifying that the company should receive an adverse internal control opinion (Rice and Weber [2012]). $MISSTATE_{i,t}$ equals one if the company’s financial statements are restated subsequent to the year covered by the internal control audit report, and zero otherwise. We identify material misstatements corresponding to all restatements announced up to May 2015, where the lag between the misstatement and the subsequent restatement announcement is typically less than one year (median lag = 180 days). We expect a positive coefficient on $MISSTATE_{i,t}$ because auditors are more likely to issue adverse opinions when companies are materially misstating their financial statements. We also control for auditor resignations ($RESIGN_{i,t}$) because they can signal the existence of internal control weaknesses. $RESIGN_{i,t}$ equals one if an auditor resignation is announced during the past 12 months, and zero otherwise. Rice and Weber [2012] find that audit fees are positively associated with reported internal control weaknesses, so we also include the natural log of audit fees ($ln(AF)_{i,t}$). The positive association between adverse opinions and audit fees may reflect additional audit effort to detect internal control weaknesses, or the auditor’s
pricing response to higher control risk. We also control for the log of non-audit service fees ($ln(NAS)_{i,t}$), since non-audit services may include consulting related to risk management and internal controls, which is likely to reduce the likelihood of an adverse opinion.

We control for client size using the natural log of the company’s market capitalization ($SIZE_{i,t}$), because smaller companies may have insufficient personnel and/or expertise to maintain high quality controls. Because poorly performing companies are more likely to have control weaknesses, we also include a loss dummy ($LOSS_{i,t}$), which takes the value one if net income is negative and zero otherwise. Multi-segment companies can face internal control problems due to the complexity of their operations, and multinationals may find it difficult to have effective internal controls due to their geographic and institutional diversity. Thus, following Ashbaugh et al. [2007] and Doyle et al. [2007], we control for the number of business segments ($SEGMENTS_{i,t}$) and include a $FOREIGN_{i,t}$ dummy, which equals one if the company has a foreign currency translation, and zero otherwise. Also following Ashbaugh et al. [2007], we control for the ratio of inventory to total assets ($INVENTORY_{i,t}$). Companies with more inventory face increased risks related to measurement and recognition issues and an increased risk of theft. We also control for sales growth ($GROWTH_{i,t}$) because the internal control system of a growing company may fail to keep pace with the rate of expansion of the company’s operations.

We control for external financing ($XTFIN_{i,t}$) using the sum of newly issued equity and debt, divided by total assets. Companies intending to raise external financing are likely to prefer clean internal control opinions. On the other hand, external financing may expose the auditor to a greater risk of reputational or litigation losses if the auditor fails to
disclose existing weaknesses. Finally, we control for mergers and acquisitions by including a dummy variable \( (M&A_{i,t}) \), which equals one if the company is involved in a merger or acquisition during the year, and zero otherwise. Acquisitions present significant challenges in internal control system integration. We also include a dummy to capture restructurings \( (RESTRUCTURE_{i,t}) \), which is coded one if the company restructures its operations during the year, and zero otherwise. Restructuring can cause control weaknesses due to the loss of critical staff and supervision problems.

II. Sensitivity Tests

COMPANY FIXED EFFECTS

We control for audit firm and year fixed effects in our primary analysis. We do not control for company fixed effects because this would necessitate dropping companies that have no time variation in the dependent variable \( (ICOP_{i,t}) \); i.e., the fixed effects logit model requires that a company receive both a clean opinion and an adverse opinion at some point during the four-year sample period. To test whether our results are sensitive to controlling for company fixed effects, we examine the sub-sample of companies that do not receive the same internal control opinion in every sample year, reducing our sample from 13,933 to just 1,023 observations. As reported in Table A, the results remain significant in this fixed effects analysis. When the fixed effects are estimated for the model of internal control reporting, the \( DEF_{IC\%_{i,t}} \) coefficient is positive and statistically significant (z-stat. = 4.14). The \( DEF_{IC\%_{i,t}} \) coefficient is also positive and statistically significant in the audit fee model (t-stat. = 4.83). In addition, we estimate

\[ \text{\textsuperscript{1}} \text{The fixed effects specification drops 12,847 observations for companies that receive clean opinions in every year, and 63 observations for companies that receive adverse opinions in every year, which reduces our sample to 1,023 = 13,933 – 12,847 – 63.} \]
random effects specifications, which do not result in sample attrition because the random effects models retain companies that lack time variation in the dependent variable. In the random effects specifications, the $DEF_{IC%_{i,t}}$ coefficients are positive and significant in both the internal control reporting model ($z$-stat. = 4.70) and the audit fee model ($t$-stat. = 3.05). Hence, our results are robust to controlling for unobserved time-invariant company characteristics.

AN ALTERNATIVE MEASURE OF INTERNAL CONTROL DEFICIENCY RATES

While our tests examine the level of deficiencies, auditors may also respond to changes in deficiency rates. Thus, in untabulated analysis we repeat our analysis in Table 4 after replacing $DEF_{IC%_{i,t}}$ with the change in the deficiency rate ($\Delta DEF_{IC%_{i,t}}$). We drop the first sample year (2010) because we do not observe the number of inspected audits before 2010, causing the sample to fall to 9,989 observations. The coefficient on $\Delta DEF_{IC%_{i,t}}$ remains positive and highly significant ($z$-statistic = 6.27). Therefore, our analysis is robust to replacing the level of the audit firm’s reported deficiency rate with the change in this rate.

AUDIT DEFICIENCIES THAT ARE UNRELATED TO INTERNAL CONTROL AUDITS

Some inspected audits have both internal control audit deficiencies and other types of audit deficiencies (primarily deficiencies in substantive testing or analytical procedures). To test whether our results are driven by the combination of internal control and other types of deficiencies, we re-estimate our Table 4 analysis after replacing $DEF_{NOT\_IC%_{i,t}}$ with $DEF_{%_{i,t}}$, where $DEF_{%_{i,t}}$ equals the total number of audits that are found to have any type of deficiency divided by the total number of audits examined during the inspection. In untabulated analysis, the coefficient on $DEF_{IC%_{i,t}}$ remains
positive and significant (z-statistic = 3.72) whereas the coefficient on $DEF_{%i,t}$ is insignificant (z-statistic = 0.42). This supports our Table 4 analysis, which finds that auditors’ internal control opinions are affected by the inspectors’ prior disclosure of internal control deficiencies but not by their prior disclosure of other types of audit deficiencies.

**USING UNSCALED INTERNAL CONTROL AUDIT DEFICIENCIES**

In untabulated analysis we find similar results if we do not scale the number of deficient audits by the number of inspected engagements. Our inferences are also unchanged if we estimate eq. (1) using the unscaled deficiency variable for the entire period for which we have auditors’ reports on internal controls (2005-2013).

**AN ALTERNATIVE POST-INSPECTION WINDOW**

As shown in Fig. 1, our tabulated analysis uses the dates of the inspection reports to measure the post-inspection windows. However, remediation could begin earlier than this because the inspectors typically discuss their findings with audit firms before releasing the inspection reports to the public. Therefore, as a robustness test, we re-estimate the models of Table 4 after constructing the post-inspection windows using the inspection completion dates rather than the inspection report dates. In untabulated analysis, we continue to find significant positive coefficients on $DEF_{IC\%i,t}$ in the full sample, and the sub-samples of Big Four auditors and Non-Big Four auditors (z-stats. = 5.26, 4.04, 3.28).

**DROPPING THE MATERIALITY DEFICIENCIES**

In 2007, the switch from AS2 to AS5 changed the threshold for classification of a material weakness from “more than remote” to “reasonably possible,” respectively. Thus, the increase in adverse opinions may result from differences between the way inspectors
and auditors interpret the new definition of a material weakness. If so, our results might be explained by a change in the way auditors classify material weaknesses, rather than an improvement in audit quality, *per se*.

We investigate this by rerunning our analysis in Table 4 after redefining $DEF_{IC\%_i,t}$ to include only the audits with deficiencies related to *tests* of internal controls, and exclude the audits with deficiencies related to the inappropriate *evaluation* of materiality. This analysis finds that the coefficient on $DEF_{IC\%_i,t}$ remains significantly positive (t-stat. = 4.44). We also repeat our analysis of audit fees in Table 9 after redefining $DEF_{IC\%_i,t}$ to include only the testing deficiencies, and we continue to find a significant positive coefficient on $DEF_{IC\%_i,t}$ (t-stat. = 2.97). Thus, our results are not explained by differences in the way that auditors and inspectors classify material weaknesses.

**DO CLIENT CHARACTERISTICS EXPLAIN THE UPWARD TREND IN ADVERSE OPINIONS?**

As discussed in Section 4.1.3, an increase in average client risk could potentially explain the increase in adverse opinions reported in Table 1. While we control for client characteristics in our multivariate tests, in this section we investigate whether such characteristics contribute to explaining the time trend observed in Table 1. Specifically, we use client characteristics to model the probability of auditors issuing adverse opinions during 2010-2013, then regress this probability on a yearly time trend. In untabulated analysis, we find a negative but insignificant coefficient on the time trend variable (t-stat. = -1.29). Thus, the upward trend in adverse opinions is not explained by clients becoming riskier over the 2010-2013 time period.
We also repeat our Table 4 analysis after dropping clients that entered or exited the sample to ensure that our results are not attributable to changes in the composition of companies. We continue to find a significant positive coefficient on $DEF_{IC\%i,t}$ (t-stat. = 4.47). In addition, we repeat our analysis of audit fees (reported in Table 9) after dropping clients who entered or exited the sample, and we continue to find a significant positive coefficient on $DEF_{IC\%i,t}$ (t-stat. = 2.57).

**CONTROLLING FOR ADVERSE INTERNAL CONTROL OPINIONS IN ESTIMATING AUDIT FEES**

A potential omitted correlated variable in our analysis of audit fees in Table 9 is the auditor’s issuance of an adverse internal control audit opinion. Thus, we rerun the Table 9 analysis after including $ICOP_{i,t}$ as a control variable. In untabulated analysis, we find a significant positive coefficient on $ICOP_{i,t}$ (t-stat. = 6.73). However, we also continue to find a significant positive coefficient on $DEF_{IC\%i,t}$ (t-stat. = 2.55). This indicates that remediation of the deficiencies identified in the inspections is associated with higher audit fees, even after controlling for the effect on audit fees of the auditor’s discovery (and reporting) of a material internal control weakness.

**III. Inherent Difficulties in Calculating “Type I Errors” for Internal Control Audit Opinions**

Calculating so-called “Type I errors” for internal control audit opinions is inherently challenging, and likely to lead to misleading inferences about the quality of internal control audits. This is because the issuance of an adverse ICOFR opinion is not a prediction by the auditor that a misstatement is likely to occur. Rather an adverse opinion indicates the presence of a material weakness that presents “a reasonable possibility” or “more than a remote possibility” that a material misstatement will not be prevented or
detected on a timely basis. Thus, when auditors issue adverse ICOFR opinions to clients that do not subsequently report a misstatement, we cannot infer that the auditor issued the “wrong” opinion. We also note that it is highly unlikely that auditors would issue an adverse ICOFR opinion to a client that does not actually have a material internal control weakness, because this would increase the likelihood of the company dismissing the auditor from the engagement (e.g., Newton et al., 2016). Therefore, auditors are strongly motivated to avoid issuing adverse opinions unnecessarily.

A simple numerical example will illustrate that Type I errors are likely to give misleading inferences about the quality of internal control audits. Suppose that the Type I error rate is calculated as the joint probability that a company receives an adverse internal control opinion ($ICOP = 1$) and its financial statements are not subsequently restated ($RESTATE = 0$). Then the probability of a Type I error is $\text{Prob}(ICOP = 1 \& RESTATE = 0)$. The key question is whether this probability is indicative of a low quality or a high quality audit. If it indicates a low quality audit, then the Type I error rate would be a useful measure of audit quality; if it captures a high quality audit, then the Type I error rate would give an incorrect inference that audit quality is low when in fact it is high.

To assess whether the Type I error rate is likely to be a good measure of audit quality, let’s consider a numerical example. Assume that:

- if the company has a material internal control weakness, a high quality auditor issues an adverse opinion with probability 90%.
- if the company does not have a material internal control weakness, a high quality auditor issues an adverse opinion with probability 1%.
- if the company has a material internal control weakness, a low quality auditor issues an adverse opinion with probability 40%.
- if the company does not have a material internal control weakness, a low quality auditor issues an adverse opinion with probability 5%.
The above probabilities reflect that companies with internal control weaknesses are more likely to receive adverse opinions, and high quality auditors are more likely to issue the “correct” opinion.

Suppose that 30% of companies have material internal control weaknesses and 80% of audits are high quality. Further, suppose that:

- if the company has a material internal control weakness, its financial statements are restated with probability 20%.
- if the company does not have a material internal control weakness, its financial statements are restated with probability 0%.

These probabilities reflect that companies only misstate their financial statements if they have material internal control weaknesses, although not all companies with material weaknesses have subsequent restatements.

Based on the above assumptions, there are 12 possible outcomes: ²

i. The company has a material weakness (30%), and audit quality is high (80%), and the auditor issues an adverse opinion (90%), and the company subsequently restates (20%).

ii. The company has a material weakness (30%), and audit quality is high (80%), and the auditor issues an adverse opinion (90%), and the company does not subsequently restate (80%).

iii. The company has a material weakness (30%), and audit quality is high (80%), and the auditor issues a clean opinion (10%), and the company subsequently restates (20%).

iv. The company has a material weakness (30%), and audit quality is high (80%), and the auditor issues a clean opinion (10%), and the company does not subsequently restate (80%).

v. The company has a material weakness (30%), and audit quality is low (20%), and the auditor issues an adverse opinion (40%), and the company subsequently restates (20%).

² There are 4 binary decisions, which result in 16 outcomes ($2^4 = 16$). However, companies without material weaknesses have a 0% probability of restatement, so the number of positive probability outcomes reduces to 12 ($2^3 + 4 = 12$).
vi. The company has a material weakness (30%), and audit quality is low (20%), and the auditor issues an adverse opinion (40%), and the company does not subsequently restate (80%).

vii. The company has a material weakness (30%), and audit quality is low (20%), and the auditor issues a clean opinion (60%), and the company subsequently restates (20%).

viii. The company has a material weakness (30%), and audit quality is low (20%), and the auditor issues a clean opinion (60%), and the company does not subsequently restate (80%).

ix. The company does not have a material weakness (70%), and audit quality is high (80%), and the auditor issues an adverse opinion (1%), and the company does not subsequently restate (100%).

x. The company does not have a material weakness (70%), and audit quality is high (80%), and the auditor issues a clean opinion (99%), and the company does not subsequently restate (100%).

xi. The company does not have a material weakness (70%), and audit quality is low (20%), and the auditor issues an adverse opinion (5%), and the company does not subsequently restate (100%).

xii. The company does not have a material weakness (70%), and audit quality is low (20%), and the auditor issues a clean opinion (95%), and the company does not subsequently restate (100%).

The probabilities for each of these 12 outcomes are as follows:

i. 4.32% = 30% x 80% x 90% x 20%.

ii. 17.28% = 30% x 80% x 90% x 80%.

iii. 0.48% = 30% x 80% x 10% x 20%.

iv. 1.92% = 30% x 80% x 10% x 80%.

v. 0.48% = 30% x 20% x 40% x 20%.

vi. 1.92% = 30% x 20% x 40% x 80%.

vii. 0.72% = 30% x 20% x 60% x 20%.

viii. 2.88% = 30% x 20% x 60% x 80%.
ix.  $0.56\% = 70\% \times 80\% \times 1\% \times 100\%$.

ox.  $55.44\% = 70\% \times 80\% \times 99\% \times 100\%$.

xi.  $0.70\% = 70\% \times 20\% \times 5\% \times 100\%$.

xii. $13.30\% = 70\% \times 20\% \times 95\% \times 100\%$.

A Type I error is assumed to occur when the auditor issues an adverse opinion and the company does not restate. This occurs in four of the 12 outcomes: i.e., ii, vi, ix, xi.

Thus, the following four outcomes would be classified as Type I errors:

ii. The company has a material weakness (30%), and audit quality is high (80%), and the auditor issues an adverse opinion (90%), and the company does not subsequently restate (80%). This occurs with probability 17.28%.

vi. The company has a material weakness (30%), and audit quality is low (20%), and the auditor issues an adverse opinion (40%), and the company does not subsequently restate (80%). This occurs with probability 1.92%.

ix. The company does not have a material weakness (70%), and audit quality is high (80%), and the auditor issues an adverse opinion (1%), and the company does not subsequently restate (100%). This occurs with probability 0.56%.

xi. The company does not have a material weakness (70%), and audit quality is low (20%), and the auditor issues an adverse opinion (5%), and the company does not subsequently restate (100%). This occurs with probability 0.70%.

The Type I error methodology *correctly* classifies outcomes vi and xi as low quality audits. In contrast, the Type I error methodology *incorrectly* classifies cases ii and ix as low quality audits. Thus, we can now calculate how often the Type I methodology gives the correct answer versus the wrong answer. The Type I methodology gives the correct answer in situations vi and xi, which occur with probabilities 1.92% and 0.70%, respectively. The Type I methodology gives the incorrect answer in situations ii and ix,
which occur with probabilities 17.28% and 0.56%, respectively. Thus, in this numerical example, the Type I methodology is incorrect more often than it is correct (i.e., 17.28% + 0.56% > 1.92% + 0.70%). The main source of error with the Type I methodology is in situation ii; i.e., the company has a material weakness, the high quality auditor correctly issues an adverse opinion, but the company does not subsequently restate. This illustrates that so-called Type I errors might actually capture high quality audits rather than low quality audits. This is because the absence of a subsequent restatement does not indicate the absence of a material internal control weakness.
Controlling for company-specific effects.

Z-statistics and t-statistics are reported in parentheses below the coefficients. Standard errors are corrected for clustering on each company. We use the same control variables (\textit{CONTROLS}) as in Tables 4 and 7 but results for the control variables are untabulated.

<table>
<thead>
<tr>
<th></th>
<th>Internal control reporting (\textit{ICOP})</th>
<th>Audit fees (\textit{Ln(AF)})</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Random effects</td>
<td>Fixed effects</td>
</tr>
<tr>
<td>\textit{DEF_IC}%\textsubscript{i,t}</td>
<td>3.19***  (4.70)</td>
<td>5.19***  (4.14)</td>
</tr>
<tr>
<td>\textit{CONTROLS}?</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Obs.</td>
<td>13,933</td>
<td>1,023</td>
</tr>
</tbody>
</table>

***, **, * denote statistically significant at the 1%, 5%, 10% levels, respectively (two-tailed tests).

See Appendix for variable definitions.