Auditing: A Journal of Practice & Theory
Vol. 33, No. 4
November 2014
pp. 197–219

American Accounting Association
DOI: 10.2308/ajpt-50792

Do Big 4 Auditors Provide Higher Audit Quality after Controlling for the Endogenous Choice of Auditor?

John Daniel Eshleman and Peng Guo

SUMMARY: Recent research suggests that Big 4 auditors do not provide higher audit quality than other auditors, after controlling for the endogenous choice of auditor. We re-examine this issue using the incidence of accounting restatements as a measure of audit quality. Using a propensity-score matching procedure similar to that used by recent research to control for clients' endogenous choice of auditor, we find that clients of Big 4 audit firms are less likely to subsequently issue an accounting restatement than are clients of other auditors. In additional tests, we find weak evidence that clients of Big 4 auditors are less likely to issue accounting restatements than are clients of Mid-tier auditors (Grant Thornton and BDO Seidman). Taken together, the evidence suggests that Big 4 auditors do perform higher quality audits.

Keywords: Big 4 auditor; audit quality; propensity-score matching; audit quality proxies.

JEL Classifications: M41; M42.

Data Availability: All data are publicly available from sources identified in the text.

INTRODUCTION

One of the earliest theories in the audit literature is that Big 4[1] auditors, due to their larger size and better training programs, provide higher audit quality than other auditors. The argument is that larger audit firms have more reputation to lose by sacrificing their independence on any given audit engagement (DeAngelo 1981). In addition, larger audit firms have

John Daniel Eshleman is an Assistant Professor at Oklahoma State University, and Peng Guo is a Ph.D. Student at Louisiana State University.

We thank Donald J. Stokes, two anonymous reviewers, Qiang Cheng, Neil Bhattacharya, Jae Bum Kim, Chee Yeow Lim, Jeff Ng, Tharindra Ranasinghe, Ken Reichelt, Jared Soileau, Yoonseok Zang, and all workshop participants at Singapore Management University. All errors that remain are our own.

Editor's note: Accepted by Donald J. Stokes.

Submitted: February 2013
Accepted: April 2014
Published Online: April 2014

[1] We use the term Big 4 to refer to the Big 5 or Big 4 accounting firms.

Downloaded from http://meridian.allenpress.com/ajpt/article-pdf/33/4/197/1603079/ajpt-50792.pdf by Case Western Reserve University, Charles Gletsu on 23 February 2021

more resources to invest in training programs, resulting in better trained auditors. Early literature provided evidence consistent with DeAngelo's theory. For example, Francis and Krishnan (1999) find that Big 4 auditors exhibit greater conservatism when issuing audit reports. There is also evidence that investors place more weight on the earnings of a firm audited by a Big 4 auditor, consistent with investors viewing the earnings as being of higher quality (Teoh and Wong 1993). In addition, it has been shown that clients of Big 4 auditors exhibit higher earnings quality in the form of a lower magnitude of discretionary accruals (Becker, DeFond, Jiambalvo, and Subramanyam 1998; Francis, Maydew, and Sparks 1999). To the extent that discretionary accruals capture opportunistic earnings management, this implies that Big 4 auditors tolerate less earnings management than other auditors.

However, there is a potential endogeneity problem with the research discussed above. Firms select their auditors and auditors decide if they will accept the firm as their client. Audit firms will tend to prefer less risky clients with higher earnings quality.[2] Thus, it is not clear from the early research (Teoh and Wong 1993; Becker et al. 1998) that Big 4 auditors have higher audit quality. Lawrence, Minutti-Meza, and Zhang (2011) correct for this endogenous choice of the firm's auditor using a propensity-score matching approach. After correcting for the endogeneity inherent in client selection, the authors provide evidence that suggests that clients of Big 4 auditors exhibit audit quality similar to the matched sample of non-Big 4 clients.[3] In a related study, Boone, Khurana, and Raman (2010) find that Big 4 auditors and Mid-tier auditors (i.e., Grant Thornton and BDO Seidman) exhibit similar audit quality, where audit quality is measured using the absolute value of discretionary accruals. The authors also find "weak evidence that the Big 4 have a higher propensity to issue going concern audit opinions for distressed companies" (Boone et al. 2010, 330). Boone et al. (2010) conclude that, among Big 4 and Mid-tier audit firm clients, there is little difference in actual audit quality.

The purpose of this paper is to re-examine whether Big 4 auditors deliver higher audit quality after controlling for the endogenous choice of auditor. In this study, we choose an audit quality proxy, which we believe better captures whether the client engaged in non-GAAP reporting. Our proxy is the likelihood of a firm issuing an accounting restatement. We construct our sample by matching non-Big 4 clients with Big 4 clients via a propensity-score matching model similar to that used by Lawrence et al. (2011) and Boone et al. (2010). We first replicate the results from Lawrence et al. (2011) and find similar results for our sample, implying that clients of the Big 4 do not receive higher quality audits than clients of the non-Big 4. However, our restatement analysis tells a different story. We find that clients of non-Big 4 auditors are significantly more likely to subsequently issue an accounting restatement than are clients of the Big 4. This result holds after controlling for a set of innate firm characteristics known to affect the likelihood of issuing a restatement. This is consistent with non-Big 4 auditors allowing a higher frequency of material misstatements than Big 4 auditors. In additional analyses, we present evidence that clients of Big 4 auditors are significantly less likely to be sanctioned by the SEC for an Accounting and Auditing Enforcement Release (AAER) than are clients of other auditors.

We then use the same propensity-score matching approach to construct a matched sample of Big 4 and Mid-tier auditors. We find that, when using the matched sample, clients of the Big 4 are significantly less likely to subsequently issue an accounting restatement than are clients of Mid-tier auditors. However, these results do not hold when using the full (non-matched) sample.

[2] In addition, small auditors are likely unable to audit the largest companies, due to capacity constraints and litigation costs.

[3] Specifically, the authors find that, among Big 4 and non-Big 4 clients, there is little difference in the magnitude of discretionary accruals, analyst forecast accuracy, and the cost of equity capital.

Downloaded from http://meridian.allenpress.com/ajpt/article-pdf/33/4/197/1603079/ajpt-50792.pdf by Case Western Reserve University, Charles Gletsu on 23 February 2021

We also construct a matched sample of Mid-tier and small auditors. We find no evidence that Mid-tier auditors provide higher audit quality than the small audit firms. Finally, we construct a matched sample of Big 4 and small auditors. We find that clients of small auditors are significantly more likely to subsequently issue an accounting restatement than are clients of the Big 4. Taken together, the evidence suggests a hierarchy of audit firms, with Big 4 auditors providing the highest audit quality, small auditors providing the lowest level of audit quality, and Mid-tier auditors providing audit quality in between the Big 4 and the small auditors.

The evidence presented in this paper is of interest to managers, audit committees, investors, creditors, and regulators. Managers and audit committees would like to know whether the Big 4 actually do provide higher quality audits. This information will help them choose an auditor. Given that Big 4 auditors earn a fee premium (Ireland and Lennox 2002), managers and audit committees must decide whether the services they receive from the auditor are worth the premium. Investors and creditors will also be interested in our results, as this will help them assess the credibility of firms' financial reports. Regulators are also interested in whether the Big 4 accounting firms actually provide higher quality audits. Recently, some have suggested that smaller firms can provide audits of similar quality. For example, the U.S. Chamber of Commerce has recently suggested that "all parties should actively encourage public companies to consider high quality firms outside the Big 4" (U.S. Chamber of Commerce 2006, 18).

This paper contributes to the audit literature by providing evidence of superior audit quality at Big 4 audit firms. Recent research by Lawrence et al. (2011) and Boone et al. (2010) suggests that differences in audit quality among Big 4 and non-Big 4 clients may be attributable to the different clienteles of Big 4 and non-Big 4 auditors. Our study complements and extends these two studies by using a different proxy for audit quality. The fact that we draw different inferences using a different audit quality proxy underscores the importance of using caution when selecting audit quality proxies. We believe focusing on tangible outputs of the audit, such as accounting restatements, provides better tests of audit quality.[4]

The rest of the paper is organized as follows. The following section contains a literature review and our hypotheses, the third section contains the research design, the fourth contains the sample selection, and the fifth reports the replication of Lawrence et al. (2011). The sixth section reports the results for the Big 4 versus non-Big 4 analysis, the seventh section reports the results for the Big 4 versus Mid-tier analysis, the eighth section contains additional analysis, and the ninth concludes.

LITERATURE REVIEW AND HYPOTHESIS DEVELOPMENT

In a seminal paper, DeAngelo (1981) makes the argument that Big 4 auditors will provide higher audit quality because they have "more to lose." Big 4 auditors have a larger client portfolio than other audit firms. If the auditor is perceived to have allowed the client to manage earnings, then investors will view the earnings quality of all clients of the Big 4 auditor unfavorably. As a result, many of the auditor's clients will demand lower audit fees and/or switch to a different auditor. Therefore, a larger audit firm has more to lose by compromising its independence. In addition, there is the argument that Big 4 auditors' large size enables them to invest in more high quality training and audit technology (DeAngelo 1981; Boone et al. 2010). The early audit literature provides evidence in support of this notion. Teoh and Wong (1993) find that clients of Big 4 auditors exhibit

[4] A disadvantage of using restatements or going concern opinions for audit quality is that these are rare events. In our sample, approximately 7 percent of all firm-years involve an accounting restatement. Therefore, this audit quality proxy cannot tell us anything about how the audit quality differs among the other 93 percent of the firm-years that had no accounting restatements. It only tells us that audit quality was worse for the 7 percent of firms that did later issue a restatement.

American Accounting Association

Downloaded from http://meridian.allenpress.com/ajpt/article-pdf/33/4/197/1603079/ajpt-50792.pdf by Case Western Reserve University, Charles Gletsu on 23 February 2021

higher earnings response coefficients, consistent with investors viewing these clients' earnings as being of higher quality. There is also evidence that Big 4 auditors report more conservatively (Francis and Krishnan 1999). Research also finds that clients of Big 4 auditors generally exhibit higher earnings quality, which is consistent with auditors tolerating less earnings management (Becker et al. 1998; Francis et al. 1999). Finally, Lennox and Pittman (2010) show that clients of Big 4 auditors are less likely to commit fraud, even when controlling for the endogenous choice of auditor.[5]

For several years, there was little research on the issue; the literature had produced a large body of evidence suggesting that Big 4 auditors provided higher quality audits than smaller auditors.[6] However, Lawrence et al. (2011) cast doubt on the superiority of Big 4 auditors. Lawrence et al. note that a firm's choice of an auditor is endogenous.[7] Firms with better performance and higher quality earnings are more likely to choose Big 4 auditors. Similarly, Big 4 auditors will prefer less risky clients with higher earnings quality. The authors control for the endogenous choice of auditor by modeling the choice of an auditor using a propensity-score matching model to match each non-Big 4 client with a Big 4 client. Using this matched sample, the authors show that clients of Big 4 auditors do not exhibit higher audit quality than clients of non-Big 4 auditors. Boone et al. (2010) is another study related to this issue. The authors examine whether clients of Big 4 auditors have higher audit quality than clients of the Mid-tier audit firms (i.e., Grant Thornton and BDO Seidman). They find that Big 4 audit firms do not perform higher quality audits than Mid-tier audit firms.

Lawrence et al. (2011) choose three proxies for the unobservable audit quality. The first proxy is an earnings quality metric, the absolute value of discretionary accruals estimated from the Jones (1991) model. The justification for using an earnings quality metric to proxy for audit quality is that higher quality auditors will tolerate less earnings management from the client. Therefore, firms with higher magnitudes of discretionary accruals are assumed to have received lower quality audits. The second proxy is the accuracy of financial analysts' forecasts (Behn, Jong-Hag, and Kang 2008). The more accurate are the forecasts, the higher is the earnings quality, which implies that the auditors must have tolerated less earnings management. The third proxy is the *ex ante* cost of equity capital. The rationale is that firms with more credible financial statements will have a lower perceived information risk, which will result in a lower cost of capital.

There are reasons to believe that the three audit quality proxies chosen by Lawrence et al. (2011) may not be capturing audit quality differences between Big 4 and non-Big 4 auditors. First, by using the magnitude of discretionary accruals to proxy for audit quality, the researcher is assuming that all discretionary accruals are equally harmful to earnings quality. This need not be the case; managers can use discretionary accruals to signal firm value (Subramanyam 1996). Second, the use of analyst forecast accuracy to proxy for audit quality relies on a very indirect link from auditor quality to earnings quality to analyst accuracy (Behn et al. 2008). It seems more reasonable to examine the relationship between the auditor and earnings quality. Finally, the cost of equity

[5] See Francis (2004) for a more complete review of the audit quality literature.

[6] Although there was little research that directly examined whether Big 4 auditors provide higher audit quality than non-Big 4 auditors, several studies of audit quality included Big 4 as a control variable. For example, Choi, Kim, Qiu, and Zang (2012) find that Big 4 auditors are associated with a lower magnitude of discretionary accruals. Asthana and Boone (2012) find that Big 4 clients are less likely to meet or beat analysts' earnings per share (EPS) forecasts by two pennies or less, and have a lower magnitude of discretionary accruals.

[7] Lawrence et al. (2011) are not the first to note that the choice of auditor is endogenous. Hogan (1997), Ireland and Lennox (2002), and Lennox and Pittman (2010) all note that the choice of auditor is endogenous, and control for this in their research designs. Early research by J. (Krish) Krishnan and J. Krishnan (1996) also controls for client risk via a two-stage approach (see also Fargher and Jiang 2008).

Downloaded from http://meridian.allenpress.com/ajpt/article-pdf/33/4/197/1603079/ajpt-50792.pdf by Case Western Reserve University, Charles Gletsu on 23 February 2021

capital cannot speak to differences in audit quality; this variable measures investors' *perceptions* of audit quality.[8]

The purpose of this study is to re-examine whether Big 4 auditors deliver higher audit quality when the endogeneity of auditor choice is controlled via a propensity-score matching approach. Our proxy for audit quality is the likelihood of a firm issuing an accounting restatement. Consistent with recent literature, we only consider restatements attributable to a failure in the application of GAAP (Newton, Wang, and Wilkins 2013).[9] Based on prior literature, we view a restatement of audited financial statements as a strong indicator that the audit of the original financial statements was of low quality (Palmrose and Scholz 2004; Kinney, Palmrose, and Scholz 2004).[10] Plumlee and Yohn (2010) classify the explanations for accounting restatements into four main groups: internal errors by the firm (57 percent), fraud (3 percent), transaction complexity (3 percent), and application of accounting standards with judgment (37 percent). As discussed by Francis et al. (2014), the external auditor bears some responsibility for allowing a firm to issue financial statements that are materially misstated because of any one of these four reasons. Ours is not the first study to test whether clients of the Big 4 are less prone to issue restatements. Lobo and Zhao (2013, Table 5, Panel B) find evidence consistent with clients of the Big 4 being significantly less likely to issue an accounting restatement. Bentley, Omer, and Sharp (2013); Carcello, Neal, Palmrose, and Scholz (2011); and Newton et al. (2013) all find that clients of the Big 4 are no less likely to restate their earnings than are other firms.[11] However, none of the four aforementioned studies controls for the endogeneity of auditor selection. Our first hypothesis, stated in alternative form, is then:

H1: Clients of Big 4 auditors have a lower likelihood of issuing an accounting restatement than clients of non-Big 4 auditors after controlling for the client's propensity to choose a Big 4 auditor.

A separate but related question is whether the Big 4 provide higher audit quality than Mid-tier auditors. Regulators have questioned whether the Big 4 actually provide higher audit quality than Mid-tier auditors. For example, Kayla Gillan, a member of the PCAOB, suggests that audit committees "consider the so-called second-tier audit firms. I dislike using that term because it implies that the firms are secondary in quality which I strongly believe is false." (Grant Thornton 2007). We therefore test the following hypothesis.

H2: Clients of the Big 4 have a lower likelihood of issuing an accounting restatement than clients of Mid-tier auditors after controlling for the client's propensity to choose a Big 4 auditor.

We consider the two audit firms Grant Thornton and BDO Seidman to be the Mid-tier auditors, following Boone et al. (2010).

---

[8] If investors view the financial statements as being more credible due to the higher quality audit, then they will assign the firm a lower cost of equity capital. However, investors do not actually know the quality of the audit, since audit quality is unobservable. This is similar to using earnings response coefficients as a proxy for audit quality.

[9] The Audit Analytics Restatement database includes some restatements that are not due to failures in the application of GAAP. Of the 12,883 restatements available on Audit Analytics at the time of this study, 543 were not due to GAAP failure. Our results remain unchanged if we include these restatements in our analysis.

[10] Dechow, Ge, and Schrand (2010) note that "a significant benefit of using the restatement sample to identify firms with earnings quality problems is a lower Type I error rate in the identification of misstatements" (Dechow et al. 2010, 374). Restatements have been used as measures of audit quality in recent literature, such as Blankley, Hurtt, and MacGregor (2012); Francis and Michas (2013); and Francis, Michas, and Yu (2014).

[11] Many recent studies on accounting restatements restrict their sample to Big 4 clients only (e.g., Blankley et al. 2012; Francis and Michas 2013).

American Accounting Association

Downloaded from http://meridian.allenpress.com/ajpt/article-pdf/33/4/197/1603079/ajpt-50792.pdf by Case Western Reserve University, Charles Gletsu on 23 February 2021

RESEARCH DESIGN

To test our hypotheses, we first control for the endogenous choice of auditor. We follow Lawrence et al. (2011) and Boone et al. (2010) and use a propensity-score matching approach to match each non-Big 4 client with a Big 4 client on the basis of observable firm characteristics. Note that the propensity-score matching approach does not control for unobservable differences between Big 4 and non-Big 4 clients. To control for unobservable differences, one would use the Heckman (1979) self-selection model (Lennox, Francis, and Wang 2012). We prefer the propensity-score matching approach because it does not require exclusion restrictions. In our setting, an exclusion restriction is a variable or set of variables highly correlated with *BIG4* but uncorrelated with audit quality. As noted by Lennox and Pittman (2010, 236):

> In the context of auditor choice, a researcher who wishes to use the Heckman model faces the often intractable task of identifying an independent variable that meets the following conditions: (a) it is exogenous, (b) it is a very powerful predictor of auditor choice in the first stage model, and (c) it does not affect the dependent variable in the second stage model.

If one does not have such a variable, then the results from the Heckman model can be extremely sensitive to minor changes in the specification of the model (Lennox et al. 2012). We therefore model the choice of auditor using the following logit regression, estimated separately for each year:

$$BIG4_{it} = \beta_0 + \beta_1 LNASSETS_{it} + \beta_2 ATURN_{it} + \beta_3 ROA_{it} + \beta_4 LEV_{it} + \beta_5 CURR_{it} + \sum CONTROLS_{it} + \varepsilon_{it}. \quad (1)$$

See Table 1 for variable definitions. Subscripts i and t indicate firm and year, respectively. Each independent variable is winsorized at the first and 99th percentile before estimating this regression.

We use a logit regression to estimate this model because it is the most often used approach for estimating propensity scores (Guo and Fraser 2010, 135). After obtaining the fitted values from estimating Equation (1), we match each non-Big 4 client, with replacement, to the Big 4 client with the closest fitted value in the same year and same two-digit SIC code industry, requiring a maximum distance of 0.01 between the two fitted values. Requiring this maximum distance results in a loss of observations when no good match exists but it ensures that we find a close match.[12] We then test our hypotheses on the resulting matched-pairs sample. In testing the second hypothesis, we first drop all small auditor clients and then estimate Equation (1) on the smaller sample to obtain a matched sample of Big 4 and Mid-tier clients. The key matching variable in the model is firm size, which we proxy for using the log of total assets (*LNASSETS*) and the sales of the firm (*ATURN*). Larger firms tend to select Big 4 auditors and the Big 4 auditors tend to prefer larger clients. In addition, Big 4 auditors will prefer less risky clients. Therefore, we include the current ratio (*CURR*) and firm leverage (*LEV*) to control for the financial distress of the client. We follow Lawrence et al. (2011) and include the client's return on assets (*ROA*). Finally, consistent with Lawrence et al. (2011), we include all control variables used in the respective audit-quality analysis (*CONTROLS*). The propensity-score matching procedure attempts to minimize the variation in the firm characteristics, such as size and risk. In effect, the procedure is attempting to create a "pseudo random" sample in which one group of firms (i.e., the Big 4 clients) is given the "treatment," while the other group (the non-Big 4 clients) is not given the treatment.

[12] Results in the paper are robust to imposing a maximum distance of 0.02, 0.03, 0.05, 0.10, 0.20, or 0.25 between the two fitted values. Since matching with replacement results in duplicate control firm observations, we replicate our main result after deleting these duplicate observations. Our inferences remain unchanged.

Downloaded from http://meridian.allenpress.com/ajpt/article-pdf/33/4/197/1603079/ajpt-50792.pdf by Case Western Reserve University, Charles Gletsu on 23 February 2021

TABLE 1

Variable Definitions
(in the order they appear)

Variable		Definition
Variables Used in Restatement Analysis		
RESTATE	=	1 if the firm subsequently issues an accounting restatement, 0 otherwise.
BIG4	=	1 if the firm has a Big 4 auditor, 0 otherwise. For years before 2002, the Big 4 is actually the Big 5 as it includes Arthur Andersen.
LNASSETS	=	The natural logarithm of total assets (AT).
ATURN	=	Total sales (SALE) divided by lagged assets (AT).
ROA	=	Income before extraordinary items (IB), divided by average total assets (AT).
LEV	=	Financial leverage, defined as long-term debt plus debt in current liabilities, all scaled by total assets ([DLTT + DLC]/AT).
CURR	=	The current ratio, defined as current assets divided by current liabilities (ACT/LCT).
BM	=	The book-to-market ratio, defined as book equity scaled by market value at fiscal year-end (CEQ/[PRCC_F × CSHO]).
FIN	=	Financing raised, defined as additional cash raised from the issuance of long-term debt (DLTIS), plus cash raised from common and preferred stock (SSTK), all scaled by total assets (AT).
EPSGROW	=	1 if the firm had positive earnings (IBQ) changes for four consecutive quarters, 0 otherwise.
EP	=	The earnings-to-price ratio, defined as income before extraordinary items, scaled by market value at fiscal year-end (IB/[PRCC_F × CSHO]).
FREEC	=	Demand for external financing, defined as operating cash flows (OANCF) less capital expenditures (CAPX), all scaled by lagged assets.
AGE	=	The natural logarithm of the number of years the firm has been listed on Compustat.
QUAL	=	1 if the firm receives a qualified audit opinion, 0 otherwise. This variable equals 1 if the Compustat variable AUOP equals 2, 4, or 5.
INFLUENCE	=	The total fees (TOTAL_FEES) received from client *i* in year *t*, divided by the total fees earned by the audit office in year *t*. Audit offices are defined at the MSA level.
LNFEE	=	The natural logarithm of audit fees (AUDIT_FEES) charged to client *i* in year *t*.
Variables Used in Replication of Lawrence et al. (2011)		
|*DACC*|	=	The absolute value of discretionary accruals estimated as the residual from the Jones (1991) model, augmented with return on assets as suggested by Kothari, Leone, and Wasley (2005), estimated by industry-year, where industries are defined using two-digit SIC codes. We require at least 20 firms per industry-year to estimate this variable. We use all firms with sufficient data on Compustat to estimate this variable.
RPEG	=	The *ex ante* cost of equity capital, calculated as in Easton (2004) and Lawrence et al. (2011).
ACCY	=	Analyst forecast accuracy, calculated as −1 times the absolute value of the consensus analyst forecast error. The consensus analyst forecast error is the median consensus EPS estimate less the actual earnings, all scaled by stock price at fiscal year-end (PRCC_F).
LOGMV	=	The natural log of firm's market value (PRCC_F × CSHO).

(continued on next page)

American Accounting Association

Downloaded from http://meridian.allenpress.com/ajpt/article-pdf/33/4/197/1603079/ajpt-50792.pdf by Case Western Reserve University, Charles Gletsu on 23 February 2021

TABLE 1 (continued)

Variable		Definition
SURP	=	Earnings surprise, calculated as change in net income (NI) scaled by the firm's market value.
LOSS	=	1 if the firm reports negative net income (NI), 0 otherwise.
ZSCORE	=	Altman's (1968) bankruptcy score, calculated as $e^X/(1 + e^X)$, where $X = -4.34 - 0.08 \times (\text{WCAP/AT}) + 0.04 \times (\text{RE/AT}) - 0.10 \times (\text{PI} + \text{XINT} - \text{IDIT})/\text{AT} - 0.22 \times (\text{PRCCF} \times \text{CSHO})/\text{LT} + 0.06 \times (\text{SALE/AT})$.
HORIZON	=	The average number of days between the forecast announcement date and the subsequent earnings announcement date.
STDROE	=	Standard deviation of net income over past 5 years.
NUMANAL	=	The number of analysts following the firm.
EL	=	Actual earnings per share during t+1.
BETA	=	Market beta estimated using daily data for the year ending on the firm's fiscal year-end.
LOGLEV	=	The natural log of the firm's leverage ([DLTT + DLC]/AT).
LOGBM	=	The natural log of the firm's book-to-market ratio (CEQ/[PRCC_F × CSHO]).
DISP	=	The standard deviation of analysts' EPS forecasts during the fiscal year-end month.
GROWTH	=	Forecasted growth, calculated as the difference between the mean analysts' two-year- and one-year-ahead earnings forecasts scaled by the one-year-ahead earning forecast.

To test the first hypothesis, we estimate the following probit model on the matched sample with standard errors clustered by firm and year (Peterson 2009):

$$\begin{aligned} RESTATE_{it} = {} & \alpha_0 + \alpha_1 BIG4_{it} + \alpha_2 LNASSETS_{it} + \alpha_3 ATURN_{it} + \alpha_4 ROA_{it} + \alpha_5 LEV_{it} \\ & + \alpha_6 CURR_{it} + \alpha_7 BM_{it} + \alpha_8 FIN_{it} + \alpha_9 EPSGROW_{it} + \alpha_{10} EP_{it} + \alpha_{11} FREEC_{it} \\ & + \alpha_{12} AGE_{it} + \alpha_{13} QUAL_{it} + \alpha_{14} INFLUENCE_{it} + \alpha_{15} LNFEE_{it} \\ & + Industry\ Fixed\ Effects + Year\ Fixed\ Effects + \varepsilon_{it}. \end{aligned} \tag{2}$$

Subscripts i and t indicate firm and year, respectively. *RESTATE* equals 1 if the client's financial statements are later restated, 0 otherwise. We choose to focus on the original period as this is the period in which the auditor overlooked misstatements, indicating lower audit quality.[13] If Big 4 auditors do perform higher quality audits, then we expect to observe a negative coefficient on *BIG4*. We control for the size of the firm (*LNASSETS* and *ATURN*), profitability (*ROA*), the financial health of the firm (*LEV* and *CURR*), the book-to-market ratio (*BM*), whether the firm raised external financing (*FIN*), capital market pressures (*EPSGROW*), the earnings-to-price ratio (*EP*), demand for external financing (*FREEC*), firm age (*AGE*), whether the firm received a qualified audit opinion (*QUAL*), and the relative importance of the client to the auditor (*INFLUENCE*). Finally, we control for the amount of audit fees charged to the client (*LNFEE*), since clients receiving higher audit fees are less likely to subsequently issue a restatement (Blankley et al. 2012; Lobo and Zhao 2013).

[13] In other words, if the client's fiscal year-end falls between the beginning of the restated period and the end of the restated period, then it is classified as a restatement year. As noted by Newton et al. (2013), focusing on the period in which the restatement is announced likely identifies client-years in which the auditor discovered misstatements, indicating higher audit quality.

Downloaded from http://meridian.allenpress.com/ajpt/article-pdf/33/4/197/1603079/ajpt-50792.pdf by Case Western Reserve University, Charles Gletsu on 23 February 2021

TABLE 2

Sample Selection

Panel A: Big 4 versus Non-Big 4 Sample Selection

All Firm-Years on Compustat with Non-Missing CIK and SIC Code During 2000–2009:	90,634
Less: Financial Services Firms (SIC codes 6000–6999)	(14,832)
Less: Firm-years with material weaknesses	(1,599)
Less: Firm-years with auditor switches	(6,395)
Less: firm-years with missing data necessary to calculate control variables	(30,320)
Big 4 versus Non-Big 4 Full Sample:	37,488
Less: observations lost due to limited number of matches	(19,944)
Less: observations lost due to matched firm being too different	(11,594)
Big 4 versus Non-Big 4 Matched Sample:	5,950

Panel B: Big 4 versus Mid-Tier Sample Selection

All Firm-Years on Compustat with non-missing CIK and SIC code during 2000–2009:	56,966
Less: Financial Services Firms (SIC codes 6000–6999)	(8,757)
Less: Firm-years with material weaknesses	(1,368)
Less: Firm-years with auditor switches	(3,125)
Less: firm-years with missing data necessary to calculate control variables	(12,370)
Big 4 versus Mid-tier Full Sample:	31,346
Less: observations lost due to limited number of matches	(26,086)
Less: observations lost due to matched firm being too different	(2,012)
Big 4 versus Mid-tier Matched Sample:	3,248

To test whether Big 4 auditors perform higher quality audits than Mid-tier auditors (H2), we first re-estimate Equation (1) using a subsample of client-years, in which the client's auditor was either a Big 4 or a Mid-tier auditor, in order to create a matched sample. We then re-estimate Equation (2) on this sample.

SAMPLE SELECTION

We obtain financial statement data from the Compustat Fundamentals Annual file and auditor and restatement data from Audit Analytics for the period 2000–2009. The sample selection differs for each hypothesis. Table 2 outlines our sample selection procedure.

To construct our sample for H1, we begin with all client-year observations on Compustat with non-missing data on Audit Analytics and non-missing SIC code for the period 2000–2009. We then delete financial services clients (SIC codes 6000–6999), clients with material weaknesses over financial reporting, and client-years involving an auditor switch. Finally, we delete observations with insufficient data to calculate the variables used in Equation (2). This leaves us with a primary sample of 37,488 client-year observations (hereafter, the "Big 4 versus non-Big 4 full sample"). We perform the propensity-score matching regression on this sample. We lose 19,944 observations due to a limited number of matches. After imposing the requirement that the two firms' fitted values do

American Accounting Association

Downloaded from http://meridian.allenpress.com/ajpt/article-pdf/33/4/197/1603079/ajpt-50792.pdf by Case Western Reserve University, Charles Gletsu on 23 February 2021

not differ by more than 0.01, we are left with 2,975 Big 4 clients and 2,975 non-Big 4 clients, giving us a total sample of 5,950 observations (hereafter, the "Big 4 versus non-Big 4 matched sample"). We deliberately end the sample two years before the most recent year of data available when we began the study (2011), to allow for a sufficient amount of time for a firm to subsequently restate its earnings.[14] We leave two years for clients to restate earnings because Cheffers, Whalen, and Usvyatsky (2010) show that the average time between an originally released financial statement and a restatement is approximately 700 days, which is slightly less than two years.[15]

We construct our sample for H2 in a similar manner. The only difference is that we delete all client-year observations in which the client is audited by a small auditor. This results in a primary sample of 31,346 client-year observations (hereafter, the "Big 4 versus Mid-tier full sample"). After performing the propensity-score matching regression (Equation (1)) and deleting matched pairs in which the fitted values differ by more than 0.01, we are left with 1,624 Big 4 client-years and 1,624 Mid-tier client-years, giving us a sample of 3,248 observations (hereafter, the "Big 4 versus Mid-tier matched sample"). In both samples, all continuous independent variables are winsorized at the first and 99th percentiles. We do not tabulate the results from estimating the propensity-score matching regression (Equation (1)) on each of our samples to save space.[16]

REPLICATION OF LAWRENCE ET AL. (2011)

In this section we replicate the results of Lawrence et al. (2011) using the full and matched samples to ensure that our results are not due to changes in audit quality during different sample periods.[17] Lawrence et al. (2011) test whether Big 4 auditors deliver higher audit quality, using the following three linear regression models:

$$|DACC_{it}| = \rho_0 + \rho_1 BIG4_{it} + \rho_2 LOGMV_{it} + \rho_3 ROA_{it} + \rho_4 LEV_{it} + \rho_5 CURR_{it} + Industry\ Fixed\ Effects + Year\ Fixed\ Effects + \varepsilon_{it}. \quad (3)$$

$$ACCY_{it} = \omega_0 + \omega_1 BIG4_{it} + \omega_2 LOGMV_{it} + \omega_3 SURP_{it} + \omega_4 LOSS_{it} + \omega_5 ZSCORE_{it} + \omega_6 HORIZON_{it} + \omega_7 STDROE_{it} + \omega_8 NUMANAL_{it} + \omega_9 EL_{it} + Industry\ Fixed\ Effects + Year\ Fixed\ Effects + \varepsilon_{it}. \quad (4)$$

$$RPEG_{it} = \varphi_0 + \varphi_1 BIG4_{it} + \varphi_2 BETA_{it} + \varphi_3 LOGLEV_{it} + \varphi_4 DISP_{it} + \varphi_5 LOGMV_{it} + \varphi_6 LOGBM_{it} + \varphi_7 GROWTH_{it} + Industry\ Fixed\ Effects + Year\ Fixed\ Effects + \varepsilon_{it}. \quad (5)$$

In all models, subscripts i and t denote firm and year, respectively. See Table 1 for variable definitions. Each model includes industry- and year-fixed effects. In Equation (3), audit quality is measured using the absolute value of discretionary accruals ($|DACC|$). In Equation (4), audit

[14] For example, if we used all observations up to the most recent year available on Audit Analytics (2011 at the time of this study), then we would not know whether a firm's 2010 earnings would be restated in later years, such as 2012.

[15] For the Big 4 versus non-Big 4 full sample, there are 1,791 restatements, all of which are attributable to a failure in the application of GAAP. Of these restatements, 1,530 (85.4 percent) resulted in a downward adjustment to income, 73 (4.1 percent) are attributable to financial fraud, 20 (1.1 percent) are attributable to clerical errors, and 183 (10.2 percent) are attributable to other significant issues. Certain restatements are attributable to more than one cause.

[16] We note that the R^2 of the propensity-score matching regression is 44.8 percent for the Big 4 versus non-Big 4 sample and 14.8 percent for the Big 4 versus Mid-tier sample. The lower explanatory power when using the Big 4 versus Mid-tier sample is likely due to the matching model R^2 being an increasing function of the heterogeneity of the two groups before the matching procedure (Peel and Makepeace 2012, 617).

[17] Lawrence et al.'s (2011) sample spans 1988–2006, while our sample covers 2000–2009.

Downloaded from http://meridian.allenpress.com/ajpt/article-pdf/33/4/197/1603079/ajpt-50792.pdf by Case Western Reserve University, Charles Gletsu on 23 February 2021

quality is measured using analyst forecast accuracy (*ACCY*). In Equation (5), audit quality is measured using the firm's cost of capital, estimated as in Easton (2004).

Table 3 reports the results of estimating Equations (3), (4), and (5) on the full and matched Big 4 versus non-Big 4 samples.[18] Panel A reports the discretionary accruals results. Using the full sample, the coefficient on *BIG4* is negative and significant. Consistent with Lawrence et al. (2011), the coefficient becomes insignificant when using the propensity-score matched sample (Coeff. = 0.058, p-value = 0.552). The goodness of fit of the model is consistent with Lawrence et al. (2011); when using the full sample, the R^2 is 0.199 compared to an R^2 of 0.14 reported in Lawrence et al. (2011). Panel B reports the analyst forecast accuracy results. Using the full sample, the coefficient on *BIG4* is significantly positive, consistent with clients of the Big 4 having higher analyst forecast accuracy. However, consistent with Lawrence et al. (2011), this coefficient becomes insignificant after matching (Coeff. = 0.004, p-value = 0.606). When using the full sample, the R^2 is 0.219, compared with an R^2 of 0.28 reported by Lawrence et al. (2011). Panel C reports the cost of capital results. Here, the coefficient on *BIG4* is significantly negative when using the full sample but becomes insignificant when using the matched sample (Coeff. = 0.001, p-value = 0.837), consistent with Lawrence et al. (2011). In sum, the results of Lawrence et al. (2011) continue to hold for our sample period. This means that any divergent results we observe when using restatements as a proxy for audit quality are likely due to the choice of audit quality proxy.

There is reason to believe that the proxies used by Lawrence et al. (2011) may not be capturing audit quality. For example, using the magnitude of discretionary accruals as a proxy for audit quality assumes all discretionary accruals, regardless of whether they are positive or negative, are equally harmful to earnings quality. However, extant research provides evidence that this is not the case. Krishnan (2003) finds that the association between discretionary accruals and future profitability is stronger for Big 4 clients compared with non-Big 4 clients. Therefore, while clients of Big 4 auditors do not exhibit significantly different discretionary accruals than clients of smaller auditors, this does not imply that the two sets of firms have similar earnings quality. Clients of the Big 4 could be using their discretionary accruals to signal future profitability.[19] The use of financial analyst forecast accuracy as a proxy for audit quality relies on the assumption that auditors affect earnings quality, which affects analyst forecast accuracy. This is a very indirect measure of audit quality; it would seem that examining an earnings quality metric would be more informative about the quality of the audit. Finally, the third proxy for audit quality chosen by Lawrence et al. (2011) is the *ex ante* cost of equity capital. The argument is that firms with better audit quality will be seen as having more credible financial statements. This credibility lowers information risk, which will lower the firm's cost of capital. However, the *ex ante* cost of equity capital only speaks to investors' perceptions of audit quality. For this measure to capture audit quality, we must assume that investors can observe audit quality and can price it efficiently.

EMPIRICAL RESULTS FOR BIG 4 VERSUS NON-BIG 4

In this section, we examine whether clients of Big 4 auditors are less likely to subsequently issue accounting restatements. Table 4 contains descriptive statistics for the full and matched Big 4 versus non-Big 4 samples. Panel A reports the descriptive statistics for the full sample of Big 4 and

[18] The number of observations differs for each test, due to missing variables.

[19] Another problem with this proxy is that it assumes that the goal of management is to book the largest discretionary accruals in any given period. This is not always the case. Consider a manager of a firm with pre-discretionary earnings (i.e., earnings before discretionary accruals) of $2.00 per share and an earnings target of $1.97 per share. In this case, the manager has no incentive to record large discretionary accruals. Finally, using the absolute value of discretionary accruals to measure audit quality assumes that the discretionary accrual model accurately dissects a firm's accruals into a normal and a discretionary component.

American Accounting Association

Downloaded from http://meridian.allenpress.com/ajpt/article-pdf/33/4/197/1603079/ajpt-50792.pdf by Case Western Reserve University, Charles Gletsu on 23 February 2021

TABLE 3

Replication of Lawrence et al. (2011)

Panel A: Discretionary Accruals Analysis

	Full Sample		Matched Sample	
Variable	**Estimate**	**p-value**	**Estimate**	**p-value**
Intercept	0.716***	0.000	1.371**	0.016
BIG4	−0.228***	0.000	0.058	0.552
LOGMV	−0.026*	0.058	−0.132*	0.090
ROA	−0.397***	0.008	−0.118	0.588
LEV	−0.050	0.620	0.200	0.562
CURR	−0.020***	0.000	−0.028**	0.039
Industry and Year FE	Yes		Yes	
n	36,153		5,907	
Adj. R^2	0.199		0.193	

Panel B: Analyst Forecast Accuracy Analysis

	Full Sample		Matched Sample	
Variable	**Estimate**	**p-value**	**Estimate**	**p-value**
Intercept	−0.047***	0.000	−0.050**	0.050
BIG4	0.005**	0.021	0.004	0.606
LOGMV	0.009***	0.000	0.020***	0.000
SURP	0.000	0.810	−0.003***	0.000
LOSS	−0.019***	0.000	−0.011*	0.077
ZSCORE	−3.047***	0.000	−4.240***	0.000
HORIZON	0.000***	0.000	−0.001***	0.000
STDROE	0.000***	0.000	0.000**	0.016
NUMANAL	0.000**	0.045	−0.003**	0.015
EL	0.005**	0.014	0.012***	0.002
Industry and Year FE	Yes		Yes	
n	19,669		2,228	
Adj. R^2	0.219		0.400	

Panel C: Cost of Capital Analysis

	Full Sample		Matched Sample	
Variable	**Estimate**	**p-value**	**Estimate**	**p-value**
Intercept	0.107***	0.000	0.168***	0.000
BIG4	−0.003*	0.064	0.001	0.837
BETA	−0.001	0.706	0.005	0.173
LOGLEV	0.002***	0.000	0.000	0.529
DISP	0.034***	0.001	0.175***	0.000
LOGMV	−0.012***	0.000	−0.021***	0.000
LOGBM	0.007***	0.000	0.006***	0.002
GROWTH	−0.005***	0.000	−0.011***	0.002

(continued on next page)

Downloaded from http://meridian.allenpress.com/ajpt/article-pdf/33/4/197/1603079/ajpt-50792.pdf by Case Western Reserve University, Charles Gletsu on 23 February 2021

TABLE 3 (continued)

	Full Sample		Matched Sample	
Variable	**Estimate**	**p-value**	**Estimate**	**p-value**
Industry and Year FE	Yes		Yes	
n	7,975		680	
Adj. R^2	0.463		0.513	

*, **, *** Denote statistical significance at the 10, 5, and 1 percent levels, respectively, using a two-tailed test.
This table displays the results of attempting to replicate the results documented in Lawrence et al. (2011). In Panel A, the dependent variable is the absolute value of discretionary accruals (|*DACC*|). In Panel B, the dependent variable is analyst forecast accuracy (*ACCY*). In Panel C, the dependent variable is the firm's *ex ante* cost of equity capital (*RPEG*). All models include industry- and year-fixed effects. p-values are based on t-statistics, which are clustered by firm and year (Peterson 2009).

non-Big 4 clients. Surprisingly, clients of the Big 4 have a higher frequency of restatements (9.3 percent compared to 6.7 percent for the non-Big 4). However, we caution against putting too much weight on this result, as we have not controlled for other factors that influence the likelihood of issuing a restatement. Compared with clients of non-Big 4 auditors, clients of the Big 4 have significantly more assets (*LNASSETS*), a significantly higher return on assets (*ROA*), and are significantly older (*AGE*). In addition, clients of the Big 4 have a lower absolute value of discretionary accruals (|*DACC*|), a lower cost of equity (*RPEG*), and greater analyst forecast accuracy (*ACCY*). Note that these three variables were not used in the propensity-score matching regression. We only report the difference in means of these variables for comparison with prior literature (Lawrence et al. 2011). The last columns of Table 4 report a parametric t-test of the difference in means and a non-parametric Kolmogorov-Smirnov (KS) test. The KS test is a non-parametric test, which tests the difference in distributions between the two groups of firms. Differences in means for indicator variables are tested using Chi-square (χ^2) tests. The difference in means and difference in distributions of each variable are significant at the 1 percent level, highlighting the importance of controlling for endogeneity.

Panel B of Table 4 reports descriptive statistics for the Big 4 versus non-Big 4 matched sample. The important thing to note is that of the 14 control variables, only three (six) have means (distributions) that are significantly different at the 5 percent level across the two groups of auditors. Further, the magnitudes of the differences are not overly large. For example, clients of the Big 4 have an average current ratio of 3.533 compared to 3.837 for clients of non-Big 4 auditors. Finally, it is worth noting that the two groups of clients do not exhibit significantly different means or distributions of the three audit quality proxies used by Lawrence et al. (2011) (|*DACC*|, *RPEG,* and *ACCY*). This is important, since, if our matching procedure produced a sample in which Big 4 clients exhibited a lower magnitude of discretionary accruals, a lower cost of equity capital, and/or higher analyst forecast accuracy, then we could not attribute our restatement results to our choice of audit quality proxy. The results would likely be attributable to our time period and/or our matching model.[20]

Table 5 reports the results of estimating the restatement model (Equation (2)). The first set of columns displays results when using the full sample. The coefficient on *BIG4* is −0.091, significant at the 5 percent level. This implies that clients of Big 4 auditors are less likely to subsequently issue an accounting restatement. The second set of columns reports the results when using the matched

[20] We do not report correlations to save space, but we note that none of the variance inflation factors on any of the variables exceeds 5, which is well below the threshold of 10 recommended by Kennedy (1992).

American Accounting Association

Downloaded from http://meridian.allenpress.com/ajpt/article-pdf/33/4/197/1603079/ajpt-50792.pdf by Case Western Reserve University, Charles Gletsu on 23 February 2021

TABLE 4

Descriptive Statistics for Big 4 versus Non-Big 4 Samples

Panel A: Descriptive Statistics for Full Sample

	Big 4 Clients (n = 28,716)				Non-Big 4 Clients (n = 8,772)				t-test Diff.	KS Diff.	χ^2
Variable	**Mean**	**Q1**	**Med.**	**Q3**	**Mean**	**Q1**	**Med.**	**Q3**	**p-value**	**p-value**	**p-value**
RESTATE	0.093	0.000	0.000	0.000	0.067	0.000	0.000	0.000	—	—	0.000
LNASSETS	6.230	4.747	6.157	7.632	3.396	2.225	3.299	4.474	0.000	0.000	
ATURN	1.095	0.497	0.900	1.432	1.236	0.422	1.008	1.683	0.000	0.000	
ROA	−0.037	−0.050	0.031	0.077	−0.166	−0.238	−0.015	0.063	0.000	0.000	
LEV	0.195	0.008	0.163	0.325	0.159	0.000	0.089	0.271	0.000	0.000	
CURR	3.102	1.313	2.036	3.397	3.823	1.229	2.144	4.020	0.000	0.000	
BM	0.665	0.275	0.476	0.779	0.767	0.235	0.508	0.943	0.000	0.000	
FIN	0.156	0.007	0.040	0.181	0.213	0.001	0.034	0.263	0.000	0.000	
EPSGROW	0.125	0.000	0.000	0.000	0.085	0.000	0.000	0.000	—	—	0.000
EP	−0.100	−0.052	0.032	0.061	−0.175	−0.172	−0.012	0.056	0.000	0.000	
FREEC	−0.026	−0.049	0.030	0.089	−0.175	−0.209	−0.027	0.066	0.000	0.000	
AGE	2.533	1.946	2.485	3.135	2.402	1.792	2.485	3.045	0.000	0.000	
QUAL	0.453	0.000	0.000	1.000	0.338	0.000	0.000	1.000	—	—	0.000
INFLUENCE	0.053	0.002	0.010	0.038	0.214	0.036	0.090	0.248	0.000	0.000	
LNFEE	13.377	12.367	13.349	14.274	11.776	11.002	11.704	12.464	0.000	0.000	
|DACC|	0.787	0.047	0.138	0.500	1.275	0.071	0.206	0.834	0.000	0.000	
RPEG	0.135	0.073	0.099	0.152	0.183	0.092	0.138	0.226	0.000	0.000	
ACCY	−0.018	−0.007	−0.002	−0.001	−0.041	−0.021	−0.005	−0.001	0.000	0.000	

Panel B: Descriptive Statistics for Matched Sample and Test of Covariate Balance between Matched Pairs (n = 2,975 pairs)

	Big 4 Clients (n = 2,975)				Non-Big 4 Clients (n = 2,975)				t-test Diff.	KS Diff.	χ^2
Variable	**Mean**	**Q1**	**Med.**	**Q3**	**Mean**	**Q1**	**Med.**	**Q3**	**p-value**	**p-value**	**p-value**
RESTATE	0.056	0.000	0.000	0.000	0.078	0.000	0.000	0.000	—	—	0.001
LNASSETS	4.055	2.860	3.871	5.101	4.040	2.813	3.941	5.151	0.743	0.370	
ATURN	1.117	0.459	0.919	1.493	1.083	0.484	0.945	1.430	0.159	0.157	
ROA	−0.137	−0.209	−0.010	0.063	−0.157	−0.247	−0.012	0.068	0.113	0.000	
LEV	0.129	0.000	0.033	0.230	0.136	0.000	0.063	0.230	0.037	0.109	
CURR	3.533	1.520	2.492	4.132	3.837	1.379	2.322	4.216	0.005	0.007	
BM	0.687	0.239	0.477	0.830	0.693	0.243	0.489	0.868	0.749	0.286	
FIN	0.185	0.003	0.026	0.217	0.201	0.004	0.035	0.259	0.058	0.004	
EPSGROW	0.121	0.000	0.000	0.000	0.106	0.000	0.000	0.000	—	—	0.065
EP	−0.190	−0.169	−0.006	0.049	−0.183	−0.168	−0.010	0.052	0.640	0.348	
FREEC	−0.111	−0.175	−0.015	0.073	−0.141	−0.189	−0.010	0.076	0.004	0.316	
AGE	2.400	1.946	2.398	2.890	2.417	1.946	2.485	2.996	0.368	0.000	
QUAL	0.384	0.000	0.000	1.000	0.384	0.000	0.000	1.000	—	—	1.000
INFLUENCE	0.134	0.000	0.004	0.035	0.145	0.030	0.066	0.150	0.081	0.000	
LNFEE	12.245	11.408	12.129	13.028	12.294	11.488	12.245	13.097	0.100	0.000	

(continued on next page)

Downloaded from http://meridian.allenpress.com/ajpt/article-pdf/33/4/197/1603079/ajpt-50792.pdf by Case Western Reserve University, Charles Gletsu on 23 February 2021

TABLE 4 (continued)

	Big 4 Clients (n = 2,975)				Non-Big 4 Clients (n = 2,975)				t-test Diff.	KS Diff.	χ^2
Variable	**Mean**	**Q1**	**Med.**	**Q3**	**Mean**	**Q1**	**Med.**	**Q3**	**p-value**	**p-value**	**p-value**
|DACC|	1.335	0.086	0.274	1.160	1.269	0.088	0.275	1.124	0.352	0.492	
RPEG	0.167	0.084	0.123	0.192	0.175	0.087	0.127	0.210	0.284	0.235	
ACCY	−0.036	−0.012	−0.004	−0.001	−0.031	−0.014	−0.004	−0.001	0.209	0.386	

This table displays descriptive statistics for the Big 4 versus non-Big 4 samples. Panel A reports descriptive statistics for the full sample while Panel B reports descriptive statistics for the matched sample. In each Panel, the mean, first quartile (Q1), median, and third quartile (Q3) are displayed. For each variable, the last two columns report on the difference in the mean and the overall distribution of values for that variable among Big 4 and non-Big 4 clients. The t-test difference p-value column reports the p-value from a test of difference in means across the Big 4 and non-Big 4 clients. The KS difference p-value is the p-value from the Kolmogorov-Smirnov test of difference in distributions across the two sets of clients (Big 4 versus non-Big 4 clients). Chi-square (χ^2) tests are used to test the difference in distributions for indicator variables.
See Table 1 for variable definitions.

sample, which better controls for the endogenous choice of auditor. This is important, since it may be the case that clients with better financial reporting quality choose Big 4 auditors. The results are similar; the coefficient on *BIG4* is −0.201 and is significant at the 1 percent level. The marginal effect of *BIG4* is −2.21 percent, implying that using a Big 4 auditor results in a 2.21 percent decrease in the probability of issuing an accounting restatement.[21] At first glance, this may appear to be a rather small effect. However, it is quite large if one considers the baseline probability of issuing an accounting restatement is 5.3 percent for the matched sample and 7.5 percent for the full sample. The baseline probability is the probability that the dependent variable equals 1 if all independent variables are held at their mean values. To summarize, the results of Table 5 offer strong support for H1.[22]

EMPIRICAL RESULTS FOR BIG 4 VERSUS MID-TIER

In the previous section, we demonstrated that clients of the Big 4 are less likely to restate their earnings than are clients of other auditors, even after controlling for the choice of auditor. In this section, we examine whether this result can be explained by the inclusion of small auditors in the sample. Specifically, we examine whether clients of Mid-tier auditors are more likely to restate their earnings than are clients of the Big 4.

Panel A of Table 6 reports descriptive statistics for the full sample of Big 4 and Mid-tier clients. The results are somewhat similar to Panel A, with clients of the Big 4 generally having more assets (*LNASSETS*), greater profitability (*ROA*), less financial distress (*LEV, BM*), and being younger (*AGE*). In addition, clients of the Big 4 tend to exhibit a lower magnitude of discretionary accruals (*|DACC|*) and a lower cost of capital (*RPEG*). The difference in means and distributions for each of these variables is significant at the 10 percent level or better, with the exception of the level of financing raised (*FIN*).

[21] It is worth noting that the marginal effect of *BIG4* is much larger when using the matched sample. Given that the variance of the estimate of treatment effects is lower in matched samples (Rosenbaum and Rubin 1983, 48), we put more weight on the marginal effect of the treatment (i.e., *BIG4*) when using the matched sample.

[22] If we use matching without replacement and a 3 percent caliper distance, as in Lawrence et al. (2011), then the coefficient on *BIG4* is −0.205 with a Z-statistic of −1.14. If we use a 10 percent caliper, then the Z-statistic increases to −1.54, which is significant at the 10 percent level using a one-tailed test.

American Accounting Association

Downloaded from http://meridian.allenpress.com/ajpt/article-pdf/33/4/197/1603079/ajpt-50792.pdf by Case Western Reserve University, Charles Gletsu on 23 February 2021

TABLE 5

Do Clients of the Big 4 Have Fewer Restatements?

	Full Sample			Matched Sample		
Variable	**Estimate**	**Marginal Effect**	**p-value**	**Estimate**	**Marginal Effect**	**p-value**
Intercept	−1.714***		0.000	−2.956***		0.000
BIG4	−0.091**	−1.35%	0.021	−0.201***	−2.21%	0.007
LNASSETS	0.067***	0.95%	0.001	0.100***	1.10%	0.003
ATURN	0.036**	0.50%	0.014	0.099***	1.20%	0.005
ROA	0.072	−0.94%	0.412	−0.153	−1.93%	0.576
LEV	−0.066	1.01%	0.410	−0.165	−1.81%	0.377
CURR	−0.005	−0.07%	0.315	−0.002	−0.03%	0.824
BM	−0.039**	−0.55%	0.038	−0.109**	−1.16%	0.025
FIN	0.040	0.55%	0.457	0.157*	1.66%	0.079
EPSGROW	−0.008	−0.12%	0.851	−0.067	−0.65%	0.584
EP	0.040	0.56%	0.362	0.052	0.60%	0.641
FREEC	0.029	0.40%	0.650	0.060	0.64%	0.632
AGE	−0.028	−0.39%	0.168	−0.029	−0.28%	0.665
QUAL	0.082***	1.17%	0.000	0.070	0.89%	0.279
INFLUENCE	−0.034	−0.48%	0.727	0.078	0.89%	0.495
LNFEE	−0.021	−0.29%	0.385	0.014	0.23%	0.714
Industry	Yes			Yes		
Year FE	Yes			Yes		
n	37,488			5,950		
Pseudo R^2	0.064			0.081		

*, **, *** Denote statistical significance at the 10, 5, and 1 percent levels, respectively, using a two-tailed test.
This table displays estimated coefficients from estimating a probit model (Equation (2)) where the dependent variable (*RESTATE*) equals 1 if the client subsequently issues an accounting restatement, 0 otherwise. The model is estimated on both the full Big 4 versus non-Big 4 sample and the matched Big 4 versus non-Big 4 sample. All independent variables are winsorized at the 1st and 99th percentiles to reduce the influence of outliers. The model includes industry- and year-fixed effects, where industries are defined using two-digit SIC codes. p-values are based on Z-statistics, which are clustered by firm and year (Peterson 2009).

Panel B of Table 6 reports descriptive statistics for the matched sample of Big 4 and Mid-tier clients. Of the 14 control variables, only three (four) exhibit significantly different means (distributions) across the two groups of auditors. One variable, which is still significantly different across the two sets of firms, is the relative importance of the client to the audit office (*INFLUENCE*). The average client of the Big 4 represents only 5.5 percent of the office's total fees, compared with 8.4 percent for the average Mid-tier client. It is worth noting that the two sets of clients do not exhibit significantly different magnitudes of discretionary accruals or significantly different costs of capital. Finally, after matching, clients of the Big 4 have significantly higher analyst forecast accuracy than clients of Mid-tier auditors (−0.022 compared to −0.029).

Panel A of Table 7 reports the main results. The first set of columns contains the results when using the full sample. Here, we find that the coefficient on *BIG4* is not significantly different from zero (p-value = 0.584). The second set of columns reports results when using the matched sample. Here, we find a significantly negative coefficient on *BIG4* (−0.266, p-value = 0.006). This suggests

Downloaded from http://meridian.allenpress.com/ajpt/article-pdf/33/4/197/1603079/ajpt-50792.pdf by Case Western Reserve University, Charles Gletsu on 23 February 2021

TABLE 6

Descriptive Statistics for Big 4 versus Mid-Tier Samples

Panel A: Descriptive Statistics for Full Sample

	Big 4 Clients (n = 28,716)				Mid-Tier Clients (n = 2,630)				t-test Diff.	KS Diff.	χ^2
Variable	**Mean**	**Q1**	**Med.**	**Q3**	**Mean**	**Q1**	**Med.**	**Q3**	**p-value**	**p-value**	**p-value**
RESTATE	0.093	0.000	0.000	0.000	0.071	0.000	0.000	0.000	—	—	0.000
LNASSETS	6.236	4.747	6.157	7.632	4.346	3.167	4.268	5.353	0.000	0.000	
ATURN	1.091	0.497	0.900	1.432	1.304	0.673	1.138	1.754	0.000	0.000	
ROA	−0.034	−0.050	0.031	0.077	−0.072	−0.122	0.011	0.071	0.000	0.000	
LEV	0.195	0.008	0.163	0.325	0.162	0.000	0.095	0.284	0.000	0.001	
CURR	3.060	1.313	2.036	3.397	3.170	1.355	2.154	3.620	0.098	0.000	
BM	0.662	0.275	0.476	0.779	0.801	0.305	0.552	0.984	0.000	0.000	
FIN	0.154	0.007	0.040	0.181	0.171	0.003	0.025	0.200	0.002	0.157	
EPSGROW	0.125	0.000	0.000	0.000	0.102	0.000	0.000	0.000	—	—	0.001
EP	−0.099	−0.052	0.032	0.061	−0.146	−0.136	0.013	0.059	0.000	0.000	
FREEC	−0.020	−0.049	0.030	0.089	−0.041	−0.090	0.013	0.086	0.000	0.000	
AGE	2.533	1.946	2.485	3.135	2.579	2.079	2.639	3.091	0.007	0.000	
QUAL	0.453	0.000	0.000	1.000	0.388	0.000	0.000	1.000	—	—	0.000
INFLUENCE	0.051	0.002	0.010	0.038	0.104	0.022	0.051	0.123	0.000	0.000	
LNFEE	13.381	12.367	13.349	14.274	12.504	11.729	12.445	13.230	0.000	0.000	
|DACC|	0.760	0.047	0.138	0.500	0.959	0.059	0.165	0.646	0.000	0.000	
RPEG	0.135	0.073	0.099	0.152	0.169	0.088	0.129	0.205	0.000	0.000	
ACCY	−0.017	−0.007	−0.002	−0.001	−0.033	−0.016	−0.004	−0.001	0.000	0.000	

Panel B: Descriptive Statistics for Matched Sample and Test of Covariate Balance between Matched Pairs (n = 1,624 pairs)

	Big 4 Clients (n = 1,624)				Mid-Tier Clients (n = 1,624)				t-test Diff.	KS Diff.	χ^2
Variable	**Mean**	**Q1**	**Med.**	**Q3**	**Mean**	**Q1**	**Med.**	**Q3**	**p-value**	**p-value**	**p-value**
RESTATE	0.049	0.000	0.000	0.000	0.079	0.000	0.000	0.000	—	—	0.000
LNASSETS	4.506	3.300	4.322	5.574	4.653	3.475	4.604	5.693	0.010	0.000	
ATURN	1.177	0.565	0.981	1.571	1.146	0.576	1.021	1.540	0.314	0.353	
ROA	−0.088	−0.156	0.007	0.070	−0.083	−0.143	0.009	0.070	0.597	0.845	
LEV	0.148	0.000	0.072	0.263	0.141	0.000	0.067	0.242	0.213	0.400	
CURR	3.344	1.495	2.290	3.843	3.390	1.420	2.296	3.935	0.702	0.137	
BM	0.721	0.242	0.495	0.874	0.723	0.291	0.530	0.895	0.943	0.005	
FIN	0.174	0.004	0.032	0.221	0.168	0.004	0.026	0.205	0.554	0.561	
EPSGROW	0.105	0.000	0.000	0.000	0.118	0.000	0.000	0.000	—	—	0.264
EP	−0.143	−0.127	0.007	0.054	−0.148	−0.130	0.009	0.054	0.771	0.708	
FREEC	−0.054	−0.120	0.011	0.084	−0.055	−0.116	0.011	0.085	0.963	0.869	
AGE	2.469	1.946	2.398	2.917	2.497	2.079	2.485	2.996	0.271	0.002	
QUAL	0.419	0.000	0.000	1.000	0.407	0.000	0.000	1.000	—	—	0.498
INFLUENCE	0.055	0.001	0.006	0.023	0.084	0.024	0.050	0.107	0.000	0.000	
LNFEE	12.588	11.717	12.529	13.380	12.674	11.844	12.613	13.451	0.022	0.057	
|DACC|	1.147	0.075	0.204	0.931	1.073	0.074	0.219	0.906	0.368	0.586	

(continued on next page)

American Accounting Association

Downloaded from http://meridian.allenpress.com/ajpt/article-pdf/33/4/197/1603079/ajpt-50792.pdf by Case Western Reserve University, Charles Gletsu on 23 February 2021

TABLE 6 (continued)

	Big 4 Clients (n = 1,624)				Mid-Tier Clients (n = 1,624)				t-test Diff.	KS Diff.	χ^2
Variable	**Mean**	**Q1**	**Med.**	**Q3**	**Mean**	**Q1**	**Med.**	**Q3**	**p-value**	**p-value**	**p-value**
RPEG	0.168	0.086	0.119	0.210	0.159	0.086	0.122	0.195	0.328	0.783	
ACCY	−0.022	−0.012	−0.003	−0.001	−0.029	−0.011	−0.003	−0.001	0.076	0.644	

This table displays descriptive statistics for the Big 4 versus Mid-tier samples. Panel A reports descriptive statistics for the full sample while Panel B reports descriptive statistics for the matched sample. In each Panel, the mean, first quartile (Q1), median, and third quartile (Q3) are displayed. For each variable, the last two columns report on the difference in the mean and the overall distribution of values for that variable among Big 4 and Mid-tier clients. The t-test difference p-value column reports the p-value from a test of difference in means across the Big 4 and Mid-tier clients. The KS difference p-value is the p-value from the Kolmogorov-Smirnov test of difference in distributions across the two sets of clients (Big 4 versus Mid-tier clients). Chi-square (χ^2) tests are used to test the difference in distributions for indicator variables.
See Table 1 for variable definitions.

that, after controlling for the endogenous choice of auditor, clients of the Big 4 are less likely to subsequently restate their earnings than are clients of Mid-tier auditors. Overall, Table 7 provides weak evidence that Big 4 auditors provide higher audit quality than Mid-tier auditors.

Medium and Small Auditors

In this section, we extend our analysis to small auditors (i.e., auditors other than the Big 4, Grant Thornton, and BDO Seidman). We estimate a propensity-score matching regression similar to Equation (1), with the exception that the *BIG4* variable is replaced by a Mid-tier indicator variable, which equals 1 for Mid-tier auditors, 0 otherwise. The full sample contains 8,772 firm-years, while the matched sample contains 1,406 firm-years. Panel B of Table 7 reports the results. We do not tabulate control variables in order to save space. Using either the full or matched sample, the results are similar. The coefficient on the Mid-tier indicator variable is negative, but not significantly different from zero. This suggests that clients of small auditors are no more likely to restate their earnings than are clients of Mid-tier auditors.[23]

ADDITIONAL ANALYSIS

Alternative Measures of Misstatements

A key assumption in our analysis is that restatements are an adequate proxy for audit quality. To test the robustness of our results, we repeat our analysis using the existence of Accounting and Auditing Enforcement Releases (AAERs) as a proxy for audit quality, following Lennox and Pittman (2010). We also examine the sensitivity of our results when using restatements that require a downward adjustment to net income.

Panel A of Table 8 reports the results of estimating Equation (2) after replacing the dependent variable with *AAER,* which equals 1 if the firm is subsequently targeted by the SEC in an

[23] We have also tested whether the Big 4 provide higher audit quality than small auditors. Using either the full or a matched sample, we find that clients of the Big 4 are significantly less likely to issue an accounting restatement than are clients of small auditors.

Downloaded from http://meridian.allenpress.com/ajpt/article-pdf/33/4/197/1603079/ajpt-50792.pdf by Case Western Reserve University, Charles Gletsu on 23 February 2021

TABLE 7

Mid-Tier and Small Auditors

Panel A: Do Clients of the Big 4 Have Fewer Restatements than Clients of Mid-Tier Firms?

	Full Sample			Matched Sample		
Variable	**Estimate**	**Marginal Effect**	**p-value**	**Estimate**	**Marginal Effect**	**p-value**
Intercept	−1.951***		0.000	−5.095***		0.000
BIG4	−0.034	−0.50%	0.584	−0.266***	−2.50%	0.006
LNASSETS	0.042	0.61%	0.091	0.058	0.51%	0.236
ATURN	0.029	0.42%	0.155	0.077	0.69%	0.276
ROA	0.130	−2.62%	0.196	−0.455	−0.49%	0.232
LEV	−0.188	1.87%	0.164	−0.553*	−4.1%	0.060
CURR	−0.005	−0.07%	0.401	−0.014	−0.13%	0.555
BM	−0.020	−0.28%	0.398	0.089	0.79%	0.241
FIN	0.041	0.60%	0.556	0.313	2.80%	0.150
EPSGROW	−0.020	−0.29%	0.658	−0.327***	−2.40%	0.001
EP	0.076	1.09%	0.120	0.304**	2.72%	0.041
FREEC	0.175	2.53%	0.064	0.421**	3.77%	0.024
AGE	0.003	0.047%	0.891	0.028	0.25%	0.719
QUAL	0.090***	1.31%	0.000	0.149*	1.37%	0.087
INFLUENCE	−0.175	−2.53%	0.172	−0.287	−2.56%	0.373
LNFEE	0.019	0.27%	0.549	0.128	1.15%	0.184
Industry and Year FE	Yes			Yes		
n	31,346			3,248		
Pseudo R^2	0.071			0.119		

Panel B: Do Clients of the Mid-Tier Have Fewer Restatements than Clients of Small Auditors?

	Full Sample			Matched Sample		
Variable	**Estimate**	**Marginal Effect**	**p-value**	**Estimate**	**Marginal Effect**	**p-value**
Intercept	−1.365***		0.001	−6.055***		0.000
MIDTIER	−0.024	−0.25%	0.776	−0.127	−0.88%	0.354
CONTROLS	Yes			Yes		
Industry and Year FE	Yes			Yes		
n	8,772			1,406		
Pseudo R^2	0.067			0.133		

*, **, *** Denote statistical significance at the 10, 5, and 1 percent levels, respectively, using a two-tailed test.
This table displays coefficients from estimating a probit model (Equation (2)) where the dependent variable (*RESTATE*) equals 1 if the client subsequently issues an accounting restatement, 0 otherwise. In Panel A, the model is estimated on both the full Big 4 versus Mid-tier sample and the matched Big 4 versus Mid-tier sample. In Panel B, the model is estimated on a sample of Mid-tier and small audit firms. All independent variables are winsorized at the 1st and 99th percentiles to reduce in the influence of outliers. The model includes industry- and year-fixed effects, where industries are defined using two-digit SIC codes. p-values are based on Z-statistics, which are clustered by firm and year (Peterson 2009).

American Accounting Association

Downloaded from http://meridian.allenpress.com/ajpt/article-pdf/33/4/197/1603079/ajpt-50792.pdf by Case Western Reserve University, Charles Gletsu on 23 February 2021

TABLE 8

Alternative Measures of Misstatements

Panel A: AAER Analysis

	Full Sample			Matched Sample		
Variable	**Estimate**	**Marginal Effect**	**p-value**	**Estimate**	**Marginal Effect**	**p-value**
Intercept	−2.595***		0.000	−11.663***		0.000
BIG4	−0.203*	−0.22%	0.092	−0.536**	−0.04%	0.034
CONTROLS	Yes			Yes		
Industry FE	Yes			Yes		
Year FE	Yes			Yes		
n	37,488			5,950		
Pseudo R^2	0.151			0.459		

Panel B: Income-Decreasing Restatements Only

	Full Sample			Matched Sample		
Variable	**Estimate**	**Marginal Effect**	**p-value**	**Estimate**	**Marginal Effect**	**p-value**
Intercept	−2.061***		0.003	−2.927***		0.000
BIG4	−0.089*	−1.16%	0.063	−0.182**	−1.70%	0.020
CONTROLS	Yes			Yes		
Industry FE	Yes			Yes		
Year FE	Yes			Yes		
n	37,488			5,950		
Pseudo R^2	0.064			0.053		

*, **, *** Denote statistical significance at the 10, 5, and 1 percent levels, respectively, using a two-tailed test.
Panel A displays coefficients from estimating Equation (2) after replacing the dependent variable with *AAER*, which equals 1 if the client is subsequently subject to an Accounting and Auditing Enforcement Release, 0 otherwise. Panel B displays coefficients from estimating Equation (2) after replacing the dependent variable with *NEGRESTATE*, which equals 1 if the firm subsequently issues an accounting restatement that results in a negative adjustment to net income, 0 otherwise. In all regressions, the control variables from Equation (2) are included. All independent variables are winsorized at the 1st and 99th percentiles to reduce in the influence of outliers. The model includes industry- and year-fixed effects, where industries are defined using two-digit SIC codes. p-values are based on Z-statistics, which are clustered by firm (Peterson 2009).

Accounting and Auditing Enforcement Release, 0 otherwise.[24] The model includes all control variables in Equation (2) but we do not report the coefficients on these controls to save space. Using either the full or the matched sample, the coefficient on *BIG4* is significantly negative, suggesting that Big 4 clients are less likely to receive an AAER.

Panel B reports the results of estimating Equation (2) after replacing the dependent variable with *NEGRESTATE,* which equals 1 if the firm subsequently issues an accounting restatement that

[24] AAER data are obtained from the Center for Financial Reporting and Management at the University of California, Berkeley. See Dechow, Ge, Larson, and Sloan (2011) for details on the dataset. The dataset contains details on AAERs up to 2011. The sample for this test is the same sample used in the main analyses.

Downloaded from http://meridian.allenpress.com/ajpt/article-pdf/33/4/197/1603079/ajpt-50792.pdf by Case Western Reserve University, Charles Gletsu on 23 February 2021

results in a downward adjustment to net income, 0 otherwise. Using either sample, the coefficient on *BIG4* is significantly negative, suggesting that clients of Big 4 auditors are receiving higher quality audits.

To alleviate concerns that the results documented here are an artifact of the matching model, we replicate all of our analyses using the matching model of Boone et al. (2010).[25] We find that the results are qualitatively similar when using this model.

CONCLUSION

Because larger firms tend to select higher quality auditors, recent research questions whether the Big 4 provide higher audit quality than smaller auditors, once the endogenous choice of auditor has been controlled for (Lawrence et al. 2011; Boone et al. 2010). In this study, we re-examine this issue. We find that clients of Big 4 auditors are less likely to subsequently restate their earnings than are clients of non-Big 4 auditors. We also find weak evidence that clients of the Big 4 are less likely to issue a restatement than are clients of Mid-tier auditors (Grant Thornton and BDO Seidman). Taken together, the evidence is consistent with Big 4 auditors delivering higher quality audits. That we find different results when using a different audit quality proxy raises an important point. Proxies that are correlated with firm fundamentals (e.g., the absolute value of discretionary accruals) may not be capturing audit quality (see also Hribar and Nichols 2007). Similarly, readers must use caution when interpreting results from tests that use proxies that are actually measuring investors' perception of audit quality (e.g., the *ex ante* cost of equity capital).

The findings in this paper are subject to a few limitations. First, our inferences depend on the ability of accounting restatements to capture audit quality. We have presented arguments that support the use of accounting restatements as a proxy for audit quality. Nevertheless, to the extent that accounting restatements arising from materially misstated financial statements are not attributable to the auditor, our proxy may measure audit quality with error. As well, the use of accounting restatements to proxy for audit quality cannot tell us anything about audit quality differences among firms that did not issue restatements. We can only infer that firms that restated their earnings received inferior audit quality than firms that did not restate. Second, a limitation of any matching procedure is that we are unable to match firms based on pre-treatment attributes (Lawrence et al. 2011). Third, the findings in this paper are silent on the source of the Big 4 auditors' superior audit quality. The higher audit quality of Big 4 auditors may stem from these auditors' higher legal liability and/or greater economic rents to lose. On the other hand, the higher audit quality may simply be the result of better training programs (DeAngelo 1981). We leave this to future research.

REFERENCES

Asthana, S. C., and J. Boone. 2012. Abnormal audit fee and audit quality. *Auditing: A Journal of Practice & Theory* 31 (3): 1–22.

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

[25] The Boone et al. (2010) model is as follows: $BIG4 = \delta_0 + \delta_1 TA + \delta_2 lnSALES + \delta_3 LEV + \delta_4 EP + \delta_5 ISSUE + \delta_6 LOSS + \varepsilon$, where $BIG4 = 1$ if the client uses a Big 4 auditor, 0 otherwise; *TA* is the absolute value of total accruals, scaled by sales; *lnSALES* is the natural log of sales revenue; *LEV* is the debt to asset ratio; *EP* is the earnings-to-price ratio; $ISSUE = 1$ if the change in equity is greater than 10 percent, 0 otherwise; and $LOSS = 1$ if net income is negative and the absolute change in net income is greater than 10 percent, 0 otherwise.

American Accounting Association

Downloaded from http://meridian.allenpress.com/ajpt/article-pdf/33/4/197/1603079/ajpt-50792.pdf by Case Western Reserve University, Charles Gletsu on 23 February 2021

Behn, B. K., C. Jong-Hag, and T. Kang. 2008. Audit quality and properties of analyst earnings forecasts. *The Accounting Review* 83 (2): 327–349.

Bentley, K., T. Omer, and N. Y. Sharp. 2013. Business strategy, financial reporting irregularities, and audit effort. *Contemporary Accounting Research* 30 (2): 780–817.

Blankley, A. I., D. N. Hurtt, and J. MacGregor. 2012. Abnormal audit fees and restatements. *Auditing: A Journal of Practice & Theory* 31 (1): 79–96.

Boone, J. P., I. Khurana, and K. K. Raman. 2010. Do the Big 4 and second-tier firms provide audits of similar quality? *Journal of Accounting and Public Policy* 29 (4): 330–352.

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

Cheffers, M. D., D. Whalen, and O. Usvyatsky. 2010. 2009 financial restatements: A nine year comparison. *Audit Analytics* (February). Available at: http://www.complianceweek.com/s/documents/AARestatements2010.pdf

Choi, J.-H., J.-B. Kim, A. Qiu, and Y. Zang. 2012. Geographic proximity between auditor and client: How does it impact audit quality? *Auditing: A Journal of Practice & Theory* 31 (2): 43–72.

DeAngelo, L. E. 1981. Auditor size and audit quality. *Journal of Accounting and Economics* 3 (3): 183–199.

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

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

Easton, P. 2004. PE ratios, PEG ratios, and estimating the implied expected rate of return on equity capital. *The Accounting Review* 79 (1): 73–95.

Fargher, N. L., and L. Jiang. 2008. Changes in the audit environment and auditors' propensity to issue going-concern opinions. *Auditing: A Journal of Practice & Theory* 27 (2): 55–77.

Francis, J. R., and J. Krishnan. 1999. Accounting accruals and auditor reporting conservatism. *Contemporary Accounting Research* 16 (1): 135–165.

Francis, J. R., E. L. Maydew, and C. Sparks. 1999. The role of Big 6 auditors in the credible reporting of accruals. *Auditing: A Journal of Practice & Theory* 18 (2): 17–34.

Francis, J. R. 2004. What do we know about audit quality? *The British Accounting Review* 36: 345–368.

Francis, J. R., and P. Michas. 2013. The contagion effect of low-quality audits. *The Accounting Review* 88 (2): 521–552.

Francis, J. R., P. Michas, and M. Yu. 2014. Office size of Big 4 auditors and client restatements. *Contemporary Accounting Research* 30 (4): 1626–1661.

Grant Thornton. 2007. *After Decades Of Consolidation, the Onset of Sarbanes-Oxley and the Fall of Andersen, the Audit Environment Is Changing*. Available at: http://www.GrantThorton.com (last accessed May 30, 2007).

Guo, S., and M. Fraser. 2010. *Propensity Score Analysis: Statistical Methods and Applications*. Thousand Oaks, CA: SAGE Publications Inc.

Heckman, J. 1979. The sample selection bias as a specification error. *Econometrica* 47 (1): 153–162.

Hogan, C. 1997. Costs and benefits of audit quality in the IPO market: A self-selection analysis. *The Accounting Review* 72 (1): 67–86.

Hribar, P., and C. Nichols. 2007. The use of unsigned earnings quality measures in tests of earnings management. *Journal of Accounting Research* 45 (5): 1017–1053.

Ireland, J. C., and C. S. Lennox. 2002. The large audit firm fee premium: A case of selectivity bias? *Journal of Accounting, Auditing, and Finance* 17 (1): 73–91.

Jones, J. J. 1991. Earnings management during import relief investigations. *Journal of Accounting Research* 29 (2): 193–228.

Kennedy, P. 1992. *A Guide to Econometrics*. Cambridge, MA: MIT Press.

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

Downloaded from http://meridian.allenpress.com/ajpt/article-pdf/33/4/197/1603079/ajpt-50792.pdf by Case Western Reserve University, Charles Gletsu on 23 February 2021

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

Krishnan, G. V. 2003. Audit quality and the pricing of discretionary accruals. *Auditing: A Journal of Practice & Theory* 22 (1): 109–126.

Krishnan, J., and J. Krishnan. 1996. The role of economic trade-offs in the audit opinion decision: An empirical analysis. *Journal of Accounting, Auditing, and Finance* 11 (4): 565–586.

Lawrence, A., M. Minutti-Meza, and P. Zhang. 2011. Can Big 4 versus non-Big 4 differences in audit-quality proxies be attributed to client characteristics? *The Accounting Review* 86 (1): 259–286.

Lennox, C., and J. A. Pittman. 2010. Big Five audits and accounting fraud. *Contemporary Accounting Research* 27 (1): 209–247.

Lennox, C., J. R. Francis, and Z. Wang. 2012. Selection models in accounting research. *The Accounting Review* 87 (2): 589–616.

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

Newton, N. J., D. Wang, and M. S. Wilkins. 2013. Does a lack of choice lead to lower quality? Evidence from auditor competition and client restatements. *Auditing: A Journal of Practice & Theory* 32 (3): 31–67.

Palmrose, Z.-V., and S. Scholz. 2004. The circumstances and legal consequences of non-GAAP reporting: Evidence from restatements. *Contemporary Accounting Research* 21 (1): 139–180.

Peel, M. J., and G. H. Makepeace. 2012. Differential audit quality, propensity score matching, and Rosenbaum bounds for confounding variables. *Journal of Business, Finance, and Accounting* 39 (5): 606–648.

Peterson, M. 2009. Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies* 22: 435–480.

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

Rosenbaum, P. R., and D. B. Rubin. 1983. The central role of propensity score in observational studies for causal effects. *Biometrika* 70 (1): 41–55.

Subramanyam, K. R. 1996. The pricing of discretionary accruals. *Journal of Accounting and Economics* 22 (1–3): 249–281.

Teoh, S. H., and T. J. Wong. 1993. Perceived auditor quality and the earnings response coefficient. *The Accounting Review* 68 (2): 346–366.

U.S. Chamber of Commerce. 2006. *Auditing: A Profession at Risk*. Washington, DC: U.S. Chamber of Commerce.

Downloaded from http://meridian.allenpress.com/ajpt/article-pdf/33/4/197/1603079/ajpt-50792.pdf by Case Western Reserve University, Charles Gletsu on 23 February 2021

American Accounting Association