THE ACCOUNTING REVIEW
Vol. 94, No. 3
May 2019
pp. 113–147

American Accounting Association
DOI: 10.2308/accr-52206

Audit Quality and Specialist Tenure

Jennifer J. Gaver
The University of Georgia

Steven Utke
University of Connecticut

ABSTRACT: We argue that the association between auditor industry specialization and audit quality depends on how long the auditor has been a specialist. We measure audit quality using absolute discretionary accruals, income-increasing discretionary accruals, and book-tax differences. Our results, based on a sample of Big 4 audit clients from 2003–2015, indicate that auditors who have only recently gained the specialist designation produce a level of audit quality that does not surpass that produced by non-specialist auditors, and is generally lower than the audit quality produced by seasoned specialists. We estimate that the seasoning process takes two to three years. In contrast to prior research that finds no effect of specialization after propensity score matching, we find that seasoned specialists generally produce higher-quality audits than other auditors even after matching. This suggests that the audit quality effect associated with seasoned industry specialist auditors is not due to differences in client characteristics.

JEL Classifications: M42.

Data Availability: Data used in this study are available from public sources identified in the text.

Keywords: audit quality; industry specialization; auditor tenure; discretionary accruals; book-tax differences; auditor changes; propensity score matching.

I. INTRODUCTION

The association between auditor industry expertise and audit quality has attracted considerable interest from accounting researchers. Results from experimental studies suggest that industry expertise enhances auditor judgments regarding error detection (Solomon, Shields, and Whittington 1999; Owhoso, Messier, and Lynch 2002), risk assessments, and audit planning (Taylor 2000; Low 2004; Hammersley 2006). Results from archival studies also suggest that industry-specialist auditors provide higher-quality audits to their clients.[1] Compared to clients of non-specialist auditors, clients of specialists have lower discretionary accruals and higher earnings response coefficients (ERCs) (Balsam, Krishnan, and Yang 2003; Krishnan 2003).

A limitation of archival research on auditor industry specialization is that specialization is not directly observable. Archival researchers typically use an auditor's within-industry market share as an indirect measure of specialization, in contrast to the experience-based measures of industry expertise used in experimental studies. Recent archival work challenges the conclusion that auditor industry expertise, measured by market share, is associated with audit quality. In particular, Minutti-Meza (2013) (hereafter, MM) finds no evidence that auditor industry expertise improves audit quality after using propensity score matching (PSM) to address functional form misspecification arising from differences between characteristics of specialist versus non-

This paper has benefited from helpful comments from Ashley Austin, Ben Ayers, Denny Beresford, Bill Buslepp (discussant), Nathan Goldman, Matt Hart, Youree Kim, workshop participants at the University of Connecticut, the 2014 AAA Southeast Regional Meeting, and the 2014 AAA Annual Meeting, and especially Mike Ricci and Quinn Swanquist. We thank Mark L. DeFond (editor) and two anonymous reviewers for their careful and constructive guidance during the review process. We gratefully acknowledge the support of The University of Georgia's Terry College of Business and J. M. Tull School of Accounting. Steven Utke also thanks the University of Connecticut School of Business for financial support.

Editor's note: Accepted by Mark L. DeFond.

Submitted: August 2015
Accepted: July 2018
Published Online: July 2018

[1] Throughout the paper, we refer to audit firms as "audit firms" or "auditors" and clients as "clients" or "firms."

specialist auditors' clients. However, DeFond, Erkens, and Zhang (2017a) and Shipman, Swanquist, and Whited (2017) show that subjective design choices underlying PSM affect the composition of the matched sample and can alter inferences from a PSM analysis.

In contrast to this recent literature focusing on econometric issues in the relation between auditor characteristics and audit quality, our paper revisits the relation between audit quality and auditor industry specialization using a more nuanced measure of specialization than market share alone. Existing market share-based industry specialization measures implicitly assume that upon achievement of a market share threshold, the auditor immediately implements industry best practices on an audit firm-wide level. In contrast, we argue that a coordinated approach to industry-wide clients develops over time (Gendron, Cooper, and Townley 2007) and propose incorporating *specialist tenure* into existing industry specialization measures. Specialist tenure refers to an auditor's experience (length of service) as the market share leader in a given *industry*. This differs from auditor tenure, which focuses on an auditor's experience on a given *client*. We refer to longer (shorter) tenured specialists as "seasoned" ("unseasoned").

The consideration of specialist tenure is important if it is possible for an audit firm to become a market share-based specialist due to exogenous events unrelated to auditor expertise. For example, mergers and acquisitions by clients, industry entry by clients, and client growth could make an audit firm the industry market share leader without any action by the audit firm itself. If it then takes time to hire and develop personnel and implement the audit firm-wide technologies needed to meet the increased demand for high-quality audits across the auditor's larger market share in that industry, the newly created specialist will not immediately function as an industry expert (see Bills, Swanquist, and Whited 2016). In this scenario, newly created (unseasoned) specialists will function more like non-specialist auditors, producing audits of lower quality compared to auditors that have had sufficient time to make the necessary investments in personnel and technology to function as experts (seasoned specialists). That is, by incorporating specialist tenure, we account for the possibility that an audit firm becomes an industry market share leader following a rapid market share increase unrelated to auditor expertise. On the other hand, if specialists arise because clients are attracted to the industry expertise *previously* developed by the audit firm, then specialist tenure is irrelevant. It is possible that there are some cases where expertise drives the auditor's market share and others where market share is a result of exogenous events unrelated to the auditor. Ignoring the latter scenario and viewing all specialists identically blurs the distinction between specialist and non-specialist auditors and is a potential explanation for the insignificant results reported by MM.

We begin our analysis by replicating the results of MM. We measure audit quality using absolute discretionary accruals (*ADA*), income-increasing discretionary accruals (*DA*), and book-tax differences (*BTD*) in our main analyses.[2] Our samples, spanning 2003–2015, consist of up to 25,901 firm-years when measuring audit quality using *ADA,* 13,863 firm-years when measuring audit quality using income-increasing *DA,* and 10,551 firm-years when measuring audit quality using *BTD*. In all cases, the firm has a Big 4 auditor in both the current and the preceding year.

We observe a positive association between a market share-based measure of auditor specialization and audit quality in the full (unmatched) *ADA* and income-increasing *DA* samples, but not in the *BTD* sample. Thus, even without matching, an industry specialization proxy based only on market share produces somewhat mixed results. Following MM, we then reexamine the relation between auditor specialization and audit quality after employing PSM. The general pattern of results (significant in the *ADA* and income-increasing *DA* samples, insignificant in the *BTD* sample) remains after matching, although the statistical significance is somewhat weaker.

DeFond et al. (2017a) and Shipman et al. (2017) show that PSM results are sensitive to design choices, so we examine the robustness of our findings across nine separate PSM specifications. In contrast to MM, we find a significant relation between market share-based industry specialization and audit quality in 44 (89) percent of our PSM samples when measuring audit quality using either *ADA* or *BTD* (using income-increasing *DA*). The impact of PSM design choices on the relation between auditor characteristics and audit quality is not the focus of our study (DeFond et al. [2017a] thoroughly address this issue). However, the results suggest that the relation between market share-based industry specialization and audit quality in the unmatched sample should not be dismissed as an artifact of differences between specialist and non-specialist auditors' clients, as argued by MM.

Our main innovation is to partition industry specialists into experienced (seasoned) and newly created (unseasoned) specialists, resulting in an industry specialization measure that reflects both market share *and* seasoning. In order to create our partition, we must first identify unseasoned specialists. This requires us to investigate the events preceding the creation of industry specialists, which is critical to our argument that market share dominance can precede the development of industry expertise. Consistent with specialist tenure likely playing an important role in auditor industry specialization, in almost all

[2] We also perform robustness tests that measure audit quality using dichotomous measures (meeting or beating analysts' forecasts, receiving a going concern opinion) and the earnings response coefficient.

cases, we find that changes in the industry market leader are due to factors external to the auditor, such as industry entry by clients and client growth, rather than auditor changes. There is little evidence that auditors attain market dominance by attracting new clients due to their pre-existing industry expertise.

We next reestimate the MM model after partitioning specialists into seasoned and unseasoned categories. We find that compared to unseasoned specialists, seasoned specialists generally provide higher audit quality. This result holds across our main audit quality measures (*ADA,* income-increasing *DA,* and *BTD*) and nearly all matching strategies, including strategies other than PSM. Further, seasoned specialists provide higher audit quality than non-specialists, again across our main audit quality measures and nearly all matching strategies.[3] In terms of economic significance, having a seasoned specialist, rather than a non-specialist, auditor improves audit quality by about 6 (42) percent using the *ADA* and income-increasing *DA* (*BTD*) measures. For all three audit quality measures and all matching strategies, the quality of audits produced by unseasoned specialists is either not statistically different from, or worse than, the quality of audits produced by non-specialist auditors. Because we find that specialists are generally created by factors external to the auditor, this suggests that auditors likely fall backward into the industry leadership position and *then* develop the tools necessary to provide higher-quality audits. Thus, expertise follows market share dominance, rather than the other way around.

We also investigate the length of time it takes for an auditor to become seasoned. While recent work investigates an auditor's learning curve on a specific client (Cameran, Francis, Marra, and Pettinicchio 2015; Cassell, Hansen, Myers, and Seidel 2017), little evidence exists on the speed with which industry knowledge is created within an audit firm. Our findings suggest that, on average, seasoning develops during the first two to three years that an auditor is the market share leader, depending on the audit quality measure. Auditors in their second or third consecutive year as industry leader often produce audit quality that is statistically indistinguishable from that of seasoned specialists, while auditors in their first one or two years as market share leader often produce audits of lower quality than seasoned specialists. Thus, the seasoning process for industry specialists appears to take longer than the relatively quick learning process on a specific client (Cassell et al. 2017). An implication is that auditors in their initial years as market share leaders have been misclassified as specialists in prior research, blurring the relation between industry expertise and audit quality. We also find some evidence that audit quality is increasing in a continuous measure of specialist tenure. Thus, we encourage researchers to consider seasoning when creating specialization measures.

We contribute to the literature on auditor industry expertise in several ways. First, we find that dominant market share is not a sufficient condition for industry expertise. Researchers must also consider the specialist's tenure. Our results are also potentially relevant for researchers outside of auditing who acknowledge the limitations of market share-based measures of expertise (e.g., Bao and Edmans 2011). Second, we question recent research suggesting that the relation between industry specialization and audit quality is due to differences in client characteristics. Our analysis suggests that these prior results may be sensitive to the specific matching technique used, echoing recent studies that suggest that PSM be used with caution (DeFond, Erkens, and Zhang 2017b; Shipman et al. 2017) or not at all (King and Nielsen 2016).[4] Although unseasoned specialists constitute only about 2 percent of total observations (11 percent of specialist observations), excluding them makes the effect of specialists on audit quality generally robust to matching. Third, we provide evidence on the process by which industry expertise is created in an audit firm. We find that it takes two to three years for an auditor to function as an industry expert, in contrast to the relatively quick learning process for an auditor on an individual client. We also find that auditors often become industry market share leaders for reasons outside of their control and then invest in industry expertise. This challenges a common assumption in the literature that clients self-select into industry-specialist auditors (Gul, Fung, and Jaggi 2009). Finally, we note that, if implemented, mandatory auditor rotation is likely to create unseasoned specialists that may need a seasoning period to function as true industry experts.

II. HYPOTHESES

We contend that an auditor's within-industry market share is an incomplete specialization measure because it ignores the importance of seasoning in developing expertise. However, seasoning alone does not make an industry expert (Ericsson, Charness, Feltovich, and Hoffman 2006). A critical mass of business, proxied for by market share, must also be present to induce an auditor to invest in industry-specific personnel and technologies. Thus, we argue that industry expertise is a multidimensional construct that requires both market share *and* seasoning.[5]

[3] However, for reasons discussed in Section V, our results vary somewhat with alternative seasoning periods (two or three years, versus the one-year results discussed here), and are generally weaker when using alternative audit quality proxies (meeting or beating analysts' forecasts, receiving a going concern opinion, or the earnings response coefficient) and specialization measures.

[4] PSM can also be problematic when partitioning a sample into multiple groups, as required for our study. We discuss and address this issue in Section V.

[5] While we are not the first to question the validity of market share-based specialization measures (see Audousset-Coulier, Jeny, and Jiang [2016] for a review), we are the first to propose within-industry seasoning as a potential improvement to existing industry specialization measures.

Gendron et al. (2007) describe the development of audit expertise as an iterative process where good practices are developed, documented, and subsequently validated *over time* through the experiences of audit firm personnel.[6] An audit firm may decide to implement these practices on an audit firm-wide level by investing in various decision aids (such as checklists), training sessions (Bédard 1989; Power 1996), and guidance from a centralized national office (Danos, Eichenseher, and Holt 1989).[7] This type of coordinated approach is more likely when the audit firm has a sufficient volume of clients that could benefit from similar practices, such as firms in the same industry. Thus, although audit firm personnel develop experience auditing specific clients, leveraging that expertise throughout the audit firm requires a commitment of resources that is unlikely to be made unless the audit firm has a large enough market share to justify investing in personnel and technologies specific to an industry.

An important point is that dominant market share in an industry can precede industry investment, as illustrated next. In the following example, based on an actual specialization change during our sample period, we assume that the home furnishing and equipment retail industry consists of three relatively large firms and two small firms. Best Buy (BBY) represents 45 percent of the market, Circuit City (CC) represents 20 percent of the market, Linens-N-Things (LIN) represents 15 percent of the market, and two small firms (Other1 and Other2) represent the remaining 20 percent of the market.[8] In Year 1, Ernst & Young (EY) audits BBY, KPMG audits CC and LIN, Deloitte audits Other1, and PricewaterhouseCoopers (PwC) audits Other2. In Year 2, BBY changes auditors from EY to Deloitte, and EY becomes the auditor of LIN. Deloitte is now the market share-based specialist not only on the audit of BBY, but also on the audit of Other1. We argue that prior to becoming the market share leader, it is unlikely that Deloitte invested in the industry heavily enough to immediately behave as an industry expert upon gaining the leadership position. After obtaining the leadership position, Deloitte has the incentive to develop and implement audit firm-wide industry guidance, which then benefits its audit clients throughout the industry.

Firm (Market Share)	Year 1	Year 2
BBY (45%)	EY (specialist)	Deloitte (specialist)
CC (20%)	KPMG	KPMG
LIN (15%)	KPMG	EY
Other1 (10%)	Deloitte	Deloitte (specialist)
Other2 (10%)	PwC	PwC

In this example, the industry specialist changed because BBY, the largest firm in the industry, changed auditors. A change could also result from rapid growth of clients relative to their industry competitors (e.g., BBY and CC could flip positions in the market), client mergers (e.g., CC could acquire LIN and Others), new firms entering the market, or existing firms leaving the market.[9] Interestingly, in the real-world case, BBY was required to change auditors due to a conflict of interest with EY (Taub 2005). BBY did not self-select to Deloitte, one of the smallest Big 4 auditors in the industry, because of Deloitte's industry expertise; if industry expertise drove the auditor selection, then BBY would have selected the auditor with the second-largest pre-change industry market share, KPMG.

As illustrated in this example, we argue that if the achievement of a dominant market share is the result of exogenous events, then a dominant market share may not indicate an industry expert; the auditor also needs time to establish audit firm-wide technologies to support the higher volume of business in the industry. Investments made to support the higher volume of business will, over time, result in higher-quality audits.[10] We, therefore, predict that audits performed by seasoned specialists will be of higher quality than audits performed by either unseasoned specialists or non-specialist auditors. Our hypotheses, in alternative form, are as follows:

H1: The quality of audits produced by seasoned industry-specialist auditors is higher than the quality of audits produced by unseasoned industry-specialist auditors.

[6] While we focus on auditors, researchers in industrial organization also examine the development and transfer of knowledge within firms (Tsai 2001; van Wijk, Jansen, and Lyles 2008; Phelps, Heidl, and Wadhwa 2012).

[7] A Big 4 senior manager confirmed that auditors distribute industry knowledge using these methods, plus email alerts.

[8] This example is simplified for expositional purposes. Having a small number of firms in an industry is not a necessary condition for the creation of unseasoned specialists. Over our sample period, unseasoned specialists serve as auditors on clients in diverse industries, some comprised of only a few firms and others having over 450 firms.

[9] In the real-world case, BBY changed auditors and LIN exited the industry because it was taken private. In Section III, we examine the causes of the specialist changes in our sample. Most are due to exogenous events unrelated to the auditor. We argue that new specialists created by exogenous events may not immediately behave as industry experts, but instead will develop expertise over time. We acknowledge that some exogenous events might be more likely to create incentives for the auditor to invest in industry expertise than others. For example, changes that require an auditor to increase the scope of its industry operations to serve new or larger clients likely provide greater incentives for investment than market share changes due solely to client divestitures or industry exits. However, parsing these differences is beyond the scope of our study and is a potentially fruitful area for future research.

[10] We acknowledge that our example does not account for small firms (that are unable to shift the industry-specialist designation themselves) selecting into specialist auditors after a large industry change creates a new specialist. To the extent that this takes time to occur, it is consistent with our hypotheses suggesting that specialization develops over time.

H2: The quality of audits produced by seasoned industry-specialist auditors is higher than the quality of audits produced by non-specialist auditors.

H3: The quality of audits produced by unseasoned industry-specialist auditors does not differ from the quality of audits produced by non-specialist auditors.

Several factors work against finding these results. Our distinction between seasoned and unseasoned specialists assumes that unseasoned specialists have not had sufficient time to make industry-specific investments that will result in higher-quality audits for all industry clients. However, individual audit personnel within the firm may have ample experience conducting audits of other clients in the industry, or the auditor could hire new personnel with industry experience. If these experiences can be quickly leveraged to produce quality audits on a larger scale, then we are less likely to find a significant difference in audit quality between seasoned and unseasoned specialists (H1).[11] Likewise, non-specialist audit firms may choose to institutionalize best practices from an industry even if they are not the market share leader in that industry, reducing the likelihood that H2 is supported. Further, if an auditor achieves a dominant market share by gradually developing expertise and increasing its presence in an industry, then the auditor may immediately behave as a specialist, reducing the likelihood that we find support for H1 or H3. Finally, our sample includes only Big 4 audit firms. If Big 4 auditors are experts in all industries (as is often claimed in their marketing materials), then we will find no differences in audit quality among Big 4 auditors using any measure of industry specialization, including measures that distinguish unseasoned from seasoned specialists.[12]

Prior studies (Geiger and Raghunandan 2002; Johnson, Khurana, and Reynolds 2002; J. Myers, L. Myers, and Omer 2003; Carcello and Nagy 2004) find that audit quality is lower when auditors are new to their clients. Therefore, an alternative explanation for a lower level of audit quality produced by unseasoned specialists is that, compared to seasoned specialists, unseasoned specialists are simply less familiar with their clients. However, many newly designated (unseasoned) specialists are *not* new to their clients.[13] In our example, Deloitte, an unseasoned specialist, is new to BBY, but not to Other1. As discussed in the next section, we include an auditor tenure variable in our model to control for its association with audit quality.

III. RESEARCH DESIGN

Regression Model and Variables

Our hypotheses relate audit quality to auditor industry expertise. As recommended by DeFond and Zhang (2014), we use multiple measures of audit quality in our tests: discretionary accruals (absolute and income-increasing) and book-tax differences. We focus on continuous measures that are available for a large number of firms because our hypotheses posit a nuanced relation between auditor expertise and audit quality, and our sample consists of firms that, as Big 4 audit clients, generally have a high level of audit quality. Alternative measures, such as restatements and going concern opinions, are likely to be relatively rare among these firms, resulting in low-power tests (DeFond and Zhang 2014).[14] However, in supplemental tests discussed in Section V, we also examine earnings response coefficients (ERC), meeting or beating analysts' earnings forecasts, and going concern opinions.

Discretionary accruals are a common audit quality proxy, and book-tax differences are incrementally useful to accruals in identifying earnings management, fraud, and restatements (Phillips, Pincus, and Rego 2003; Ettredge, Sun, Lee, and Anandarajan 2008; Badertscher, Phillips, Pincus, and Rego 2009; Chi, Pincus, and Teoh 2014). We estimate discretionary accruals using the performance-adjusted Jones model (Kothari, Leone, and Wasley 2005), and measure book-tax differences as

[11] A newly created specialist faces an increased scale of production and may not have enough personnel with relevant industry expertise to serve both the new (or larger) clients and the existing clients in the industry. In this case, professionals from other industries will be assigned to the audits, decreasing the average level of industry expertise across all of the auditor's clients in the industry (similar to the findings of Bills et al. 2016).

[12] See, for example, Ernst & Young's list of Global Industry Centers, in which they have "invested . . . for sharing industry-focused knowledge and experience" (available at: https://www.ey.com/us/en/industries). This list covers nearly all industries.

[13] While auditor tenure relates to repetition on a particular client, specialist tenure (seasoning) relates to repetition as the leader in a particular industry. If adding one, or a few, large client(s) causes an auditor to become a specialist, then that auditor is unseasoned not only for the new client(s), but also for all of its continuing (longer-tenured) clients in the industry. It is, therefore, possible to be an unseasoned specialist on a client with which the auditor has a long-term relationship. In our *ADA* sample (described in Section IV), unseasoned specialists are more likely than seasoned specialists to be in their first year with a client (3.7 percent of unseasoned observations involve new clients versus 1.4 percent of seasoned observations). Both seasoned and unseasoned auditors have relatively long auditor-client relationships, with average auditor tenures of 14.21 and 14.51 years, respectively. In untabulated tests, discussed in Section V, we find that our results are robust to examining seasoned and unseasoned specialists excluding auditor changes, ruling out the possibility that auditors that are new to their clients drive our results.

[14] Consistent with most prior research, we define audit quality based on audit outputs, reasoning that audit failure occurs when financial statements are misstated (DeFond and Zhang 2014). The Public Company Accounting Oversight Board (PCAOB) and some recent research (Bell, Causholli, and Knechel 2015) define audit quality using audit inputs. However, audit inputs are generally unobservable and are not necessarily correlated with the quality of audit outputs (Bell et al. 2015, footnote 4).

grossed-up deferred tax expense scaled by pre-tax income (Chi et al. 2014). Following prior research, we assume that audit quality is negatively associated with the absolute value of firms' discretionary accruals (*ADA*), income-increasing discretionary accruals (*DA*), and book-tax differences (*BTD*).[15] Appendix A provides variable definitions.

We employ a model adapted from MM that regresses audit quality (*AQ*, measured as *ADA*, income-increasing *DA*, or *BTD*) on a dichotomous market share-based measure of auditor industry specialization, control variables, and year fixed effects.[16] That is, before examining the effect of seasoning, we examine the effect of industry specialization using an existing market share-based specialization proxy. We classify an auditor as a specialist (*SPECIALIST*) when its U.S. national-level market share is the highest in a given industry-year *and also* more than 10 percent higher than the next-largest competitor, with industry based on two-digit SIC codes (Reichelt and Wang 2010).[17] This model is presented in Equation (1).

$$\begin{aligned} AQ_{it} = {} & \alpha_0 + \beta_1 SPECIALIST_{it} + \beta_2 LOGMKT_{it} + \beta_3 LEV_{it} + \beta_4 ROAL_{it} + \beta_5 ROA_{it} + \beta_6 LOSS_{it} + \beta_7 CFO_{it} + \beta_8 BTM_{it} \\ & + \beta_9 ABS\_AC\_LAG_{it} + \beta_{10} GROWTH_{it} + \beta_{11} ALTMAN_{it} + \beta_{12} STDEARN_{it} + \beta_{13} TENURE_{it} + \beta_t YEAR\_FE_t + v_{it} \end{aligned} \tag{1}$$

If market share-based industry-specialist auditors provide higher audit quality than non-specialist auditors, then β_1 will be negative.

In Equation (1), auditors are classified as industry experts based only on market share. However, we argue that a critical mass of business in an industry is a necessary, but not sufficient, condition to establish auditor expertise and that "time in grade" as the industry leader is also needed. To test this assertion, we replace *SPECIALIST* in Equation (1) with two variables: *SEASONED* and *UNSEASONED*. *UNSEASONED* takes on the value of 1 when the auditor is in its first year of being classified as a specialist in a given industry, and is 0 otherwise.[18] *SEASONED* takes on the value of 1 when the auditor is classified as a specialist, and *UNSEASONED* is coded 0. A 0 coding on both *UNSEASONED* and *SEASONED* indicates that the auditor is not an industry specialist. The revised model is given in Equation (2).

$$\begin{aligned} AQ_{it} = {} & \alpha_0 + \beta_1 SEASONED_{it} + \beta_2 UNSEASONED_{it} + \beta_3 LOGMKT_{it} + \beta_4 LEV_{it} + \beta_5 ROAL_{it} + \beta_6 ROA_{it} + \beta_7 LOSS_{it} \\ & + \beta_8 CFO_{it} + \beta_9 BTM_{it} + \beta_{10} ABS\_AC\_LAG_{it} + \beta_{11} GROWTH_{it} + \beta_{12} ALTMAN_{it} + \beta_{13} STDEARN_{it} + \beta_{14} TENURE_{it} \\ & + \beta_t YEAR\_FE_t + v_{it} \end{aligned} \tag{2}$$

If seasoned specialists outperform non-specialists (H2), then the coefficient on *SEASONED* (β_1) in Equation (2) will be negative. If seasoned specialists outperform unseasoned specialists (H1), then the coefficient on *UNSEASONED* (β_2) will exceed the coefficient on *SEASONED* (β_1) in Equation (2). Finally, if the quality of audits produced by unseasoned industry specialists does not differ from the quality of audits produced by non-specialist auditors (H3), then the coefficient on *UNSEASONED* (β_2) in Equation (2) will be statistically indistinguishable from zero.[19]

[15] We focus on income-increasing, as opposed to income-decreasing or signed discretionary accruals, consistent with prior research (Becker, DeFond, Jiambalvo, and Subramanyam 1998).

[16] The model includes controls for the size (*LOGMKT*), profitability (*ROA, ROAL, CFO,* and *LOSS*), risk (*LEV, ALTMAN,* and *STDEARN*), and growth profile (*BTM* and *GROWTH*) of the client and the length of the client's relationship with the auditor (*TENURE*). The absolute value of lagged accruals (*ABS_AC_LAG*) controls for scale effects in the audit quality variables that are not addressed by other variables in the model. We do not include MM's *BIG4* control variable because all of our sample firms are Big 4 clients. All other control variables correspond to MM's model. Variable definitions are provided in Appendix A. Consistent with MM and Reichelt and Wang (2010), we do not include industry fixed effects in the accruals models because accruals are estimated by industry. We include industry fixed effects (two-digit SIC) in the book-tax differences model.

[17] Palmrose (1986) and Neal and Riley (2004) define the minimum market share for specialization as 1.2 times the inverse of the number of Big N auditors. In our study, this is 30 percent (1.2/4 = 0.30). In order to clearly distinguish between firms audited by specialist and non-specialist auditors, we exclude observations if their auditor is classified as a specialist under the 30 percent rule, but not under our main specialist definition (industry market share leader by at least 10 percent). Except for the *BTD* tests, our inferences are unchanged if we include these firm-years in our analysis. Regarding the *BTD* tests, analysis suggests that specialists measured under the alternative definition are particularly good at constraining *BTD*. As such, including the 30 percent specialists as non-specialists in the *BTD* robustness test adds measurement error to the non-specialists category, as these auditors behave as true specialists.

[18] Although we initially define unseasoned specialists as auditors in their first year as market share leader, the seasoning process could take more than one year. In Section V, we investigate the length of the seasoning period.

[19] In general, prior research does not differentiate audit quality within specialist auditors. An exception is Cahan, Jeter, and Naiker (2011), who find that clients of specialist auditors with a small number of large clients have lower discretionary accruals compared to clients of specialist auditors with a large number of small clients. However, as discussed in more detail below, client size is correlated with discretionary accruals and is, therefore, a confounding factor relative to the auditor specialist-audit quality relation. Our paper differentiates specialists using an attribute (seasoning) other than client size. We also employ multiple matching techniques to control for functional form misspecification in the relation between auditor specialization and audit quality.

TABLE 1

Reasons for Industry Specialist Auditor Changes, By Year

Panel A: Reason for Changes by Year

Year	Number of Industries with Unseasoned Auditors	Auditor Change to Unseasoned Auditor—Major	Auditor Change to Unseasoned Auditor—Minor	Merger by Unseasoned Auditor Client	Divestiture by Client	Industry Entry
2003	8	1	0	1	0	1
2004	9	0	0	0	1	3
2005	8	3	0	1	0	1
2006	9	2	0	0	0	1
2007	4	0	1	2	0	0
2008	1	0	0	0	0	0
2009	6	1	0	1	0	3
2010	5	0	0	1	0	0
2011	7	1	1	0	1	2
2012	8	1	0	0	1	1
2013	3	0	1	1	1	0
2014	7	3	0	0	0	1
2015	1	0	0	0	0	0
2003–2015	76	12	3	7	4	13
Frequency		11.01%	2.75%	6.42%	3.67%	11.93%

(continued on next page)

Reasons for Specialist Changes

Unseasoned specialists are created when auditors' within-industry market shares change. Our study of specialist tenure, which requires identifying unseasoned specialists, presents an opportunity to examine the events preceding market share changes. To our knowledge, ours is the first study to investigate and document this phenomenon. Our analysis encompasses all specialist changes occurring over our sample period (2003–2015). We first calculate U.S. national-level industry specialization using firms with the minimum required data in Compustat: U.S. headquarters location, sales, SIC, and auditor. We then identify 76 industry-years with an unseasoned specialist. For each of these industries, we obtain the industry and auditor composition in the year of the specialist change, as well as the prior year, and then assess the reasons for the change. Based on our review, we identify nine main events associated with specialist changes. Table 1 presents our results, broken out by year.

In order for there to be a relationship between specialist tenure and audit quality, it must be possible for market share dominance to precede the development of industry expertise. This occurs when a new specialist arises from industry events unrelated to the auditor. We find that in most cases (64 out of 76 industry-years; see Table 1), the unseasoned specialist auditor became the market leader due to exogenous events, such as client growth, industry entry, and mergers. The three most common drivers of specialist changes are client entry to or exit from the industry and client sales growth. About one-third of specialist changes are caused by multiple factors. Cases where auditors became specialists solely because clients switched to the new specialist are relatively rare (eight out of 76 industry-years). In four additional cases, an auditor change large enough to affect the specialist designation occurred, along with other industry events affecting the specialist designation, most commonly, client sales growth and industry entry. Finally, in three cases, the new specialist was created as a result of a minor auditor change, too small to change the specialist on its own, occurring in combination with some other event. Even in these cases, we found no clear evidence that clients changed auditors because the new auditor was the industry specialist. In fact, in industries where a new specialist was created, we found nine cases of large clients switching to a non-specialist auditor, suggesting that clients switching auditors do not herd toward the new specialist.[20] Thus, we find little evidence that audit firms first develop industry

[20] One reason that firms may be reluctant to herd to the industry market leader is fear that proprietary information could leak to a competitor via a common auditor (e.g., Coke and Pepsi never share an auditor). However, very few firms disclose a reason for changing auditors. When an explanation for an auditor change is offered by our sample firms, it is most often a simple statement that the audit is periodically put out to bid as a good governance practice.

TABLE 1 (continued)

Panel B: Reason for Changes by Year (continued)

Year	Industry Exit	Former Client Acquired	Client Sales Growth	Major Auditor Change Unrelated to Unseasoned Auditor	Industries with Multiple Causes	Percent of Industries with Multiple Causes
2003	4	0	3	1	3	37.50%
2004	3	1	3	1	2	22.22%
2005	1	1	3	0	2	25.00%
2006	3	1	5	0	3	33.33%
2007	0	1	2	1	2	50.00%
2008	1	0	0	0	0	0.00%
2009	0	0	3	1	2	33.33%
2010	2	0	3	0	1	20.00%
2011	2	0	4	1	2	28.57%
2012	2	1	4	2	3	37.50%
2013	1	0	1	1	1	33.33%
2014	1	3	1	1	2	28.57%
2015	1	0	0	0	0	0.00%
2003–2015	21	8	32	9	23	30.26%
Frequency	19.27%	7.34%	29.36%	8.26%		

We review all auditor specialist changes over the period from 2003 to 2015. The sample consists of all firms with the minimum required Compustat variables: U.S. headquarters location, sales, SIC, and auditor. We identify nine major reasons for the new specialist obtaining the market leadership role. Merger by Unseasoned Auditor Client indicates that the new industry-specialist's client grew through acquisition, increasing the new specialist's market share above the threshold. Divestiture by Client indicates that some firm in the industry experienced a divestiture that allowed the new specialist to capture the necessary market share. Industry Entry indicates that some new firm joined the industry (e.g., via initial public offering [IPO]), altering market share in a way that creates a new specialist. Relatedly, Industry Exit indicates that some firm previously in the industry exited (e.g., via bankruptcy or privatization). Former Client Acquired indicates that a client not associated with the new specialist was acquired. Similar to Divestiture by Client, this indicates that the overall industry size decreased, pushing the new specialist above the market share-based specialization threshold. Auditor Change to Unseasoned Auditor—Major indicates a change to the new specialist that, in itself, is large enough to alter the specialist designation. The Minor version indicates a change to the new specialist that was not large enough itself to change specialization, but changed specialization when combined with other events. Client Sales Growth refers to sales growth or declines by firms in the industry that, combined, are large enough to change the specialist designation. Major Auditor Change Unrelated to Unseasoned Auditor indicates that a client changed auditors, resulting in a large enough market share change to trigger the creation of a new specialist, even though the new specialist was not involved in the auditor change. Industries with Multiple Causes indicates cases where an industry experienced more than one of the events detailed here. Note that the sum across columns exceeds the 76 total industry-years experiencing a specialist change because of the occurrence of multiple causes in some industries.

expertise and then attract clients based on this expertise. This provides initial evidence that seasoning likely affects industry-specialist audit quality; if an auditor finds itself in the position of market share leader before developing the tools to provide higher-quality audits, then the audit quality produced by new specialists will differ from that of seasoned specialists.

IV. DATA

Sample

To investigate the association between industry specialization and audit quality, we identify 68,002 firm-year observations from Compustat during the 2003 through 2015 period with a Big 4 auditor in the current and prior year. We limit the sample to Big 4 clients in order to avoid any self-selection bias that could arise from auditor choice (Francis 2011) and because only Big 4 auditors tend to be industry specialists (DeFond and Zhang 2014). Further, the Big 4 audit over 95 percent of firms, by market value, over our sample period. We begin the sample in 2003 to exclude observations influenced by the Enron scandal or the collapse of Arthur Andersen. We also exclude firms audited by Arthur Andersen in the prior year because Blouin, Grein, and Rountree (2007) report that some firms retained their Arthur Andersen team under the new auditor, making it difficult to attribute the reporting quality of the client to the characteristics of the current auditor. As shown in Table 2, of the 68,002 firm-year observations with Big 4 auditors in the current and prior year, 25,901 (13,863) have sufficient data to estimate Equations

TABLE 2
Sample Selection

Data Restrictions	**Audit Quality Proxy** *ADA* **and** *DA*	**Income-Increasing** *DA*	*BTD*
Compustat firm-year observations 2003–2015	118,105	118,105	118,105
Less:			
Firm-years not audited by Big 4 in the current year	(44,861)	(44,861)	(44,861)
Firm-years not audited by Big 4 in the prior year	(5,242)	(5,242)	(5,242)
Total Audited by Big 4 in Current and Prior Year	68,002	68,002	68,002
Less:			
Firm-years with missing data[a]	(42,101)	(42,101)	(57,451)
Firm-years with non-positive *DA*	NA	(12,038)	NA
Total for Unmatched Regressions	25,901	13,863	10,551
Less:			
Firm-years not matched[b]	(13,748)	(7,348)	(4,600)
Total for Propensity Score Matched Regressions	12,153	6,515	5,951
Plus:			
Control firm-years matched to multiple treatment firm-years for main propensity score matched sample[c]	5,799	3,273	3,628
Full Propensity Score Matched Sample	17,952	9,788	9,579

[a] Observations are deleted when there are insufficient data to calculate the audit quality proxy or to estimate the propensity score model using the specialist status of the auditor, control variables, and year fixed effects.
[b] Observations are matched by propensity score, within common support, with replacement, using a caliper distance of 0.03 and a one-to-three match. The propensity of choosing a specialist auditor is predicted using a logistic regression of the auditor's specialist status on variables related to the client's level of earnings quality and year and industry fixed effects.
[c] Matching with replacement allows a single control firm-year to be matched to several treatment firms. Thus, the control firm-year can appear in the sample more than once. We use WLS to address the fact that not all observations in our regressions are unique (Hill and Reiter 2006).

(1) and (2) when absolute discretionary accruals (income-increasing *DA*) are used to measure audit quality. Likewise, 10,551 firm-year observations have sufficient data to estimate Equations (1) and (2) when audit quality is measured using book-tax differences (*BTD*).

Following MM, we examine the effect of industry specialization on audit quality using unmatched samples, as well as samples matched on the propensity of choosing a specialist auditor. We estimate a propensity score using a logistic regression where the dependent variable is the specialist indicator variable and the independent variables are the control variables in Equation (1), including industry and year indicator variables. Appendix A presents our logistic propensity score model. Roberts and Whited (2013) suggest matching with replacement to reduce bias and selecting multiple matches within a reasonable range (caliper) to increase precision, so we match with replacement and allow each treatment firm to match with up to three control firms. We use weighted least squares (WLS) regressions to account for the fact that when matching with replacement, not all of our observations are unique (Dehejia and Wahba 2002; Hill and Reiter 2006; Stuart 2010; DeFond et al. 2017a; Shipman et al. 2017).[21] We cluster standard errors by firm to further address the presence of repeated firms (deHaan, Hodge, and Shevlin 2013; DeFond et al. 2017a). After matching, we are left in Table 2 with 12,153 unique

[21] Ordinary least squares (OLS) has no mechanism for addressing repeated (non-unique) observations. Thus, WLS is required. Observations are matched by propensity score, within common support, using a caliper distance of 0.03. We weight observations as described in Hill and Reiter (2006, 2232). The propensity score matching and weighted least squares codes used in this paper are available on Steven Utke's SSRN page at: https://ssrn.com/abstract=2964990. Code for calculating specialist tenure is also available at: https://ssrn.com/abstract=3212035. We thank Marcelo Coca-Perraillon for assisting with SAS PSM code (see Coca-Perraillon 2006), and William Thomas for making code available online (available at: http://www.biostat.umn.edu/%7Ewill/6470stuff/Class25-12/PSmatching.sas).

observations in the *ADA* sample, 6,515 unique observations in the income-increasing *DA* sample, and 5,951 unique observations in the *BTD* sample. The full PSM samples, which include non-unique control firms that are matched to multiple treatment firms, consist of 17,952, 9,788, and 9,579 observations for the *ADA,* income-increasing *DA,* and *BTD* samples, respectively.

Although the use of PSM facilitates comparison to MM's findings, PSM has drawbacks. Broadly, PSM does not ensure that matched firms are similar (King, Nielsen, Coberley, Pope, and Wells 2011; King and Nielsen 2016; DeFond et al. 2017a; McMullin and Schonberger 2018).[22] As a result, PSM often selects random matches and is unable to resolve the functional form misspecification (i.e., covariate imbalance) that it is intended to resolve (King et al. 2011; King and Nielsen 2016). To illustrate, we review recently published papers in the top three accounting journals that use PSM as their primary method of analysis.[23] Across all studies, 17 percent of covariates exhibit potential imbalance after PSM, indicated by a statistical difference in means and/or medians.[24] Further, PSM involves matching treatment firms to a subset of firms identified as control firms, and discarding the remaining firms. This can result in low-power tests. An additional limitation of PSM that is particularly relevant for our study is that PSM cannot account for covariate imbalance across multiple groups. This is not a problem for MM, who makes one comparison of audit quality between specialist and non-specialist auditors. Our research design, on the other hand, requires us to make three comparisons of audit quality: between seasoned and unseasoned specialists; between seasoned specialists and non-specialist auditors; and between unseasoned specialists and non-specialists. PSM is limited in its ability to address multiple groups, and the fact that PSM discards data makes it difficult to perform PSM in the relatively smaller subgroups that arise when making comparisons within multiple groups.

Recognizing the limitations of PSM for our research setting, we employ two non-PSM matching strategies. The first, entropy balancing, is a covariate balancing method introduced by Hainmueller (2012) and implemented in Hainmueller and Xu (2013). Unlike PSM, entropy balancing is an "equal percent bias reducing" matching method, which ensures that covariate imbalance improves after matching. Entropy balancing achieves this by using an iterative process to reweight control sample observations until the means of the control sample covariates approximately equal those in the treatment sample. In contrast to PSM, entropy balancing does not discard observations, which increases power, and does not generate random matches (King et al. 2011).[25] King and Nielsen (2016) also suggest that matching on specific firm characteristics can be a preferable strategy compared to PSM. Accordingly, our second alternative to PSM is simply to limit the sample to large multinational firms, which likely require auditors with similar capabilities for addressing complex international issues. Limiting the sample to these large firms also addresses the concern that the inclusion of small firms can bias results (Bamber, Christensen, and Gaver 2000; Givoly, Hayn, and Lourie 2016; Hou, Xue, and Zhang 2018).

In summary, our primary tests use the unmatched sample, one-to-three PSM with replacement, entropy balancing, and a sample matched on firm characteristics. We refer to the results from these tests as our "main results."[26] In supplementary analysis, we also investigate the impact of alternative PSM design choices, as suggested by Shipman et al. (2017), varying the number of treatment to control firms (one-to-one, one-to-two, one-to-five, and one-to-ten) and matching with and without replacement.[27] PSM performance tends to deteriorate as the number of covariates increases (DeFond et al. 2017a), so we also use a simplified PSM model based only on size, industry, and year. Finally, we perform robustness checks on our samples that match on specific firm characteristics, examining large firms and multinational firms separately.[28]

---

[22] For example, assume that we estimate a propensity score using *ROA, GROWTH,* and *LEV*. For simplicity, assume that the estimated beta coefficients are all 1. Firm T has a propensity score of 0.25, with *ROA* of 0.10, *GROWTH* of 0.05, and *LEV* of 0.10. Firm C also has a propensity score of 0.25 with *ROA* of −0.20, *GROWTH* of 0.20, and *LEV* of 0.25. Despite identical propensity scores, these firms are clearly very different (King and Nielsen 2016). Examining covariate balance in a PSM sample only assures covariate balance, on average, across all firms, not match by match.

[23] We thank Jonathan Shipman, Quinn Swanquist, and Rob Whited for providing this list of 27 propensity score matching papers reviewed in Shipman et al. (2017).

[24] The potential imbalance ranges from zero to 67 percent of covariates. However, many studies do not report balance statistics, or only test imbalance at the mean or median, but not both. While determining imbalance is subjective (Shipman et al. 2017), we use an objective test based on statistical differences to facilitate our review.

[25] McMullin and Schonberger (2018) provide an excellent discussion and application of entropy balancing in an accounting setting. We implement entropy balancing as described in their paper.

[26] In our main results, we perform entropy balancing in our entire sample in order to facilitate comparisons to MM. However, entropy balancing in the full sample does not resolve the inability to balance covariates across multiple groups, a relevant issue in our setting. In additional robustness tests reported in Section V, we split our sample to focus only on the two groups that we compare in each of our hypotheses, and drop the third group, creating three separate subsamples (seasoned versus unseasoned, seasoned versus non-specialist, and unseasoned versus non-specialist). We then perform entropy balancing within each subsample. Our main results are robust to this approach.

[27] This includes MM's sampling choice of one-to-one matching without replacement.

[28] Entropy balancing leaves little discretion in design choices (McMullin and Schonberger 2018), limiting the usefulness of robustness checks. We use the default settings in Hainmueller and Xu's (2013) "ebalance" Stata macro for our entropy balancing. We thank Jens Hainmueller, Jeff McMullin, Bryce Schonberger, and Yiqing Xu for making entropy balancing code available online.

Descriptive Statistics and Univariate Results

We present descriptive statistics in Table 3, Table 4, and Table 5. Panels A and B of Table 3 report means, medians, and standard deviations for variables in the unmatched (Panel A) and matched (Panel B) *ADA* samples.[29] In both panels, about 20 percent of firm-year observations involve a seasoned specialist auditor, 2.5 percent involve an unseasoned specialist auditor, and the remaining 77.5 percent have a non-specialist auditor. This means that approximately 11 percent of specialist auditors are in their first year as the industry leader (2.5 percent/(2.5 percent + 20 percent)). Table 5, which uses the unmatched *ADA* sample, shows that the representation of each specialist auditor type varies over the sample years. The highest percentage of observations with an unseasoned industry specialist occurs in 2011 (5.37 percent) and the lowest occurs in 2015 (0 percent). We observe similar results in the *BTD* sample (untabulated).

Panels C and D of Table 3 present means and medians for all variables in the *ADA* sample after partitioning observations by auditor specialization. Panel C reports findings for the unmatched sample and Panel D reports findings for the matched sample. Univariate tests of differences in means and medians provide initial insight into our hypotheses.[30] H1 is supported when measuring audit quality using absolute discretionary accruals (*ADA*). That is, the mean and median values of *ADA* are significantly higher (audit quality is worse) when the auditor is an unseasoned specialist compared to when the auditor is a seasoned specialist. This is true in both the unmatched (Panel C) and matched (Panel D) samples. Likewise, H2 is supported when audit quality is measured as *ADA*. Specifically, both mean and median values of *ADA* are significantly lower (audit quality is better) when the auditor is a seasoned specialist compared to cases when the auditor is a non-specialist. This is true in both unmatched (Panel C) and matched (Panel D) samples. In the unmatched sample, mean and median values of *ADA* do not differ significantly between firm-years with non-specialist auditors compared to those with unseasoned specialist auditors. In the matched sample, unseasoned specialists provide *lower* audit quality than non-specialist auditors when audit quality is measured as *ADA*.

Table 4, where *BTD* is the measure of audit quality, reports means and medians after partitioning observations by auditor specialization.[31] Panel A reports findings for the unmatched sample and Panel B reports findings for the matched sample. Univariate tests of differences in means and medians provide consistent support for all of our hypotheses in both matched and unmatched samples. Mean and median values of *BTD* are significantly higher (audit quality is worse) when the auditor is an unseasoned specialist, compared to when the auditor is a seasoned specialist, in both the unmatched (Panel A) and matched (Panel B) samples, consistent with H1. Likewise, in both unmatched and matched samples, mean and median values of *BTD* are significantly higher when the auditor is a non-specialist, compared to cases where there is a seasoned specialist auditor, supporting H2. Finally, there is no significant difference in mean and median values of *BTD* between unseasoned specialists and non-specialist auditors, in both unmatched and matched samples, consistent with H3.

Table 6 presents Pearson and Spearman correlations based on the unmatched *ADA* sample. Correlations involving absolute discretionary accruals (*ADA*) or book-tax differences (*BTD*) suggest that audit quality is higher when the auditor is a specialist (*SPECIALIST*). Further, consistent with our hypotheses, absolute discretionary accruals (*ADA*) and book-tax differences (*BTD*) are negatively correlated with the presence of seasoned specialists (*SEASONED*), but not unseasoned specialists (*UNSEASONED*). Table 6 also reports significant correlations between audit quality variables and the control variables. These variables are taken into account in our multivariate tests, which are discussed in the next section.

V. MULTIVARIATE RESULTS

Replication of Minutti-Meza (2013)

Following MM, we first examine the effect of auditor industry specialization in our unmatched samples.[32] Column (1) in Table 7 presents the results. Panels A and B, C and D, and E and F of Table 7 present results for the *ADA*, income-increasing *DA*, and *BTD*

[29] The statistics for *DA* in Table 3 and Table 6 are based on the larger *ADA* sample, rather than the subsample of firms with income-increasing *DA*. Thus, *DA* in these tables includes both income-increasing and income-decreasing *DA*. We focus on income-increasing *DA* in our multivariate analyses and defer discussion of results to Section V.

[30] Table 3 (Table 4), Panels C and D (A and B) also provide covariate balance details. While assessment of covariate balance is subjective (Shipman et al. 2017), we generally find improved covariate balance in our matched samples. Because we use nine PSM samples, three samples matched on specific firm characteristics, and one entropy balanced sample, we believe it is unlikely that potential imbalance in any one sample drives our results. We also address potential covariate imbalance by including appropriate control variables in the second-stage audit quality model (Shipman et al. 2017).

[31] For the sake of brevity, we do not report descriptive statistics for the full *BTD* sample.

[32] We cannot provide an exact replication of MM because, as discussed in Sections III and IV, we have a different sample period (2003–2015), we limit our sample to clients of Big 4 auditors, and we use alternative measures of audit quality that are more appropriate for our research questions.

TABLE 3

Descriptive Statistics
Absolute Discretionary Accruals (*ADA*) Sample

Panel A: Unmatched Sample[b]

Variable[a]	n	Mean	Median	Std. Dev.
ADA	25,901	0.0608	0.0376	0.0730
DA	25,901	0.0004	0.0043	0.0889
SPECIALIST	25,901	0.1740	0.0000	0.3791
SEASONED	25,901	0.1550	0.0000	0.3619
UNSEASONED	25,901	0.0190	0.0000	0.1364
SEASONED2	25,901	0.1359	0.0000	0.3427
UNSEASONED2	25,901	0.0381	0.0000	0.1915
SEASONED3	25,901	0.1204	0.0000	0.3254
UNSEASONED3	25,901	0.0536	0.0000	0.2252
AU_CHANGE	25,901	0.0298	0.0000	0.1699
LOGMKT	25,901	7.0985	7.1175	2.0651
LEV	25,901	0.2920	0.2586	0.2201
ROAL	25,901	0.0012	0.0368	0.1702
ROA	25,901	0.0037	0.0368	0.1590
LOSS	25,901	0.2793	0.0000	0.4486
CFO	25,901	0.0722	0.0867	0.1352
BTM	25,901	0.5210	0.4513	0.7472
ABS_AC_LAG	25,901	0.0874	0.0620	0.0898
GROWTH	25,901	0.1341	0.0728	0.3711
ALTMAN	25,901	59.7805	3.7499	305.5100
STDEARN	25,901	223.1195	39.1577	565.9071
TENURE	25,901	0.7467	1.0000	0.4349

(continued on next page)

samples, respectively.[33] For brevity, we only report coefficients on our variable of interest, *SPECIALIST*. Coefficients on the control variables are generally consistent with prior research. If industry specialists provide higher audit quality than non-specialist auditors, then the coefficients on *SPECIALIST* will be reliably negative. The results are consistent with these predictions in the *ADA* (Panels A and B of Table 7) and income-increasing *DA* (Panels C and D of Table 7) samples, but not in the *BTD* (Panels E and F of Table 7) sample. Thus, even without matching, an industry specialization proxy based only on market share produces somewhat mixed results. This supports our argument that a purely market share-based proxy is a noisy measure of industry specialization.

MM argues that market share-based measures such as *SPECIALIST* result in a biased test of the relation between auditor industry expertise and audit quality. Acquiring a large client increases an auditor's market share. Client size is correlated with financial reporting quality, which is the basis for our audit quality measures. Thus, client size is correlated with both the variable of interest, (market share-based) *SPECIALIST,* and the dependent variable, audit quality (measured by *ADA,* income-increasing *DA,* and *BTD*), resulting in model misspecification. MM uses PSM to address this misspecification, matching clients of specialist and non-specialist auditors on relevant observable dimensions other than the treatment and outcome variables.

The results of reestimating Equation (1) with our main PSM sample (one-to-three with replacement) are reported in column (2) of Table 7.[34] Again, Panels A and B, C and D, and E and F of Table 7 present results for the *ADA,* income-increasing *DA,* and *BTD* samples, respectively. The coefficient on *SPECIALIST* in Panels E–F (*BTD*) is insignificant, while the coefficients in Panels A–B and C–D (*ADA* and income-increasing *DA*) remain significant, although the statistical significance of the result weakens relative to the results for the unmatched sample reported in column (1). This is largely consistent with MM, who concludes that "[a]fter matching

[33] As discussed earlier, the only signed *DA* measure examined in our main regression results is income-increasing *DA,* following Becker et al. (1998). However, we also performed all analyses presented in Tables 7 through 9 using signed *DA* as the dependent variable. Results are largely consistent with, although somewhat weaker than, the *ADA* and income-increasing *DA* results. Regarding income-decreasing *DA,* we note that prior literature (Reichelt and Wang 2010) finds no evidence that *SPECIALIST* auditors, using our definition, constrain income-decreasing accruals. Consistent with their results, we find no evidence that industry specialists in our sample constrain income-decreasing accruals, and virtually no significant differences across any groups of specialists after performing our partition on seasoning. These results (untabulated) are consistent across matched and unmatched samples.

[34] We use WLS when control firms are selected with replacement, and OLS when control firms are selected without replacement (Stuart 2010).

TABLE 3 (continued)

Panel B: Matched Sample[c]

Variable[a]	n	Mean	Median	Std. Dev.
ADA	17,952	0.0514	0.0332	0.0604
DA	17,952	0.0015	0.0030	0.0733
SPECIALIST	17,952	0.2500	0.0000	0.4330
SEASONED	17,952	0.2228	0.0000	0.4161
UNSEASONED	17,952	0.0272	0.0000	0.1626
LOGMKT	17,952	7.2704	7.3103	1.9651
LEV	17,952	0.3035	0.2785	0.2092
ROAL	17,952	0.0239	0.0380	0.1208
ROA	17,952	0.0258	0.0379	0.1119
LOSS	17,952	0.2176	0.0000	0.4126
CFO	17,952	0.0856	0.0856	0.0945
BTM	17,952	0.5427	0.4765	0.7611
ABS_AC_LAG	17,952	0.0762	0.0579	0.0704
GROWTH	17,952	0.1022	0.0667	0.2562
ALTMAN	17,952	31.8034	3.5414	156.5163
STDEARN	17,952	226.7203	41.1148	583.7800
TENURE	17,952	0.8147	1.0000	0.3886

Panel C: Unmatched Sample Partitioned by Auditor Specialization[b]

Variable[a]		Non-Specialists n = 21,395	Non-Specialists versus Specialists	All Specialists (Seasoned and Unseasoned) n = 4,506	Non-Specialists versus Seasoned Specialists	Seasoned Specialists n = 4,015	Non-Specialists versus Unseasoned Specialists	Unseasoned Specialists n = 491	Seasoned versus Unseasoned Specialists
ADA	Mean	0.063	***	0.050	***	0.049		0.060	***
	Median	0.039	***	0.031	***	0.030		0.039	***
DA	Mean	0.000		0.001		0.001		0.004	
	Median	0.004		0.005		0.005		0.007	
LOGMKT	Mean	7.053	***	7.317	***	7.332		7.190	
	Median	7.061	***	7.361	***	7.379		7.202	
LEV	Mean	0.290	***	0.303	***	0.304		0.292	
	Median	0.253	***	0.283	***	0.286		0.257	
ROAL	Mean	-0.003	***	0.023	***	0.023	***	0.020	
	Median	0.036	***	0.038	***	0.038	***	0.047	*
ROA	Mean	-0.001	***	0.024	***	0.023	***	0.031	
	Median	0.036	***	0.038	***	0.037	***	0.048	***
LOSS	Mean	0.292	***	0.218	***	0.218	***	0.220	
	Median	0.000	***	0.000	***	0.000	***	0.000	
CFO	Mean	0.070	***	0.084	***	0.084	***	0.087	
	Median	0.087	*	0.084		0.084	**	0.089	
BTM	Mean	0.529	***	0.482	***	0.487	***	0.444	
	Median	0.452	*	0.449		0.453	**	0.419	**
ABS_AC_LAG	Mean	0.090	***	0.075	***	0.074	**	0.081	*
	Median	0.064	***	0.056	***	0.055	**	0.057	
GROWTH	Mean	0.141	***	0.100	***	0.096		0.135	***
	Median	0.075	***	0.064	***	0.061		0.078	***
ALTMAN	Mean	64.168	***	38.950	***	38.452	*	43.019	
	Median	3.760	**	3.700	***	3.618	***	4.244	***
STDEARN	Mean	224.741		215.421		218.361		191.376	
	Median	38.822		40.556		39.955		42.100	
TENURE	Mean	0.732	***	0.815	***	0.819	***	0.784	*
	Median	1.000	***	1.000	***	1.000	**	1.000	*

(continued on next page)

TABLE 3 (continued)

Panel D: Matched Sample Partitioned by Auditor Specialization[c]

Variable[a]		Non-Specialists n = 13,464	Non-Specialists versus Specialists	All Specialists (Seasoned and Unseasoned) n = 4,488	Non-Specialists versus Seasoned Specialists	Seasoned Specialists n = 4,000	Non-Specialists versus Unseasoned Specialists	Unseasoned Specialists n = 488	Seasoned versus Unseasoned Specialists
ADA	Mean	0.052	**	0.050	***	0.048	**	0.059	***
	Median	0.034	***	0.031	***	0.030	***	0.039	***
DA	Mean	0.002		0.001		0.001		0.005	
	Median	0.002		0.005		0.005		0.007	
LOGMKT	Mean	7.252	**	7.326	**	7.342		7.199	
	Median	7.288	*	7.365	**	7.387		7.215	
LEV	Mean	0.304		0.302		0.303		0.292	
	Median	0.277		0.282		0.286		0.257	
ROAL	Mean	0.024		0.024		0.025		0.023	
	Median	0.038	**	0.038		0.038	**	0.047	*
ROA	Mean	0.026		0.026		0.025		0.033	
	Median	0.038	*	0.038		0.037	***	0.048	***
LOSS	Mean	0.218		0.218		0.218		0.219	
	Median	0.000		0.000		0.000		0.000	
CFO	Mean	0.086		0.086		0.085		0.088	
	Median	0.086		0.085		0.084		0.089	
BTM	Mean	0.565	***	0.476	***	0.481	***	0.436	
	Median	0.487	***	0.449	***	0.452	***	0.419	**
ABS_AC_LAG	Mean	0.077	**	0.074	***	0.073		0.080	*
	Median	0.058	***	0.055	***	0.055		0.057	
GROWTH	Mean	0.105	***	0.093	***	0.089	*	0.126	***
	Median	0.068	**	0.064	***	0.061	**	0.078	***
ALTMAN	Mean	32.166		30.714		30.243		34.582	
	Median	3.510		3.703		3.620	***	4.243	***
STDEARN	Mean	228.953		220.022		223.230		193.733	
	Median	41.171		40.556		40.028		42.073	
TENURE	Mean	0.814		0.816		0.820		0.785	*
	Median	1.000		1.000		1.000		1.000	*

***, **, * Indicate significant differences in means or medians between the indicated groups at the 1 percent, 5 percent, or 10 percent level, respectively, using two-tailed tests.
[a] Variable definitions are provided in Appendix A. Variables are winsorized at 1 percent and 99 percent.
[b] Panels A and C present results for the unmatched *ADA* sample of 25,901 firm-year observations that have Big 4 auditors in the current and in the prior year and meet certain data requirements for the years 2003–2015.
[c] Panels B and D present results for the propensity score matched *ADA* sample of 17,952 firm-year observations. Observations are matched by propensity score, within common support, with replacement, using a caliper distance of 0.03 and a one-to-three match. The propensity of choosing a specialist auditor is predicted using a logistic regression of the auditor's specialist status on variables related to the client's level of earnings quality and year and industry fixed effects.

clients of specialist and non-specialist auditors on a number of dimensions . . . there is no evidence of differences in commonly used audit-quality proxies between these two groups of auditors . . . the auditor's within-industry market share is not a reliable indicator of audit quality" (Minutti-Meza 2013, 779–780). Columns (5) through (12) report coefficients using alternative PSM specifications, in addition to our main PSM results reported in column (2). When audit quality is measured using *ADA* (Panels A and B of Table 7) or *BTD* (Panels E and F of Table 7), about 56 percent (5/9) of the coefficients on *SPECIALIST* are insignificant, consistent with MM. However, only 11 percent (1/9) of results are consistent with MM in the income-increasing *DA* (Panels C and D of Table 7) sample. We note that many of the alternative PSM design choices examined in columns (5) through (12) lead to results more consistent with MM than those of our main PSM design choice.

Columns (3) and (4) of Table 7 present results using non-PSM matching methods. Results are consistent with the unmatched sample. That is, specialization improves audit quality when measured with *ADA* and income-increasing *DA* (Panels A and C of Table 7), but not *BTD* (Panels E and F of Table 7). This is echoed in columns (13) and (14) (Panels B, D, and F of Table 7), which present results using additional firm characteristic matching approaches.

TABLE 4

Descriptive Statistics
Book-Tax Differences (*BTD*) Sample

Panel A: Unmatched Sample Partitioned by Auditor Specialization[b]

Variable[a]		Non-Specialists n = 8,089	Non-Specialists versus Specialists	All Specialists (Seasoned and Unseasoned) n = 2,462	Non-Specialists versus Seasoned Specialists	Seasoned Specialists n = 2,127	Non-Specialists versus Unseasoned Specialists	Unseasoned Specialists n = 335	Seasoned versus Unseasoned Specialists
BTD	Mean	0.077	**	0.044	***	0.032		0.121	**
	Median	0.039	***	0.028	***	0.026		0.042	*
LOGMKT	Mean	7.202	***	7.630	***	7.649	***	7.510	
	Median	7.179	***	7.576	***	7.573	***	7.604	
LEV	Mean	0.258		0.259		0.257		0.268	
	Median	0.227		0.234		0.232	*	0.240	
ROAL	Mean	0.055	***	0.065	***	0.064	***	0.069	
	Median	0.056	***	0.063	***	0.064	***	0.061	
ROA	Mean	0.074	**	0.076		0.076	**	0.080	
	Median	0.060	***	0.065	***	0.065	***	0.068	
LOSS	Mean	0.011		0.009		0.009		0.009	
	Median	0.000		0.000		0.000		0.000	
CFO	Mean	0.118		0.117		0.117		0.120	
	Median	0.109		0.113		0.112		0.117	
BTM	Mean	0.475	**	0.456	**	0.454		0.468	
	Median	0.414	**	0.395	***	0.390		0.420	
ABS_AC_LAG	Mean	0.074	***	0.064	***	0.063		0.068	
	Median	0.057	***	0.052	***	0.051		0.052	
GROWTH	Mean	0.128	***	0.104	***	0.101		0.122	*
	Median	0.087	***	0.075	***	0.074		0.085	**
ALTMAN	Mean	84.242	***	55.226	***	57.132	**	43.126	
	Median	5.346		5.574	*	5.706		4.846	*
STDEARN	Mean	122.529	***	175.228	***	177.457	**	161.072	
	Median	29.192	***	37.564	***	37.051	***	41.393	
TENURE	Mean	0.790	***	0.841	***	0.847		0.806	*
	Median	1.000	***	1.000	***	1.000		1.000	*

(continued on next page)

Tests of H1

Main Test of H1

An explanation for the insignificant coefficient on *SPECIALIST* in some specifications discussed previously is that *SPECIALIST* is a noisy proxy for industry expertise. Specifically, if the quality of audits produced by unseasoned specialists does not differ from the quality of audits produced by non-specialist auditors, then *SPECIALIST* contains "false positives" where an auditor is coded as an industry expert when, in fact, it is not. We address this problem by separating unseasoned specialists from seasoned specialists.

Table 8, Panels A and B, C and D, and E and F present the results of estimating Equation (2), where *SEASONED* and *UNSEASONED* replace *SPECIALIST*. If, as predicted by H1, seasoned specialists outperform unseasoned specialists, then the coefficient on *UNSEASONED* (β_2) will exceed the coefficient on *SEASONED* (β_1). That is, the difference (*Difference*) between these coefficients ($\beta_1 - \beta_2$) will be reliably negative. Row three (*Difference*) of columns (1) through (4) (Panels A, C, and E of Table 7) presents our main results. H1 is generally supported in all four columns across all three audit quality measures (*ADA*, income-increasing *DA*, and *BTD*).[35] This provides evidence that duration of experience as the dominant

[35] In order to rule out the possibility that unseasoned auditors that are new to their clients drive our results, we remove observations involving auditor changes from both *SEASONED* and *UNSEASONED* and create a separate variable for auditor changes. Our main results for H1, H2, and H3 are robust to this modification.

TABLE 4 (continued)

Panel B: Matched Sample Partitioned by Auditor Specialization[c]

Variable[a]		Non-Specialists n = 7,183	Non-Specialists versus Specialists	All Specialists (Seasoned and Unseasoned) n = 2,396	Non-Specialists versus Seasoned Specialists	Seasoned Specialists n = 2,069	Non-Specialists versus Unseasoned Specialists	Unseasoned Specialists n = 327	Seasoned versus Unseasoned Specialists
BTD	Mean	0.073	*	0.047	***	0.036		0.123	***
	Median	0.046	***	0.028	***	0.025		0.042	**
LOGMKT	Mean	7.522	***	7.643	***	7.661		7.524	
	Median	7.423	***	7.591	***	7.583		7.604	
LEV	Mean	0.268	**	0.259	**	0.257		0.269	
	Median	0.234		0.235	*	0.234		0.244	
ROAL	Mean	0.066		0.066		0.065		0.070	
	Median	0.060	*	0.063	*	0.064		0.060	
ROA	Mean	0.074	*	0.076		0.076	*	0.080	
	Median	0.061	***	0.065	***	0.065	**	0.068	
LOSS	Mean	0.008		0.009		0.009		0.009	
	Median	0.000		0.000		0.000		0.000	
CFO	Mean	0.116		0.117		0.117		0.120	
	Median	0.105	***	0.113	**	0.112		0.117	
BTM	Mean	0.488	***	0.456	***	0.454		0.472	
	Median	0.426	***	0.394	***	0.388		0.422	*
ABS_AC_LAG	Mean	0.063		0.063		0.063		0.066	
	Median	0.051		0.052		0.051		0.053	
GROWTH	Mean	0.107		0.101	**	0.098		0.120	**
	Median	0.077		0.075	**	0.074		0.085	**
ALTMAN	Mean	44.245		41.660		43.149		32.236	
	Median	5.200	***	5.560	***	5.700		4.798	**
STDEARN	Mean	162.023	*	182.782	**	184.663		170.878	
	Median	36.772		38.017		37.477		41.460	
TENURE	Mean	0.841		0.845		0.849		0.820	
	Median	1.000		1.000		1.000		1.000	

***, **, * Indicate significant differences in means or medians between the indicated groups at the 1 percent, 5 percent, or 10 percent level, respectively, using two-tailed tests.

[a] Variable definitions are provided in Appendix A. Variables are winsorized at 1 percent and 99 percent.

[b] Panel A presents results for the unmatched *BTD* sample of 10,551 firm-year observations that have Big 4 auditors in the current and in the prior year and meet certain data requirements for the years 2003–2015.

[c] Panel B presents results for the propensity score matched *BTD* sample of 9,579 firm-year observations. Observations are matched by propensity score, within common support, with replacement, using a caliper distance of 0.03 and a one-to-three match. The propensity of choosing a specialist auditor is predicted using a logistic regression of the auditor's specialist status on variables related to the client's level of earnings quality and year and industry fixed effects.

industry auditor is associated with audit quality. Importantly, this experience effect is apparent even after matching on client attributes.[36]

We assess the robustness of this finding by varying the matching approach, as well as the length of the seasoning period. Columns (5) through (14) in Table 8 show the results of estimating Equation (2) using different matching strategies. The tenor of our findings is unchanged: *Difference* is generally negative across sampling approaches for all three audit quality measures. Row six reports the estimated *Difference* after redefining unseasoned auditors as those in their first two years as industry experts (*UNSEASONED2*). Likewise, row nine reports the estimated *Difference* when auditors in their first three years as industry experts are considered unseasoned (*UNSEASONED3*). In general, lengthening the seasoning period does not alter our conclusions across the two accruals-based audit quality proxies. However, *Difference* is generally insignificant in the *BTD*

[36] Our accruals sample includes some non-U.S. firms, which, by definition, cannot have specialist auditors. Removing these firms from the sample does not change our inferences.

TABLE 5
Specialist Auditor Type by Year[a]

Year	Number of Observations	Non-Specialist Auditor (%)	Unseasoned Specialist Auditor (%)	Seasoned Specialist Auditor (%)
2003	2,448	80.60%	1.80%	17.61%
2004	2,247	82.60%	1.60%	15.80%
2005	2,097	83.36%	3.67%	12.97%
2006	2,092	82.17%	2.25%	15.58%
2007	2,031	83.26%	0.98%	15.76%
2008	1,933	83.14%	0.05%	16.81%
2009	1,759	84.42%	1.36%	14.21%
2010	1,791	85.99%	0.78%	13.23%
2011	1,806	81.95%	5.37%	12.68%
2012	1,921	82.04%	2.13%	15.83%
2013	1,873	81.74%	3.10%	15.16%
2014	2,000	81.20%	1.60%	17.20%
2015	1,903	82.24%	0.00%	17.76%
2003–2015	25,901	82.60%	1.90%	15.50%

[a] The sample consists of 25,901 firm-year observations that have Big 4 auditors in the current and in the prior year and meet data requirements for the *ADA* sample for the years 2003–2015.

sample when the seasoning period is greater than one year (Panel C). We return to this point below in our examination of the length of the seasoning period.[37]

Alternative Test of H1

An alternative approach for testing H1 is to restrict the sample to specialist clients, and directly compare unseasoned specialists to seasoned specialists using Equation (3).

$$\begin{aligned} AQ_{it} = {} & \alpha_0 + \beta_1 UNSEASONED_{it} + \beta_2 LOGMKT_{it} + \beta_3 LEV_{it} + \beta_4 ROAL_{it} + \beta_5 ROA_{it} + \beta_6 LOSS_{it} + \beta_7 CFO_{it} + \beta_8 BTM_{it} \\ & + \beta_9 ABS\_AC\_LAG_{it} + \beta_{10} GROWTH_{it} + \beta_{11} ALTMAN_{it} + \beta_{12} STDEARN_{it} + \beta_{13} TENURE_{it} + \beta_t YEAR\_FE_t + v_{it} \end{aligned} \tag{3}$$

An advantage of this approach is that it does not require the design choices associated with matching. H1 is supported if β_1 is positive.

The results of estimating Equation (3) support the inferences from Table 8. We find that when audit quality is measured using *ADA* or income-increasing *DA,* β_1 is positive for seasoning periods of one, two, and three years (untabulated). For *BTD,* β_1 is significant for a one-year seasoning period, but insignificant for seasoning periods of two or three years (untabulated), echoing the results in Table 8. On balance, these results support H1.

Length of the Seasoning Period

To further investigate the length of the seasoning process, we return to Equation (2) and replace *UNSEASONED* with three dummy variables that switch on when an auditor is in their first, second, or third year as a specialist (*UNSEASONED1_Dum, UNSEASONED2_Dum,* and *UNSEASONED3_Dum,* respectively). *SEASONED3* is coded 1 when the auditor has been the industry leader for more than three years. This results in Equation (4).

$$\begin{aligned} AQ_{it} = {} & \alpha_0 + \beta_1 SEASONED3_{it} + \beta_2 UNSEASONED1\_Dum_{it} + \beta_3 UNSEASONED2\_Dum_{it} + \beta_4 UNSEASONED3\_Dum_{it} \\ & + \beta_5 LOGMKT_{it} + \beta_6 LEV_{it} + \beta_7 ROAL_{it} + \beta_8 ROA_{it} + \beta_9 LOSS_{it} + \beta_{10} CFO_{it} + \beta_{11} BTM_{it} + \beta_{12} ABS\_AC\_LAG_{it} \\ & + \beta_{13} GROWTH_{it} + \beta_{14} ALTMAN_{it} + \beta_{15} STDEARN_{it} + \beta_{16} TENURE_{it} + \beta_t YEAR\_FE_t + v_{it} \end{aligned} \tag{4}$$

[37] Lengthening the seasoning period increases the percentage of observations that are classified as unseasoned specialists in each regression. For the unmatched *ADA* sample, 1.90 percent of observations are classified as unseasoned specialists under the one-year seasoning period (*UNSEASONED*), compared to 3.81 percent for the two-year period (*UNSEASONED2*) and 5.36 percent for the three-year period (*UNSEASONED3*).

TABLE 6

Correlations
Unmatched Sample[a,b]

Panel A: Correlation Variables *ADA* to *ROAL*

Variable[c]	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1) *ADA*		**−0.101**	**−0.065**	**−0.068**	**−0.070**	−0.002	**−0.268**	**0.026**	**−0.285**
(2) *DA*	**0.036**		**−0.069**	0.004	0.002	0.006	**0.032**	**−0.025**	**−0.047**
(3) *BTD*	**−0.029**	**−0.052**		**−0.021**	**−0.029**	0.015	**−0.031**	**0.070**	−0.013
(4) *SPECIALIST*	**−0.077**	0.001	**−0.032**		**0.933**	**0.303**	**0.049**	**0.023**	**0.058**
(5) *SEASONED*	**−0.082**	−0.002	**−0.038**	**0.933**		**−0.060**	**0.048**	**0.024**	**0.054**
(6) *UNSEASONED*	0.002	0.006	0.008	**0.303**	**−0.060**		0.006	0.000	**0.015**
(7) *LOGMKT*	**−0.251**	0.010	**−0.052**	**0.049**	**0.050**	0.006		**−0.044**	**0.380**
(8) *LEV*	**−0.045**	−0.012	**0.103**	**0.040**	**0.040**	0.005	**0.038**		**−0.089**
(9) *ROAL*	**−0.138**	**−0.045**	**−0.074**	**0.040**	**0.033**	**0.022**	**0.403**	**−0.130**	
(10) *ROA*	**−0.138**	**0.170**	**−0.089**	**0.038**	**0.029**	**0.028**	**0.432**	**−0.161**	**0.679**
(11) *LOSS*	**0.230**	**−0.184**	**−0.028**	**−0.062**	**−0.058**	**−0.018**	**−0.421**	**0.063**	**−0.523**
(12) *CFO*	**−0.150**	**−0.330**	**0.037**	0.012	0.008	0.012	**0.369**	**−0.078**	**0.556**
(13) *BTM*	**−0.064**	**0.067**	**0.094**	−0.012	−0.007	**−0.015**	**−0.178**	**−0.140**	**−0.098**
(14) *ABS_AC_LAG*	**0.154**	**−0.134**	**0.069**	**−0.060**	**−0.059**	**−0.012**	**−0.191**	**0.036**	**−0.286**
(15) *GROWTH*	**0.049**	**0.014**	**0.020**	**−0.032**	**−0.037**	0.010	**0.066**	**−0.018**	**0.066**
(16) *ALTMAN*	−0.003	0.011	**−0.140**	**−0.015**	**−0.022**	**0.018**	**0.134**	**−0.767**	**0.367**
(17) *STDEARN*	**−0.130**	0.000	0.008	0.002	0.003	−0.001	**0.688**	**0.154**	**0.111**
(18) *TENURE*	**−0.086**	0.010	−0.016	**0.072**	**0.071**	0.012	**0.148**	**−0.018**	**0.099**

Panel B: Correlation Variables *ROA* to *TENURE*

Variable[c]	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
(1) *ADA*	**−0.324**	**0.255**	**−0.243**	**−0.075**	**0.285**	**0.160**	**0.064**	**−0.062**	**−0.102**
(2) *DA*	**0.276**	**−0.220**	**−0.246**	**0.053**	**−0.071**	0.005	−0.009	**0.017**	**0.017**
(3) *BTD*	**−0.053**	**−0.034**	**0.060**	**0.019**	**0.048**	**0.022**	**−0.036**	−0.010	−0.002
(4) *SPECIALIST*	**0.059**	**−0.062**	**0.040**	**−0.024**	**−0.065**	**−0.042**	**−0.031**	−0.006	**0.072**
(5) *SEASONED*	**0.053**	**−0.058**	**0.036**	**−0.020**	**−0.064**	**−0.044**	**−0.030**	−0.004	**0.071**
(6) *UNSEASONED*	**0.024**	**−0.018**	**0.015**	**−0.014**	−0.010	0.000	−0.008	−0.008	0.012
(7) *LOGMKT*	**0.418**	**−0.423**	**0.372**	**−0.070**	**−0.245**	−0.009	**−0.037**	**0.439**	**0.143**
(8) *LEV*	**−0.106**	**0.111**	**−0.064**	**−0.244**	**0.075**	**0.027**	**−0.245**	0.002	**−0.029**
(9) *ROAL*	**0.715**	**−0.491**	**0.696**	**0.113**	**−0.475**	**−0.143**	**−0.015**	**0.054**	**0.104**
(10) *ROA*		**−0.681**	**0.779**	**0.069**	**−0.279**	**−0.027**	**0.013**	**0.079**	**0.094**
(11) *LOSS*	**−0.777**		**−0.496**	0.003	**0.261**	−0.010	0.007	**−0.051**	**−0.089**
(12) *CFO*	**0.646**	**−0.495**		**0.024**	**−0.163**	**−0.056**	0.003	**0.086**	**0.084**
(13) *BTM*	**−0.185**	**0.018**	**−0.179**		**−0.122**	**−0.045**	**−0.038**	−0.007	0.004
(14) *ABS_AC_LAG*	**−0.176**	**0.222**	**0.070**	**−0.111**		**0.043**	**0.041**	**0.030**	**−0.097**
(15) *GROWTH*	**0.213**	**−0.158**	**0.138**	**−0.146**	**−0.014**		**0.068**	**−0.081**	**−0.139**
(16) *ALTMAN*	**0.455**	**−0.283**	**0.316**	**−0.182**	**−0.066**	**0.178**		**−0.052**	**−0.037**
(17) *STDEARN*	**0.127**	**−0.116**	**0.139**	−0.007	−0.006	**−0.136**	**−0.186**		**0.035**
(18) *TENURE*	**0.079**	**−0.089**	**0.065**	0.011	**−0.081**	**−0.109**	**0.028**	**0.121**	

[a] This sample consists of 25,901 firm-year observations (10,551 for *BTD*) that have Big 4 auditors in the current and prior years and meet certain data requirements for the years 2003–2015.
[b] Pearson (Spearman) correlations are above (below) the diagonal. Bold coefficients are significant at the 5 percent level using two-tailed tests.
[c] Variables are defined in Appendix A.

Table 9 reports our main results from estimating Equation (3) using the three alternative measures of audit quality (*ADA* and income-increasing *DA* in Panel A of Table 7, *BTD* in Panel B of Table 7). The estimated differences in the coefficients on the seasoned and unseasoned dummy variables are reported in rows 5 through 7 of Table 9. A significantly negative difference means that an unseasoned specialist provides lower audit quality than a seasoned specialist. As in Table 8, the results vary somewhat with the audit quality measure. In general, however, they suggest that the seasoning process extends beyond the first

TABLE 7

Estimated Coefficients and p-values from the Regressions of Audit Quality Measures on Auditor Industry Specialization and Control Variables

Panel A: Absolute Discretionary Accruals

$$ADA_{it} = \alpha_0 + \beta_1 SPECIALIST_{it} + \beta_2 LOGMKT_{it} + \beta_3 LEV_{it} + \beta_4 ROAL_{it} + \beta_5 ROA_{it} + \beta_6 LOSS_{it} + \beta_7 CFO_{it} + \beta_8 BTM_{it} + \beta_9 ABS\_AC\_LAG_{it} + \beta_{10} GROWTH_{it} + \beta_{11} ALTMAN_{it} + \beta_{12} STDEARN_{it} + \beta_{13} TENURE_{it} + \beta_t YEAR\_FE_t + v_{it}$$

		Main Results				Alternative Matching Results		
Variable	**Pred.**	**(1) Unmatched Sample Coeff. (p-value)**	**(2) Main Matched Sample Coeff. (p-value)**	**(3) Entropy Balanced Coeff. (p-value)**	**(4) MNCs with MVE > 500 Coeff. (p-value)**	**(5) 1:1 with Replacement Coeff. (p-value)**	**(6) 1:2 with Replacement Coeff. (p-value)**	**(7) 1:5 with Replacement Coeff. (p-value)**
SPECIALIST	−	−0.005***	−0.002*	−0.006***	−0.003**	−0.003**	−0.002	−0.001
		(0.000)	(0.088)	(0.000)	(0.011)	(0.015)	(0.103)	(0.185)
Control Variables		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes
n		25,901	12,153	25,901	10,379	7,910	10,340	14,735
Adjusted R^2		0.191	0.153	0.148	0.132	0.158	0.151	0.142

Panel B: Absolute Discretionary Accruals (continued)

		Alternative Matching Results						
Variable	**Pred.**	**(8) 1:10 with Replacement Coeff. (p-value)**	**(9) 1:1 without Replacement Coeff. (p-value)**	**(10) 1:2 without Replacement Coeff. (p-value)**	**(11) 1:3 without Replacement Coeff. (p-value)**	**(12) Size-Industry-Year Coeff. (p-value)**	**(13) MVE > $500 Coeff. (p-value)**	**(14) MNCs Coeff. (p-value)**
SPECIALIST	−	−0.001	−0.002**	−0.001	−0.002**	−0.001	−0.005***	−0.003***
		(0.155)	(0.030)	(0.199)	(0.046)	(0.123)	(0.000)	(0.003)
Control Variables		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes
n		18,393	8,924	12,188	14,198	12,281	17,371	14,145
Adjusted R^2		0.143	0.163	0.151	0.159	0.154	0.149	0.168

***, **, * Indicate significance at the 1 percent, 5 percent, and 10 percent levels, respectively.
Huber-White robust standard errors are clustered by firm and are used to control for heteroscedasticity and serial correlation. When predictions are made, p-values are one-tailed. We present results for a number of different matching techniques in columns (2)–(14). Where applicable, the number after the colon in the column heading represents the number of control firms. We match with replacement (with Replacement) and without replacement (without Replacement), perform a reduced match using only size, industry, and year (1:3 with replacement), and perform exact matches with firms with market value of equity (MVE) greater than $500 million and with multinational firms (MNCs). Regressions for samples matched with replacement are estimated using WLS (Hill and Reiter 2006).
Variables are defined in Appendix A.

(continued on next page)

year. Results using *BTD* as the measure of audit quality (columns (9) to (12)) indicate that seasoning is complete in the second year as market share leader; this explains the insignificant *Difference* in Table 8 for two- and three- year seasoning periods when using *BTD* to measure audit quality. Results using income-increasing *DA* (columns (5) to (8)) indicate that the auditor does not perform as a true expert until the third year. Results using *ADA* (columns (1) to (4)) suggest that even after three years, unseasoned auditors do not perform as true experts. However, in the *ADA* sample, the size and statistical significance of the difference in audit quality between seasoned and unseasoned auditors is lower after two or three years compared to the initial year of industry specialization.

TABLE 7 (continued)

Panel C: Income-Increasing (Positive) Discretionary Accruals

$$DA_{it} = \alpha_0 + \beta_1 SPECIALIST_{it} + \beta_2 LOGMKT_{it} + \beta_3 LEV_{it} + \beta_4 ROAL_{it} + \beta_5 ROA_{it} + \beta_6 LOSS_{it} + \beta_7 CFO_{it} + \beta_8 BTM_{it} + \beta_9 ABS\_AC\_LAG_{it} + \beta_{10} GROWTH_{it} + \beta_{11} ALTMAN_{it} + \beta_{12} STDEARN_{it} + \beta_{13} TENURE_{it} + \beta_t YEAR\_FE_t + v_{it}$$

		Main Results				Alternative Matching Results		
Variable	**Pred.**	**(1) Unmatched Sample Coeff. (p-value)**	**(2) Main Matched Sample Coeff. (p-value)**	**(3) Entropy Balanced Coeff. (p-value)**	**(4) MNCs with MVE > 500 Coeff. (p-value)**	**(5) 1:1 with Replacement Coeff. (p-value)**	**(6) 1:2 with Replacement Coeff. (p-value)**	**(7) 1:5 with Replacement Coeff. (p-value)**
SPECIALIST	−	−0.005***	−0.002**	−0.005***	−0.003***	−0.002*	−0.002*	−0.002*
		(0.000)	(0.047)	(0.000)	(0.006)	(0.088)	(0.050)	(0.067)
Control Variables		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes
n		13,863	6,515	13,863	5,637	4,254	5,532	7,965
Adjusted R^2		0.569	0.518	0.527	0.467	0.514	0.532	0.504

Panel D: Income-Increasing (Positive) Discretionary Accruals (continued)

		Alternative Matching Results						
Variable	**Pred.**	**(8) 1:10 with Replacement Coeff. (p-value)**	**(9) 1:1 without Replacement Coeff. (p-value)**	**(10) 1:2 without Replacement Coeff. (p-value)**	**(11) 1:3 without Replacement Coeff. (p-value)**	**(12) Size-Industry-Year Coeff. (p-value)**	**(13) MVE > $500 Coeff. (p-value)**	**(14) MNCs Coeff. (p-value)**
SPECIALIST	−	−0.002*	−0.003**	−0.002*	−0.002**	−0.002	−0.005***	−0.005***
		(0.058)	(0.019)	(0.066)	(0.024)	(0.102)	(0.000)	(0.000)
Control Variables		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes
n		9,952	4,762	6,362	7,497	6,637	9,439	7,549
Adjusted R^2		0.504	0.513	0.530	0.529	0.533	0.497	0.532

***, **, * Indicate significance at the 1 percent, 5 percent, and 10 percent levels, respectively.
Huber-White robust standard errors are clustered by firm and are used to control for heteroscedasticity and serial correlation. When predictions are made, p-values are one-tailed. We present results for a number of different matching techniques in columns (2)–(14). Where applicable, the number after the colon in the column heading represents the number of control firms. We match with replacement (with Replacement) and without replacement (without Replacement), perform a reduced match using only size, industry, and year (1:3 with replacement), and perform exact matches with firms with market value of equity (MVE) greater than $500 million and with multinational firms (MNCs). Regressions for samples matched with replacement are estimated using WLS (Hill and Reiter 2006).
Variables are defined in Appendix A.

(continued on next page)

An alternative approach for assessing the length of the seasoning period is to use a continuous measure of specialist tenure, rather than the indicator variables described above.[38] In the unmatched *ADA* and income-increasing *DA* samples, we find a negative coefficient on the continuous specialist tenure variable (untabulated), indicating that audit quality is positively related to specialist tenure. In the PSM (entropy balanced) sample, this result is apparent in the *ADA* and *BTD* (*ADA* and income-increasing *DA*) samples. We suggest that audit quality is unlikely to increase linearly with specialist tenure, and this could explain these mixed results.

[38] We note that a concurrent working paper, Barnes (2015), examines specialist tenure using a continuous measure of industry specialization and unmatched samples.

TABLE 7 (continued)

Panel E: Book-Tax Differences

$$BTD_{it} = \alpha_0 + \beta_1 SPECIALIST_{it} + \beta_2 LOGMKT_{it} + \beta_3 LEV_{it} + \beta_4 ROAL_{it} + \beta_5 ROA_{it} + \beta_6 LOSS_{it} + \beta_7 CFO_{it} + \beta_8 BTM_{it} + \beta_9 ABS\_AC\_LAG_{it} + \beta_{10} GROWTH_{it} + \beta_{11} ALTMAN_{it} + \beta_{12} STDEARN_{it} + \beta_{13} TENURE_{it} + \beta_t YEAR\_FE_t + v_{it}$$

		Main Results				Alternative Matching Results		
Variable	Pred.	(1) Unmatched Sample Coeff. (p-value)	(2) Main Matched Sample Coeff. (p-value)	(3) Entropy Balanced Coeff. (p-value)	(4) MNCs with MVE > 500 Coeff. (p-value)	(5) 1:1 with Replacement Coeff. (p-value)	(6) 1:2 with Replacement Coeff. (p-value)	(7) 1:5 with Replacement Coeff. (p-value)
SPECIALIST	−	−0.020 (0.132)	−0.022 (0.109)	−0.019 (0.144)	−0.003 (0.444)	−0.017 (0.200)	−0.023 (0.104)	−0.021 (0.118)
Control Variables		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes
n		10,551	5,951	10,551	6,112	4,097	5,149	6,931
Adjusted R^2		0.069	0.080	0.067	0.043	0.085	0.081	0.072

Panel F: Book-Tax Differences (continued)

		Alternative Matching Results						
Variable	Pred.	(8) 1:10 with Replacement Coeff. (p-value)	(9) 1:1 without Replacement Coeff. (p-value)	(10) 1:2 without Replacement Coeff. (p-value)	(11) 1:3 without Replacement Coeff. (p-value)	(12) Size-Industry-Year Coeff. (p-value)	(13) MVE > \$500 Coeff. (p-value)	(14) MNCs Coeff. (p-value)
SPECIALIST	−	−0.022* (0.096)	−0.027* (0.065)	−0.024* (0.091)	−0.021 (0.126)	−0.033** (0.042)	−0.015 (0.199)	−0.020 (0.165)
Control Variables		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes
n		8,266	4,572	5,970	6,799	6,024	7,882	7,515
Adjusted R^2		0.067	0.076	0.071	0.076	0.072	0.092	0.036

***, **, * Indicate significance at the 1 percent, 5 percent, and 10 percent levels, respectively.
Huber-White robust standard errors are clustered by firm and are used to control for heteroscedasticity and serial correlation. When predictions are made, p-values are one-tailed. We present results for a number of different matching techniques in columns (2)–(14). Where applicable, the number after the colon in the column heading represents the number of control firms. We match with replacement (with Replacement) and without replacement (without Replacement), perform a reduced match using only size, industry, and year (1:3 with replacement), and perform exact matches with firms with market value of equity (MVE) greater than \$500 million and with multinational firms (MNCs). Regressions for samples matched with replacement are estimated using WLS (Hill and Reiter 2006). Industry fixed effects are included by two-digit SIC in the *BTD* regressions.
Variables are defined in Appendix A.

Tests of H2

If seasoned specialists outperform non-specialists, as predicted by H2, then the coefficient on *SEASONED* (β_1) in Equation (2) will be negative. In general, the results in Table 8 support H2, although they are sensitive to the audit quality measure, the matching strategy, and the seasoning period. When audit quality is measured as *ADA* (Panels A and B in Table 8) and income-increasing *DA* (Panels C and D in Table 8), β_1 is reliably negative for all seasoning periods using the matched and unmatched samples.[39] When

[39] Again, the main results in the accruals sample are generally robust to excluding non-U.S. firms, although the statistical significance of *SEASONED* weakens slightly in some of the PSM samples.

TABLE 8

Estimated Coefficients and p-values from the Regressions of Audit Quality Measures on Auditor Industry Specialization and Control Variables Where Specialists are Classified as Seasoned or Unseasoned

Panel A: Absolute Discretionary Accruals

$$ADA_{it} = \alpha_0 + \beta_1 SEASONED_{it} + \beta_2 UNSEASONED_{it} + \beta_3 LOGMKT_{it} + \beta_4 LEV_{it} + \beta_5 ROAL_{it} + \beta_6 ROA_{it} + \beta_7 LOSS_{it} + \beta_8 CFO_{it} + \beta_9 BTM_{it} + \beta_{10} ABS\_AC\_LAG_{it} + \beta_{11} GROWTH_{it} + \beta_{12} ALTMAN_{it} + \beta_{13} STDEARN_{it} + \beta_{14} TENURE_{it} + \beta_t YEAR\_FE_t + v_{it}$$

		Main Results				Alternative Matching Results		
Variable	**Pred.**	**(1) Unmatched Sample Coeff. (p-value)**	**(2) Main Matched Sample Coeff. (p-value)**	**(3) Entropy Balanced Coeff. (p-value)**	**(4) MNCs with MVE > 500 Coeff. (p-value)**	**(5) 1:1 with Replacement Coeff. (p-value)**	**(6) 1:2 with Replacement Coeff. (p-value)**	**(7) 1:5 with Replacement Coeff. (p-value)**
Primary Analysis								
SEASONED	–	–0.006***	–0.003**	–0.007***	–0.004***	–0.004***	–0.003**	–0.002**
		(0.000)	(0.014)	(0.000)	(0.003)	(0.002)	(0.020)	(0.041)
UNSEASONED	?	0.004	0.007**	0.003	0.002	0.006*	0.007**	0.007**
		(0.208)	(0.013)	(0.263)	(0.534)	(0.054)	(0.020)	(0.011)
Difference	–	–0.010***	–0.010***	–0.010***	–0.005**	–0.010***	–0.009***	–0.009***
		(0.000)	(0.000)	(0.000)	(0.027)	(0.000)	(0.001)	(0.001)
Alternative Seasoning Periods								
SEASONED2	–	–0.007***	–0.003***	–0.007***	–0.004***	–0.004***	–0.003**	–0.003**
		(0.000)	(0.008)	(0.000)	(0.002)	(0.001)	(0.012)	(0.022)
UNSEASONED2	?	0.001	0.004*	0.000	0.000	0.002	0.003	0.004**
		(0.816)	(0.085)	(0.918)	(0.864)	(0.341)	(0.109)	(0.047)
Difference	–	–0.008***	–0.007***	–0.007***	–0.004**	–0.007***	–0.007***	–0.007***
		(0.001)	(0.001)	(0.001)	(0.033)	(0.002)	(0.002)	(0.001)
SEASONED3	–	–0.008***	–0.004***	–0.008***	–0.005***	–0.005***	–0.004***	–0.003***
		(0.000)	(0.002)	(0.000)	(0.000)	(0.000)	(0.003)	(0.005)
UNSEASONED3	?	0.000	0.003*	–0.001	0.001	0.002	0.003*	0.004**
		(0.972)	(0.061)	(0.743)	(0.799)	(0.327)	(0.075)	(0.026)
Difference	–	–0.008***	–0.007***	–0.007***	–0.005***	–0.007***	–0.007***	–0.007***
		(0.000)	(0.000)	(0.000)	(0.009)	(0.000)	(0.000)	(0.000)
Control Variables		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes
n		25,901	12,153	25,901	10,379	7,910	10,340	14,735
Adjusted R^2		0.191	0.153	0.148	0.132	0.158	0.151	0.142

(continued on next page)

BTD is the audit quality measure (Panels E and F in Table 8), β_1 is negative for most samples and seasoning periods. If we gauge economic significance as the average β_1 across all 14 specifications, *SEASONED* auditors reduce *ADA* by 6.17 percent (–0.004/0.061), income-increasing *DA* by 6.43 percent (–0.004/0.057), and *BTD* by 42.42 percent (–0.029/0.069). We note that *BTD* has a much higher standard deviation than the other measures, resulting in larger, but less precise, estimates.

In sum, we find that *SEASONED* specialists generally provide higher audit quality than non-specialists across all of the audit quality measures in unmatched samples. The same is true for most of the matched samples. This is despite the fact that PSM likely reduces the power of our tests (Shipman et al. 2017) and may be unnecessary in a sample of Big 4 clients with fairly homogeneous characteristics (Francis 2011). Entropy balancing, which is not subject to the random matching or power reduction problems of PSM, generally provides stronger support for our hypotheses than PSM. Our finding of an industry specialization effect in matched samples after separating unseasoned from seasoned specialists suggests that the MM result is

TABLE 8 (continued)

Panel B: Absolute Discretionary Accruals (continued)

		Alternative Matching Results						
Variable	**Pred.**	**(8) 1:10 with Replacement Coeff. (p-value)**	**(9) 1:1 without Replacement Coeff. (p-value)**	**(10) 1:2 without Replacement Coeff. (p-value)**	**(11) 1:3 without Replacement Coeff. (p-value)**	**(12) Size-Industry-Year Coeff. (p-value)**	**(13) MVE > $500 Coeff. (p-value)**	**(14) MNCs Coeff. (p-value)**
Primary Analysis								
SEASONED	–	−0.002**	−0.003***	−0.002**	−0.003***	−0.002**	−0.006***	−0.004***
		(0.029)	(0.004)	(0.044)	(0.006)	(0.024)	(0.000)	(0.000)
UNSEASONED	?	0.007***	0.006**	0.007**	0.006**	0.007**	0.002	0.003
		(0.009)	(0.032)	(0.011)	(0.036)	(0.013)	(0.458)	(0.376)
Difference	–	−0.010***	−0.010***	−0.009***	−0.009***	−0.010***	−0.008***	−0.007***
		(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.010)
Alternative Seasoning Periods								
SEASONED2	–	−0.003**	−0.004***	−0.002**	−0.004***	−0.003**	−0.007***	−0.005***
		(0.015)	(0.002)	(0.028)	(0.002)	(0.012)	(0.000)	(0.000)
UNSEASONED2	?	0.004**	0.003	0.004*	0.003	0.004*	0.000	0.001
		(0.045)	(0.245)	(0.068)	(0.123)	(0.067)	(0.893)	(0.720)
Difference	–	−0.007***	−0.006***	−0.006***	−0.007***	−0.007***	−0.007***	−0.006**
		(0.001)	(0.002)	(0.003)	(0.002)	(0.001)	(0.000)	(0.010)
SEASONED3	–	−0.003***	−0.005***	−0.003***	−0.004***	−0.004***	−0.008***	−0.006***
		(0.003)	(0.001)	(0.010)	(0.001)	(0.003)	(0.000)	(0.000)
UNSEASONED3	?	0.004**	0.002	0.004*	0.003	0.004*	0.000	0.000
		(0.027)	(0.218)	(0.056)	(0.137)	(0.054)	(0.844)	(0.912)
Difference	–	−0.007***	−0.007***	−0.007***	−0.007***	−0.007***	−0.007***	−0.006***
		(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.005)
Control Variables		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes
n		18,393	8,924	12,188	14,198	12,281	17,371	14,145
Adjusted R^2		0.143	0.163	0.151	0.159	0.154	0.149	0.168

***, **, * Indicate significance at the 1 percent, 5 percent, and 10 percent levels, respectively.
Huber-White robust standard errors are clustered by firm and are used to control for heteroscedasticity and serial correlation. When predictions are made, p-values are one-tailed. We present results for a number of different matching techniques in columns (2)–(14). Where applicable, the number after the colon in the column heading represents the number of control firms. We match with replacement (with Replacement) and without replacement (without Replacement), perform a reduced match using only size, industry, and year (1:3 with replacement), and perform exact matches with firms with market value of equity (MVE) greater than $500 million and with multinational firms (MNCs). Regressions for samples matched with replacement are estimated using WLS (Hill and Reiter 2006).
Variables are defined in Appendix A.

(continued on next page)

attributable to mixing seasoned specialists with unseasoned specialists, who act more like non-specialists. We contend that testing the association between auditor industry expertise and audit quality requires consideration of specialist tenure. However, we acknowledge that our tests are not robust across all audit quality proxies, as discussed below in our analysis of alternative audit quality measures.

Analysis of H3

If the quality of audits produced by unseasoned specialists does not differ from the quality of audits produced by non-specialists, as stated in H3, then the coefficients on *UNSEASONED* (β_2) in Equation (2) will be statistically indistinguishable from zero. In general, we find little evidence in Panels A and B, C and D, and E and F of Table 8 that unseasoned specialist audit quality differs from that of non-specialists. There are a few significantly positive values of β_2 in Panels A and B (*ADA*) and E (*BTD*) in Table 8, indicating that audit quality is *lower* when the auditor is an unseasoned specialist rather than a non-specialist. Thus, audit quality is not statistically better, and possibly worse, when the auditor is an unseasoned specialist rather

TABLE 8 (continued)

Panel C: Income-Increasing (Positive) Discretionary Accruals

$$DA_{it} = \alpha_0 + \beta_1 SEASONED_{it} + \beta_2 UNSEASONED_{it} + \beta_3 LOGMKT_{it} + \beta_4 LEV_{it} + \beta_5 ROAL_{it} + \beta_6 ROA_{it} + \beta_7 LOSS_{it} + \beta_8 CFO_{it} + \beta_9 BTM_{it} + \beta_{10} ABS\_AC\_LAG_{it} + \beta_{11} GROWTH_{it} + \beta_{12} ALTMAN_{it} + \beta_{13} STDEARN_{it} + \beta_{14} TENURE_{it} + \beta_t YEAR\_FE_t + v_{it}$$

		Main Results				Alternative Matching Results		
Variable	Pred.	(1) Unmatched Sample Coeff. (p-value)	(2) Main Matched Sample Coeff. (p-value)	(3) Entropy Balanced Coeff. (p-value)	(4) MNCs with MVE > 500 Coeff. (p-value)	(5) 1:1 with Replacement Coeff. (p-value)	(6) 1:2 with Replacement Coeff. (p-value)	(7) 1:5 with Replacement Coeff. (p-value)
Primary Analysis								
SEASONED	−	−0.006***	−0.003**	−0.006***	−0.004***	−0.003**	−0.003**	−0.002**
		(0.000)	(0.017)	(0.000)	(0.002)	(0.041)	(0.021)	(0.024)
UNSEASONED	?	0.002	0.003	0.001	0.000	0.003	0.003	0.003
		(0.527)	(0.319)	(0.674)	(0.980)	(0.320)	(0.405)	(0.231)
Difference	−	−0.008***	−0.006**	−0.007**	−0.004	−0.006**	−0.005**	−0.006**
		(0.006)	(0.031)	(0.011)	(0.115)	(0.033)	(0.043)	(0.024)
Alternative Seasoning Periods								
SEASONED2	−	−0.007***	−0.003***	−0.007***	−0.004***	−0.003**	−0.003**	−0.003**
		(0.000)	(0.009)	(0.000)	(0.001)	(0.024)	(0.012)	(0.011)
UNSEASONED2	?	0.000	0.002	0.000	0.000	0.002	0.001	0.002
		(0.969)	(0.361)	(0.891)	(0.884)	(0.404)	(0.480)	(0.216)
Difference	−	−0.007***	−0.005***	−0.006***	−0.004**	−0.005**	−0.005**	−0.005***
		(0.002)	(0.009)	(0.004)	(0.038)	(0.013)	(0.015)	(0.005)
SEASONED3	−	−0.006***	−0.003**	−0.006***	−0.004***	−0.003**	−0.003**	−0.003**
		(0.000)	(0.013)	(0.000)	(0.004)	(0.027)	(0.016)	(0.017)
UNSEASONED3	?	−0.002	0.000	−0.002	−0.002	0.000	0.000	0.001
		(0.174)	(0.876)	(0.162)	(0.255)	(0.810)	(0.995)	(0.711)
Difference	−	−0.004**	−0.003**	−0.004**	−0.002	−0.004**	−0.003*	−0.003**
		(0.024)	(0.037)	(0.024)	(0.194)	(0.034)	(0.050)	(0.031)
Control Variables		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes
n		13,863	6,515	13,863	5,637	4,254	5,532	7,965
Adjusted R^2		0.569	0.518	0.527	0.467	0.514	0.532	0.504

(continued on next page)

than a non-specialist. Worse audit quality could occur if a newly created industry specialist does not have the wherewithal to meet the increased scale demanded by an expanded client base (Bills et al. 2016).

Alternative Definition of Specialist

Neal and Riley (2004) urge researchers to carefully consider their choice of specialization measure in studies of auditor industry expertise. Accordingly, we examine the sensitivity of our results to an alternative definition of specialization. Palmrose (1986) and Neal and Riley (2004) specify the minimum market share for specialization as 1.2 times the inverse of the number of Big N auditors. In our study, this is 30 percent (1.2/4 = 0.30). We reestimate our main results from Equations (1) and (2) using the 30 percent measure of specialization instead of our original definition (industry market share leader by at least 10 percent). The results (untabulated) are generally consistent with our hypotheses, although somewhat weaker overall. First, across both matched and unmatched samples and for all three audit quality measures (*ADA*, income-increasing *DA*, and *BTD*), the coefficient on *SPECIALIST* (Equation (1)) is generally negative. The coefficient on *SEASONED* (Equation (2)) is also negative, supporting H2. However, the coefficient on *UNSEASONED* (Equation (2)) also tends to be negative, especially at longer seasoning periods. Thus, we generally find evidence consistent with H3 (no statistical difference between the audit

TABLE 8 (continued)

Panel D: Income-Increasing (Positive) Discretionary Accruals (continued)

		Alternative Matching Results						
Variable	**Pred.**	**(8) 1:10 with Replacement Coeff. (p-value)**	**(9) 1:1 without Replacement Coeff. (p-value)**	**(10) 1:2 without Replacement Coeff. (p-value)**	**(11) 1:3 without Replacement Coeff. (p-value)**	**(12) Size-Industry-Year Coeff. (p-value)**	**(13) MVE > $500 Coeff. (p-value)**	**(14) MNCs Coeff. (p-value)**
Primary Analysis								
SEASONED	–	−0.002**	−0.003***	−0.002**	−0.003***	−0.002**	−0.006***	−0.006***
		(0.020)	(0.007)	(0.022)	(0.005)	(0.035)	(0.000)	(0.000)
UNSEASONED	?	0.003	0.002	0.003	0.004	0.004	0.001	0.002
		(0.270)	(0.496)	(0.253)	(0.220)	(0.164)	(0.695)	(0.616)
Difference	–	−0.006**	−0.005**	−0.006**	−0.007**	−0.007**	−0.007**	−0.007**
		(0.029)	(0.041)	(0.027)	(0.014)	(0.016)	(0.012)	(0.011)
Alternative Seasoning Periods								
SEASONED2	–	−0.003***	−0.004***	−0.003**	−0.004***	−0.003**	−0.007***	−0.006***
		(0.009)	(0.004)	(0.011)	(0.003)	(0.016)	(0.000)	(0.000)
UNSEASONED2	?	0.002	0.001	0.002	0.002	0.003	0.000	0.000
		(0.242)	(0.663)	(0.290)	(0.351)	(0.195)	(0.823)	(0.895)
Difference	–	−0.005***	−0.005**	−0.005***	−0.005***	−0.006***	−0.007***	−0.006***
		(0.006)	(0.016)	(0.009)	(0.006)	(0.006)	(0.001)	(0.005)
SEASONED3	–	−0.003**	−0.004***	−0.003**	−0.003***	−0.003**	−0.007***	−0.006***
		(0.015)	(0.006)	(0.021)	(0.009)	(0.028)	(0.000)	(0.000)
UNSEASONED3	?	0.000	−0.001	0.000	0.000	0.001	−0.002	−0.003
		(0.774)	(0.770)	(0.874)	(0.801)	(0.688)	(0.250)	(0.110)
Difference	–	−0.003**	−0.003**	−0.003*	−0.003*	−0.003**	−0.005***	−0.003
		(0.038)	(0.048)	(0.054)	(0.072)	(0.042)	(0.007)	(0.106)
Control Variables		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes
n		9,952	4,762	6,362	7,497	6,637	9,439	7,549
Adjusted R^2		0.504	0.513	0.530	0.529	0.533	0.497	0.532

***, **, * Indicate significance at the 1 percent, 5 percent, and 10 percent levels, respectively.
Huber-White robust standard errors are clustered by firm and are used to control for heteroscedasticity and serial correlation. When predictions are made, p-values are one-tailed. We present results for a number of different matching techniques in columns (2)–(14). Where applicable, the number after the colon in the column heading represents the number of control firms. We match with replacement (with Replacement) and without replacement (without Replacement), perform a reduced match using only size, industry, and year (1:3 with replacement), and perform exact matches with firms with market value of equity (MVE) greater than $500 million and with multinational firms (MNCs). Regressions for samples matched with replacement are estimated using WLS (Hill and Reiter 2006).
Variables are defined in Appendix A.

(continued on next page)

quality produced by unseasoned specialists compared to non-specialists) only for a one-year seasoning period. This is consistent with auditors learning over time, as discussed above. H1 (seasoned specialists provide higher audit quality than unseasoned specialists) is only supported when measuring audit quality using income-increasing *DA*. There are two explanations for these somewhat weaker results. First, unlike our main specialization measure, which ensures that an auditor is a market leader by at least 10 percent, the 30 percent rule may be too stringent for some industries and too liberal for others if auditors group just above and below the 30 percent threshold in certain industries. Second, and relatedly, seasoning may occur faster when measured by the 30 percent threshold if some auditors operate close to the specialist level just before achieving the formal specialist designation.

Further analysis sheds light on why changing the specialist definition weakens results. We find that unseasoned specialists defined using our primary specialization measure have larger mean increases in market share over the prior period compared to unseasoned specialists defined using the 30 percent measure (16.7 percent versus 14.3 percent). Auditors experiencing rapid industry market share growth are less likely to have an adequate supply of industry experts to meet staffing needs in the short term and, therefore, are more likely to function as non-specialist auditors (Bills et al. 2016). In contrast, as discussed in Section

TABLE 8 (continued)

Panel E: Book-Tax Differences

$$BTD_{it} = \alpha_0 + \beta_1 SEASONED_{it} + \beta_2 UNSEASONED_{it} + \beta_3 LOGMKT_{it} + \beta_4 LEV_{it} + \beta_5 ROAL_{it} + \beta_6 ROA_{it} + \beta_7 LOSS_{it} + \beta_8 CFO_{it} + \beta_9 BTM_{it} + \beta_{10} ABS\_AC\_LAG_{it} + \beta_{11} GROWTH_{it} + \beta_{12} ALTMAN_{it} + \beta_{13} STDEARN_{it} + \beta_{14} TENURE_{it} + \beta_t YEAR\_FE_t + v_{it}$$

		Main Results				Alternative Matching Results		
Variable	**Pred.**	**(1) Unmatched Sample Coeff. (p-value)**	**(2) Main Matched Sample Coeff. (p-value)**	**(3) Entropy Balanced Coeff. (p-value)**	**(4) MNCs with MVE > 500 Coeff. (p-value)**	**(5) 1:1 with Replacement Coeff. (p-value)**	**(6) 1:2 with Replacement Coeff. (p-value)**	**(7) 1:5 with Replacement Coeff. (p-value)**
Primary Analysis:								
SEASONED	–	−0.027*	−0.030*	−0.030*	−0.019	−0.024	−0.031**	−0.028*
		(0.077)	(0.054)	(0.059)	(0.189)	(0.121)	(0.050)	(0.067)
UNSEASONED	?	0.020	0.031	0.031	0.079**	0.029	0.030	0.024
		(0.576)	(0.400)	(0.372)	(0.041)	(0.419)	(0.413)	(0.497)
Difference	–	−0.047	−0.061*	−0.061*	−0.099***	−0.053*	−0.061*	−0.052*
		(0.106)	(0.053)	(0.050)	(0.009)	(0.070)	(0.050)	(0.085)
Alternative Seasoning Periods:								
SEASONED2	–	−0.025	−0.031*	−0.029*	−0.011	−0.025	−0.031*	−0.026*
		(0.108)	(0.066)	(0.074)	(0.326)	(0.126)	(0.063)	(0.094)
UNSEASONED2	?	−0.007	0.002	0.003	0.015	0.006	−0.001	−0.004
		(0.795)	(0.948)	(0.916)	(0.618)	(0.827)	(0.981)	(0.867)
Difference	–	−0.018	−0.033	−0.032	−0.026	−0.031	−0.030	−0.022
		(0.268)	(0.139)	(0.135)	(0.230)	(0.139)	(0.148)	(0.232)
SEASONED3	–	−0.029*	−0.038**	−0.034*	−0.018	−0.032*	−0.037**	−0.032*
		(0.089)	(0.041)	(0.056)	(0.240)	(0.080)	(0.042)	(0.066)
UNSEASONED3	?	−0.006	0.005	0.001	0.018	0.009	0.001	−0.001
		(0.804)	(0.851)	(0.976)	(0.503)	(0.716)	(0.970)	(0.977)
Difference	–	−0.023	−0.042*	−0.035	−0.036	−0.042*	−0.038*	−0.031
		(0.209)	(0.073)	(0.109)	(0.136)	(0.069)	(0.090)	(0.140)
Control Variables		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes
n		10,551	5,951	10,551	6,112	4,097	5,149	6,931
Adjusted R^2		0.069	0.080	0.067	0.043	0.085	0.081	0.072

(continued on next page)

II, auditors experiencing slower market share growth have more time to marshal needed resources for the increase in scale, and are more likely to have been operating close to the critical mass of business threshold prior to earning the specialist designation. This suggests that first-year specialists based on our primary definition of specialization are more likely to be truly unseasoned than first-year specialists based on the 30 percent measure. We believe that the results using our primary specialization measure are stronger because that measure allows a more powerful test of our hypotheses.

Alternative Audit Quality Measures

In this section, we examine our hypotheses using three additional audit quality measures. First, consistent with MM, we use measures that capture the likelihood of (1) meeting or beating analysts' forecasts, or (2) receiving a going concern opinion. Because these are binary variables, we reestimate Equations (1) and (2) as logistic regressions. We supplement the meet or beat regression with the natural logarithm of the firm's analyst following and the standard deviation of analysts' forecasts, following MM. We estimate the going concern model in distressed firms (Reichelt and Wang 2010). We include industry fixed effects in both models, following MM.

TABLE 8 (continued)

Panel F: Book-Tax Differences (continued)

		Alternative Matching Results						
Variable	**Pred.**	**(8) 1:10 with Replacement Coeff. (p-value)**	**(9) 1:1 without Replacement Coeff. (p-value)**	**(10) 1:2 without Replacement Coeff. (p-value)**	**(11) 1:3 without Replacement Coeff. (p-value)**	**(12) Size-Industry-Year Coeff. (p-value)**	**(13) MVE > $500 Coeff. (p-value)**	**(14) MNCs Coeff. (p-value)**
Primary Analysis:								
SEASONED	−	−0.030*	−0.036**	−0.030*	−0.027*	−0.039**	−0.028*	−0.031*
		(0.052)	(0.027)	(0.061)	(0.084)	(0.029)	(0.068)	(0.078)
UNSEASONED	?	0.025	0.028	0.007	0.012	0.002	0.055	0.036
		(0.485)	(0.448)	(0.847)	(0.755)	(0.965)	(0.130)	(0.374)
Difference	−	−0.054*	−0.064**	−0.037	−0.038	−0.040	−0.083**	−0.067*
		(0.075)	(0.044)	(0.166)	(0.166)	(0.156)	(0.013)	(0.062)
Alternative Seasoning Periods:								
SEASONED2	−	−0.030*	−0.040**	−0.029*	−0.028*	−0.038**	−0.025	−0.021
		(0.064)	(0.022)	(0.079)	(0.096)	(0.042)	(0.105)	(0.193)
UNSEASONED2	?	−0.001	0.007	−0.012	−0.005	−0.020	0.009	−0.018
		(0.983)	(0.805)	(0.665)	(0.855)	(0.480)	(0.744)	(0.562)
Difference	−	−0.030	−0.047*	−0.017	−0.023	−0.018	−0.034	−0.003
		(0.163)	(0.051)	(0.286)	(0.235)	(0.285)	(0.128)	(0.468)
SEASONED3	−	−0.035*	−0.045**	−0.032*	−0.032*	−0.042**	−0.036**	−0.018
		(0.051)	(0.017)	(0.075)	(0.081)	(0.035)	(0.048)	(0.236)
UNSEASONED3	?	−0.001	0.001	−0.013	−0.005	−0.017	0.015	−0.022
		(0.967)	(0.959)	(0.615)	(0.854)	(0.496)	(0.542)	(0.421)
Difference	−	−0.034	−0.046*	−0.019	−0.027	−0.025	−0.050**	0.004
		(0.124)	(0.051)	(0.260)	(0.190)	(0.205)	(0.037)	(0.454)
Control Variables		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes
n		8,266	4,572	5,970	6,799	6,024	7,882	7,515
Adjusted R^2		0.067	0.076	0.071	0.076	0.072	0.092	0.036

***, **, * Indicate significance at the 1 percent, 5 percent, and 10 percent levels, respectively.
Huber-White robust standard errors are clustered by firm and are used to control for heteroscedasticity and serial correlation. When predictions are made, p-values are one-tailed. We present results for a number of different matching techniques in columns (2)–(14). Where applicable, the number after the colon in the column heading represents the number of control firms. We match with replacement (with Replacement) and without replacement (without Replacement), perform a reduced match using only size, industry, and year (1:3 with replacement), and perform exact matches with firms with market value of equity (MVE) greater than $500 million and with multinational firms (MNCs). Regressions for samples matched with replacement are estimated using WLS (Hill and Reiter 2006). Industry fixed effects are included by two-digit SIC in the *BTD* regression.
Variables are defined in Appendix A.

When audit quality is measured as the likelihood of meeting or beating analysts' forecasts, the coefficient on *SEASONED3* (specialists with more than three years of seasoning) is generally negative (p-values $\leq$ 0.13), and there is a detectable difference in the audit quality produced by seasoned and unseasoned specialists. In the subsample consisting of specialists only, we find some evidence that seasoned specialists provide higher-quality audits than unseasoned specialists. Overall, while weaker than our main results, tests using the likelihood of meeting or beating estimates are generally consistent with our three hypotheses.[40] This is distinct from MM, who finds no evidence of a specialist effect on the likelihood of meeting or beating estimates when seasoned and unseasoned specialists are not separated.

Analyses using going concern opinions yield no evidence that specialists (regardless of tenure) have an effect on audit quality. However, going concern opinions may not be an appropriate audit quality proxy for our setting. First, going concern

[40] As discussed in Section III, the weak results could be a result of using a dichotomous, rather than continuous, audit quality measure in our nuanced setting.

TABLE 9

Estimated Coefficients and p-values from the Regressions of Audit Quality Measures on Auditor Industry Specialization and Control Variables Where Specialists are Classified as Seasoned or Unseasoned Separately for One, Two, or Three Years

$$AQ_{it} = \alpha_0 + \beta_1 SEASONED3_{it} + \beta_2 UNSEASONED1\_Dum_{it} + \beta_3 UNSEASONED2\_Dum_{it} + \beta_4 UNSEASONED3\_Dum_{it} + \beta_5 LOGMKT_{it} + \beta_6 LEV_{it} + \beta_7 ROAL_{it} + \beta_8 ROA_{it} + \beta_9 LOSS_{it} + \beta_{10} CFO_{it} + \beta_{11} BTM_{it} + \beta_{12} ABS\_AC\_LAG_{it} + \beta_{13} GROWTH_{it} + \beta_{14} ALTMAN_{it} + \beta_{15} STDEARN_{it} + \beta_{16} TENURE_{it} + \beta_t YEAR\_FE_t + v_{it}$$

Panel A: Regressions Where Specialists Are Classified as Seasoned or Unseasoned Separately for One, Two, or Three Years

		Dependent Variable: *ADA*				Dependent Variable: *Income-Increasing DA*			
Variable	**Pred.**	**(1) Unmatched Sample Coeff. (p-value)**	**(2) Main Matched Sample Coeff. (p-value)**	**(3) Entropy Balanced Coeff. (p-value)**	**(4) MNCs with MVE > 500 Coeff. (p-value)**	**(5) Unmatched Sample Coeff. (p-value)**	**(6) Main Matched Sample Coeff. (p-value)**	**(7) Entropy Balanced Coeff. (p-value)**	**(8) MNCs with MVE > 500 Coeff. (p-value)**
SEASONED3	–	−0.008***	−0.004***	−0.008***	−0.005***	−0.006***	−0.003**	−0.006***	−0.004***
		(0.000)	(0.002)	(0.000)	(0.000)	(0.000)	(0.013)	(0.000)	(0.004)
UNSEASONED1_Dum	?	0.004	0.007**	0.003	0.002	0.002	0.003	0.001	0.000
		(0.210)	(0.014)	(0.269)	(0.536)	(0.526)	(0.320)	(0.670)	(0.978)
UNSEASONED2_Dum	?	−0.003	0.000	−0.004	−0.001	−0.002	0.001	−0.002	−0.001
		(0.332)	(0.937)	(0.171)	(0.742)	(0.500)	(0.753)	(0.469)	(0.774)
UNSEASONED3_Dum	?	−0.001	0.003	−0.001	0.001	−0.008***	−0.004	−0.008***	−0.006**
		(0.654)	(0.327)	(0.649)	(0.806)	(0.002)	(0.189)	(0.006)	(0.022)
$\beta_1 - \beta_2$	–	−0.011***	−0.011***	−0.011***	−0.007**	−0.008***	−0.006**	−0.008***	−0.004
		(0.000)	(0.000)	(0.000)	(0.012)	(0.005)	(0.025)	(0.009)	(0.113)
$\beta_1 - \beta_3$	–	−0.005**	−0.004*	−0.004*	−0.004	−0.005**	−0.004*	−0.005**	−0.003
		(0.045)	(0.063)	(0.074)	(0.112)	(0.040)	(0.060)	(0.044)	(0.111)
$\beta_1 - \beta_4$	–	−0.006**	−0.007**	−0.007**	−0.006*	0.002	0.000	0.001	0.002
		(0.031)	(0.015)	(0.027)	(0.051)	(0.240)	(0.437)	(0.343)	(0.229)
Control Variables		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry Fixed Effects		No	No	No	No	No	No	No	No
n		25,901	12,153	25,901	10,379	13,863	6,515	13,863	5,637
Adjusted R^2		0.191	0.154	0.150	0.133	0.570	0.518	0.528	0.467

(continued on next page)

TABLE 9 (continued)

Panel B: Regressions Where Specialists Are Classified as Seasoned or Unseasoned Separately for One, Two, or Three Years (continued)

		Dependent Variable: *BTD*			
Variable	**Pred.**	**(9) Unmatched Sample Coeff. (p-value)**	**(10) Main Matched Sample Coeff. (p-value)**	**(11) Entropy Balanced Coeff. (p-value)**	**(12) MNCs with MVE > 500 Coeff. (p-value)**
SEASONED3	–	−0.030*	−0.038**	−0.035*	−0.019
		(0.087)	(0.041)	(0.051)	(0.233)
UNSEASONED1_Dum	?	0.020	0.032	0.032	0.079**
		(0.575)	(0.385)	(0.359)	(0.042)
UNSEASONED2_Dum	?	−0.034	−0.027	−0.027	−0.057
		(0.327)	(0.432)	(0.447)	(0.160)
UNSEASONED3_Dum	?	−0.005	0.009	−0.007	0.024
		(0.914)	(0.823)	(0.861)	(0.593)
$\beta_1 - \beta_2$	–	−0.049	−0.070**	−0.067**	−0.098**
		(0.108)	(0.043)	(0.044)	(0.015)
$\beta_1 - \beta_3$	–	0.005	−0.011	−0.009	0.038
		(0.449)	(0.382)	(0.408)	(0.197)
$\beta_1 - \beta_4$	–	−0.025	−0.047	−0.028	−0.042
		(0.572)	(0.274)	(0.261)	(0.179)
Control Variables		Yes	Yes	Yes	Yes
Year Fixed Effects		Yes	Yes	Yes	Yes
Industry Fixed Effects		Yes	Yes	Yes	Yes
n		10,551	5,951	10,551	6,112
Adjusted R^2		0.069	0.080	0.067	0.043

***, **, * Indicate significance at the 1 percent, 5 percent, and 10 percent levels, respectively.

Huber-White robust standard errors are clustered by firm and are used to control for heteroscedasticity and serial correlation. When predictions are made, p-values are one-tailed. The matched sample regression is estimated using WLS (Hill and Reiter 2006). Industry fixed effects are not included in the accruals models because the dependent variable is industry-adjusted (Minutti-Meza 2013). For book-tax differences, industry fixed effects are included by two-digit SIC.

Variables are defined in Appendix A.

opinions are rare and issued to distressed clients, resulting in low statistical power (DeFond and Zhang 2014). This is a particular concern for our study of clients of Big 4 auditors, which likely include fewer distressed firms compared to samples in prior studies that include non-Big 4 clients. Second, Butler, Leone, and Willenborg (2004) find no evidence that going concern opinions are associated with earnings management, raising questions about their validity as a proxy for audit quality. Chu, Fogel-Yaari, and Zhang (2016) also suggest that going concern reports are frequently issued in error, and audit quality models using going concern opinions yield inconsistent estimates, further undermining the validity of going concern opinions as an audit quality proxy.

Our final measure of audit quality is the ERC. ERCs are widely used in the audit quality literature and are estimable for a wide range of firms (Teoh and Wong 1993; Balsam et al. 2003; Lim and Tan 2008). They differ from our other measures of audit quality in that they are based on market perceptions and, thus, are an indirect audit quality measure.

We regress the cumulative abnormal return around annual earnings announcement (*CAR*) on unexpected earnings (*UE*), *SPECIALIST,* the interaction between unexpected earnings and the specialist variable (*SPECIALIST* ∗ *UE*), and control variables from prior auditor specialization studies (Balsam et al. 2003; Lim and Tan 2008). This model is presented in Equation (5a). If specialists produce higher-quality audits, then we expect a stronger market reaction to the earnings surprises of their clients, indicated by a positive coefficient on *SPECIALIST* ∗ *UE:*

$$\begin{aligned} CAR_{it} = {} & \alpha_0 + \beta_1 UE_{it} + \beta_2 SPECIALIST_{it} + \beta_3 SPECIALIST * UE_{it} + \beta_4 BTM_{it} + \beta_5 BTM * UE_{it} + \beta_6 VOLATILITY_{it} \\ & + \beta_7 VOLATILITY * UE_{it} + \beta_8 LEV_{it} + \beta_9 LEV * UE_{it} + \beta_{10} LOGMKT_{it} + \beta_{11} LOGMKT * UE_{it} + \beta_{12} LOSS_{it} \\ & + \beta_{13} LOSS * UE_{it} + \beta_t YEAR\_FE_t + \beta_k YEAR\_FE * UE_t + \nu_{it} \end{aligned} \tag{5a}$$

When we estimate Equation (5a), we observe a positive coefficient on *SPECIALIST* ∗ *UE* only in the unmatched and entropy balanced samples.

We then we replace *SPECIALIST* in Equation (5a) with two variables: *SEASONED* and *UNSEASONED*. Both variables are interacted with *UE*. This model is presented in Equation (5b):

$$\begin{aligned} CAR_{it} = {} & \alpha_0 + \beta_1 UE_{it} + \beta_2 SEASONED_{it} + \beta_3 SEASONED * UE_{it} + \beta_4 UNSEASONED_{it} + \beta_5 UNSEASONED * UE_{it} \\ & + \beta_6 BTM_{it} + \beta_7 BTM * UE_{it} + \beta_8 VOLATILITY_{it} + \beta_9 VOLATILITY * UE_{it} + \beta_{10} LEV_{it} + \beta_{11} LEV * UE_{it} \\ & + \beta_{12} LOGMKT_{it} + \beta_{13} LOGMKT * UE_{it} + \beta_{14} LOSS_{it} + \beta_{15} LOSS * UE_{it} + \beta_t YEAR\_FE_t + \beta_k YEAR\_FE \\ & * UE_t + \nu_{it} \end{aligned} \tag{5b}$$

If seasoned specialists outperform non-specialists (H2), then the coefficient on *SEASONED* ∗ *UE* (β_3) will be positive. If seasoned specialists outperform unseasoned specialists (H1), then the coefficient on *SEASONED* ∗ *UE* (β_3) will exceed the coefficient on *UNSEASONED* ∗ *UE* (β_5). Finally, if the quality of audits produced by unseasoned industry specialists does not differ from the quality of audits produced by non-specialist auditors (H3), then the coefficient on *UNSEASONED* ∗ *UE* (β_5) will be indistinguishable from zero. When we estimate Equation (5b), we observe a positive coefficient on *SEASONED* ∗ *UE* in the unmatched and entropy balanced samples. This provides limited support for H2. On the other hand, we find little evidence to support H1 or H3 when audit quality is gauged by the earnings response coefficient.

PSM When Using Partitioning Variables

As discussed earlier, a concern with using PSM in our setting is that we perform comparisons across multiple groups, but PSM cannot account for covariate imbalance across each group. One approach to address this issue is to perform PSM separately for each group. However, the issues inherent in PSM (e.g., sample size reduction) are especially problematic when splitting samples into smaller groups. Given the limitations of PSM, we take two approaches to address possible covariate imbalance across our partitions (untabulated). First, as discussed above in our alternative test of H1, we test H1 (seasoned specialists provide higher-quality audits than unseasoned specialists) in the subsample of specialists only. As shown in Tables 3 and 4, there are relatively few differences in covariates between seasoned and unseasoned specialists. Results of this analysis are consistent with our main results. Second, we split our sample to focus only on the two groups that we compare in each of our hypotheses, and drop the third group, creating three separate subsamples (seasoned versus unseasoned, seasoned versus non-specialist, and unseasoned versus non-specialist). We then perform entropy balancing within each subsample. Entropy balancing ensures covariate balance and avoids discarding data and random matching inherent in PSM. The results of these analyses are consistent with those presented in our main analyses. The implication is that differences in client characteristics are unlikely to drive the relation between industry specialization and audit quality.

VI. CONCLUSION

Our study provides archival evidence that auditor industry expertise is generally associated with audit quality. However, a dominant market share *by itself* does not make an industry expert. Auditors who find themselves in the dominant industry position for the first time produce a level of audit quality that is indistinguishable from, or worse than, that produced by non-specialist auditors, and lower than the audit quality produced by seasoned specialists. This pattern of results generally holds even after using a multitude of matching techniques, which calls into question prior research attributing the specialization effect to differences in client characteristics (MM). Our evidence suggests that the seasoning process takes up to three years, at which time, unseasoned specialists produce audits of a similar quality to seasoned specialists. However, we caution that our results are not robust to all audit quality measures.

To our knowledge, ours is one of the first papers to differentiate audit quality *within* specialist auditors and to estimate the speed with which knowledge is created and assimilated by an audit firm. We also provide initial evidence on what causes auditors to become specialists, finding that industry market share leaders are often created from exogenous events and subsequently begin to act as specialists over time. Thus, expertise follows market share dominance, rather than the other way around. This challenges a common assumption in the literature that clients self-select into industry specialist auditors (Gul et al. 2009).

Our approach differs somewhat from earlier city-level specialization studies because our sample is limited to clients of the Big 4, which, compared to smaller auditors, have a more centralized national focus. Thus, our conclusions relate to Big 4 national industry specialists; examining industry specialization at the city level is beyond the scope of our paper. We look to future research to provide additional evidence on these important topics using city-level specialization (e.g., Barnes 2015).

REFERENCES

Altman, E. I. 1983. *Corporate Financial Distress: A Complete Guide to Predicting, Avoiding, and Dealing with Bankruptcy*. New York, NY: Wiley.

Altman, E. I. 2013. Predicting financial distress of companies: Revisiting the Z-Score and ZETA models. In *Handbook of Research Methods and Applications in Empirical Finance*, edited by Bell, A. R., C. Brooks, and M. Prokopczuk, 428–456. Cheltenham, U.K.: Edward Elgar.

Audousset-Coulier, S., A. Jeny, and L. Jiang. 2016. The validity of auditor industry specialization measures. *Auditing: A Journal of Practice & Theory* 35 (1): 139–161. https://doi.org/10.2308/ajpt-51176

Badertscher, B. A., J. D. Phillips, M. Pincus, and S. O. Rego. 2009. Earnings management strategies and the trade-off between tax benefits and detection risk: To conform or not to conform? *The Accounting Review* 84 (1): 63–97. https://doi.org/10.2308/accr.2009.84.1.63

Balsam, S., J. Krishnan, and J. S. Yang. 2003. Auditor industry specialization and earnings quality. *Auditing: A Journal of Practice & Theory* 22 (2): 71–97. https://doi.org/10.2308/aud.2003.22.2.71

Bamber, L. S., T. E. Christensen, and K. M. Gaver. 2000. Do we really "know" what we think we know? A case study of seminal research and its subsequent overgeneralization. *Accounting, Organizations and Society* 25 (2): 103–129. https://doi.org/10.1016/S0361-3682(99)00027-6

Bao, J., and A. Edmans. 2011. Do investment banks matter for M&A returns? *Review of Financial Studies* 24 (7): 2286–2315. https://doi.org/10.1093/rfs/hhr014

Barnes, M. 2015. *The Consideration of Prior Experience in Auditor Industry Specialization*. Doctoral dissertation, The University of Oklahoma.

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. https://doi.org/10.1111/j.1911-3846.1998.tb00547.x

Bédard, J. 1989. Expertise in auditing: Myth or reality? *Accounting, Organizations and Society* 14 (1/2): 113–131. https://doi.org/10.1016/0361-3682(89)90037-8

Bell, T. B., M. Causholli, and W. R. Knechel. 2015. Audit firm tenure, non-audit services, and internal assessments of audit quality. *Journal of Accounting Research* 53 (3): 461–509. https://doi.org/10.1111/1475-679X.12078

Bills, K. L., Q. T. Swanquist, and R. L. Whited. 2016. Growing pains: Audit quality and office growth. *Contemporary Accounting Research* 33 (1): 288–313. https://doi.org/10.1111/1911-3846.12122

Blouin, J., B. M. Grein, and B. R. Rountree. 2007. An analysis of forced auditor change: The case of former Arthur Andersen clients. *The Accounting Review* 82 (3): 621–650. https://doi.org/10.2308/accr.2007.82.3.621

Butler, M., A. J. Leone, and M. Willenborg. 2004. An empirical analysis of auditor reporting and its association with abnormal accruals. *Journal of Accounting and Economics* 37 (2): 139–165. https://doi.org/10.1016/j.jacceco.2003.06.004

Cahan, S. F., D. C. Jeter, and V. Naiker. 2011. Are all industry specialist auditors the same? *Auditing: A Journal of Practice & Theory* 30 (4): 191–222. https://doi.org/10.2308/ajpt-10181

Cameran, M., J. R. Francis, A. Marra, and A. Pettinicchio. 2015. Are there adverse consequences of mandatory auditor rotation? Evidence from the Italian experience. *Auditing: A Journal of Practice & Theory* 34 (1): 1–24. https://doi.org/10.2308/ajpt-50663

Carcello, J. V., and A. L. Nagy. 2004. Audit firm tenure and fraudulent financial reporting. *Auditing: A Journal of Practice & Theory* 23 (2): 55–69. https://doi.org/10.2308/aud.2004.23.2.55

Cassell, C. A., J. C. Hansen, L. A. Myers, and T. A. Seidel. 2017. Does the timing of auditor changes affect audit quality? Evidence from the initial year of the audit engagement. *Journal of Accounting, Auditing & Finance* (forthcoming). https://doi.org/10.1177/0148558X17726241

Chi, S. S., M. Pincus, and S. H. Teoh. 2014. Mispricing of book-tax differences and the trading behavior of short sellers and insiders. *The Accounting Review* 89 (2): 511–543. https://doi.org/10.2308/accr-50644

Chu, L., H. Fogel-Yaari, and P. Zhang. 2016. *Is an Auditor's Propensity to Issue Going Concern Audit Reports a Valid Proxy for Audit Quality?* Working paper, University of Toronto.

Coca-Perraillon, M. 2006. *Matching with Propensity Scores to Reduce Bias in Observational Studies*. NorthEast SAS Users Group (NESUG). Available at: https://www.lexjansen.com/nesug/nesug06/an/da13.pdf

Danos, P., J. W. Eichenseher, and D. L. Holt. 1989. Specialized knowledge and its communication in auditing. *Contemporary Accounting Research* 6 (1): 91–109. https://doi.org/10.1111/j.1911-3846.1989.tb00746.x

Davis, L. R., B. S. Soo, and G. M. Trompeter. 2009. Auditor tenure and the ability to meet or beat earnings forecasts. *Contemporary Accounting Research* 26 (2): 517–548. https://doi.org/10.1506/car.26.2.8

DeFond, M., and J. Zhang. 2014. A review of archival auditing research. *Journal of Accounting and Economics* 58 (2/3): 275–326. https://doi.org/10.1016/j.jacceco.2014.09.002

DeFond, M., D. H. Erkens, and J. Zhang. 2017a. Do client characteristics really drive the Big N audit quality effect? New evidence from propensity score matching. *Management Science* 63 (11): 3628–3649. https://doi.org/10.1287/mnsc.2016.2528

DeFond, M., D. H. Erkens, and J. Zhang. 2017b. The Big N effect persists after matching on client characteristics: A response to Lawrence, Minutti-Meza, and Zhang 2017. *Management Science* 63 (11): 3652–3653.

deHaan, E., F. Hodge, and T. Shevlin. 2013. Does voluntary adoption of a clawback provision improve financial reporting quality? *Contemporary Accounting Research* 30 (3): 1027–1062. https://doi.org/10.1111/j.1911-3846.2012.01183.x

Dehejia, R. H., and S. Wahba. 2002. Propensity-score matching methods for nonexperimental causal studies. *Review of Economics and Statistics* 84 (1): 151–161. https://doi.org/10.1162/003465302317331982

Ericsson, K. A., N. Charness, P. J. Feltovich, and R. R. Hoffman. 2006. *The Cambridge Handbook of Expertise and Expert Performance*. Cambridge, U.K.: Cambridge University Press.

Ettredge, M. L., L. Sun, P. Lee, and A. A. Anandarajan. 2008. Is earnings fraud associated with higher deferred tax and/or book minus tax levels? *Auditing: A Journal of Practice & Theory* 27 (1): 1–33. https://doi.org/10.2308/aud.2008.27.1.1

Francis, J. R. 2011. A framework for understanding and researching audit quality. *Auditing: A Journal of Practice & Theory* 30 (2): 125–152. https://doi.org/10.2308/ajpt-50006

Geiger, M. A., and K. Raghunandan. 2002. Auditor tenure and audit reporting failures. *Auditing: A Journal of Practice & Theory* 21 (1): 67–78. https://doi.org/10.2308/aud.2002.21.1.67

Gendron, Y., D. J. Cooper, and B. Townley. 2007. The construction of auditing expertise in measuring government performance. *Accounting, Organizations and Society* 32 (1/2): 101–129. https://doi.org/10.1016/j.aos.2006.03.005

Givoly, D., C. Hayn, and B. Lourie. 2016. *Does the Tail Wag the Dog? Small-Firm-Bias in Capital Market Research*. Working paper, The Pennsylvania State University, University of California, Los Angeles, and University of California, Irvine.

Gul, F. A., S. Y. K. Fung, and B. Jaggi. 2009. Earnings quality: Some evidence on the role of auditor tenure and auditors' industry expertise. *Journal of Accounting and Economics* 47 (3): 265–287. https://doi.org/10.1016/j.jacceco.2009.03.001

Hainmueller, J. 2012. Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis* 20 (1): 25–46. https://doi.org/10.1093/pan/mpr025

Hainmueller, J., and Y. Xu. 2013. Ebalance: A Stata package for entropy balancing. *Journal of Statistical Software* 54 (7): 1–18. https://doi.org/10.18637/jss.v054.i07

Hammersley, J. S. 2006. Pattern identification and industry-specialist auditors. *The Accounting Review* 81 (2): 309–336. https://doi.org/10.2308/accr.2006.81.2.309

Hill, J., and J. P. Reiter. 2006. Interval estimation for treatment effects using propensity score matching. *Statistics in Medicine* 25 (13): 2230–2256. https://doi.org/10.1002/sim.2277

Hou, K., C. Xue, and L. Zhang. 2018. Replicating anomalies. *Review of Financial Studies* (forthcoming).

Johnson, V. E., I. K. Khurana, and J. K. Reynolds. 2002. Audit-firm tenure and the quality of financial reports. *Contemporary Accounting Research* 19 (4): 637–660. https://doi.org/10.1506/LLTH-JXQV-8CEW-8MXD

King, G., and R. Nielsen. 2016. *Why Propensity Scores Should Not Be Used for Matching*. Working paper, Harvard University and Massachusetts Institute of Technology.

King, G., R. Nielsen, C. Coberley, J. E. Pope, and A. Wells. 2011. *Comparative Effectiveness of Matching Methods for Causal Inference*. Working paper, Harvard University and Healthways, Inc.

Kothari, S. P., A. J. Leone, and C. E. Wasley. 2005. Performance matched discretionary accrual measures. *Journal of Accounting and Economics* 39 (1): 163–197. https://doi.org/10.1016/j.jacceco.2004.11.002

Krishnan, G. V. 2003. Does Big 6 auditor industry expertise constrain earnings management? *Accounting Horizons* 17 (Supplement): 1–16. https://doi.org/10.2308/acch.2003.17.s-1.1

Lim, C.-Y., and H.-T. Tan. 2008. Non-audit service fees and audit quality: The impact of auditor specialization. *Journal of Accounting Research* 46 (1): 199–246. https://doi.org/10.1111/j.1475-679X.2007.00266.x

Low, K.-Y. 2004. The effects of industry specialization on audit risk assessments and audit-planning decisions. *The Accounting Review* 79 (1): 201–219. https://doi.org/10.2308/accr.2004.79.1.201

McMullin, J. L., and B. Schonberger. 2018. *Entropy-Balanced Discretionary Accruals*. Working paper, Indiana University and University of Rochester.

Minutti-Meza, M. 2013. Does auditor industry specialization improve audit quality? *Journal of Accounting Research* 51 (4): 779–817. https://doi.org/10.1111/1475-679X.12017

Myers, J. N., L. A. Myers, and T. C. Omer. 2003. Exploring the term of the auditor-client relationship and the quality of earnings: A case for mandatory auditor rotation? *The Accounting Review* 78 (3): 779–799. https://doi.org/10.2308/accr.2003.78.3.779

Neal, T. L., and R. R. Riley, Jr. 2004. Auditor industry specialist research design. *Auditing: A Journal of Practice & Theory* 23 (2): 169–177. https://doi.org/10.2308/aud.2004.23.2.169

Owhoso, V. E., W. F. Messier, and J. G. Lynch. 2002. Error detection by industry specialized teams during sequential audit review. *Journal of Accounting Research* 40 (3): 883–900. https://doi.org/10.1111/1475-679X.00075

Palmrose, Z.-V. 1986. Audit fees and auditor size: Further evidence. *Journal of Accounting Research* 24 (1): 97–110. https://doi.org/10.2307/2490806

Phelps, C., R. Heidl, and A. Wadhwa. 2012. Knowledge, networks, and knowledge networks: A review and research agenda. *Journal of Management* 38 (4): 1115–1166. https://doi.org/10.1177/0149206311432640

Phillips, J., M. Pincus, and S. O. Rego. 2003. Earnings management: New evidence based on deferred tax expense. *The Accounting Review* 78 (2): 491–521. https://doi.org/10.2308/accr.2003.78.2.491

Power, M. 1996. Making things auditable. *Accounting, Organizations and Society* 21 (2/3): 289–315. https://doi.org/10.1016/0361-3682(95)00004-6

Reichelt, K. J., and D. Wang. 2010. National and office-specific measures of auditor industry expertise and effects on audit quality. *Journal of Accounting Research* 48 (3): 647–686. https://doi.org/10.1111/j.1475-679X.2009.00363.x

Roberts, M. R., and T. M. Whited. 2013. Endogeneity in empirical corporate finance. In *Handbook of the Economics of Finance, Volume 2A*, edited by Constantinides, G. M., M. Harris, and R. M. Stulz, 493–572. Amsterdam, The Netherlands: Elsevier.

Shipman, J. E., Q. T. Swanquist, and R. L. Whited. 2017. Propensity score matching in accounting research. *The Accounting Review* 92 (1): 213–244. https://doi.org/10.2308/accr-51449

Solomon, I., M. D. Shields, and O. R. Whittington. 1999. What do industry-specialist auditors know? *Journal of Accounting Research* 37 (1): 191–208. https://doi.org/10.2307/2491403

Stuart, E. A. 2010. Matching methods for causal inference: A review and a look forward. *Statistical Science* 25 (1): 1–21. https://doi.org/10.1214/09-STS313

Taub, S. 2005. Best Buy parts with Ernst & Young. *CFO.com* (January 5). Available at: http://ww2.cfo.com/accounting-tax/2005/01/best-buy-parts-with-ernst-young/

Taylor, M. H. 2000. The effects of industry specialization on auditors' inherent risk assessments and confidence judgments. *Contemporary Accounting Research* 17 (4): 693–712. https://doi.org/10.1506/3LDH-AV52-0F4W-H4BB

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

Tsai, W. 2001. Knowledge transfer in intraorganizational networks: Effects of network position and absorptive capacity on business unit innovation and performance. *Academy of Management Journal* 44 (5): 996–1004.

Utke, S. 2018. Miscodings in Compustat's auditor variable: Issues, identification, and correction. *Advances in Accounting* 43: 56–59. https://doi.org/10.1016/j.adiac.2018.09.002.

van Wijk, R., J. J. P. Jansen, and M. A. Lyles. 2008. Inter- and intra-organizational knowledge transfer: A meta-analytic review and assessment of its antecedents and consequences. *Journal of Management Studies* 45 (4): 830–853. https://doi.org/10.1111/j.1467-6486.2008.00771.x

APPENDIX A

Variable Definitions and Propensity Score Model

Dependent Variables

AQ_{it} represents one of the following three audit quality proxies: *ADA,* income-increasing *DA,* or *BTD*.

ADA_{it} is the absolute value of the residual (ε_{it}) from Equation (A1), based on Kothari, Leone, and Wasley (2005), for firm *i* in year *t*. We estimate the regression annually for each industry based on two-digit SIC codes, requiring at least ten observations per industry and excluding firms where the absolute value of total accruals scaled by total assets exceeds 1, following Kothari et al. (2005).

$$AC_{it} = \alpha + \beta_0(1/AT_{it-1}) + \beta_1 \Delta REV_{it} + \beta_2 PPE_{it} + \beta_3 ROA_{it-1} + \varepsilon_{it} \quad \text{(A1)}$$

where:

AC_{it} is total accruals for firm i in year t, defined as net income from continuing operations minus operating cash flow scaled by total assets at the end of year t–1;

AT_{it-1} is total assets for firm i at the end of year t–1;

ΔREV_{it} is the change in revenue for firm i at the end of year t scaled by total assets at the end of year t–1;

PPE_{it} is net property, plant, and equipment for firm i at the end of year t scaled by total assets at the end of year t–1; and

ROA_{it-1} is net income for firm i in year t–1 scaled by total assets at the start of year t–1.

DA_{it} is the signed value of the residual from Equation (A1). When using DA as our audit quality proxy in our regressions, we limit our sample to firms with income-increasing DA ($DA > 0$).

BTD_{it} is the grossed-up deferred tax expense scaled by pre-tax income (Chi et al. 2014). Deferred tax expense equals deferred federal tax plus deferred foreign tax. If missing, then deferred tax equals Compustat deferred tax. If still missing, then deferred tax equals total tax expense minus current tax expense. To gross-up deferred tax expense, we divide deferred tax expense by the U.S. tax rate, and multiply by 1 minus the rate. We exclude foreign firms, utilities, financials, and loss firms.

CAR_{it} is the size-adjusted three-day abnormal return around the annual earnings announcement (Lim and Tan 2008).

Variables of Interest

$SPECIALIST_{it}$ takes on the value of 1 when an auditor has a market share that is the highest in a given industry *and also* more than 10 percent higher than the next-largest competitor during the year, and is 0 otherwise. Each auditor's market share is defined as the sum of sales in an industry-year for each auditor, divided by the sum of sales across all auditors in the industry-year. We measure specialization at the U.S. national level, where U.S. auditors are determined by the client's headquarters location. We correct Compustat's auditor variable as described in Utke (2018).

$SEASONED_{it}$ takes on the value of 1 when the auditor is classified as a specialist *and UNSEASONED* is coded 0; otherwise, *SEASONED* is 0.

$UNSEASONED_{it}$ takes on the value of 1 when the auditor is in its first year of being classified as a specialist in a given industry, and is 0 otherwise.

$SEASONED2_{it}$ takes on the value of 1 when the auditor is classified as a specialist *and UNSEASONED2* is coded 0; otherwise, *SEASONED2* is 0.

$UNSEASONED2_{it}$ takes on the value of 1 when the auditor is in its first or second year of being classified as a specialist in a given industry, and is 0 otherwise.

$SEASONED3_{it}$ takes on the value of 1 when the auditor is classified as a specialist *and UNSEASONED3* is coded 0; otherwise, *SEASONED3* is 0.

$UNSEASONED3_{it}$ takes on the value of 1 when the auditor is in its first, second, or third year of being classified as a specialist in a given industry, and is 0 otherwise.

$UNSEASONED1\_Dum_{it}$ takes on the value of 1 when the auditor is in its first year of being classified as a specialist in a given industry, and is 0 otherwise.

$UNSEASONED2\_Dum_{it}$ takes on the value of 1 when the auditor is in its second year of being classified as a specialist in a given industry, and is 0 otherwise.

$UNSEASONED3\_Dum_{it}$ takes on the value of 1 when the auditor is in its third year of being classified as a specialist in a given industry, and is 0 otherwise.

Control Variables

$LOGMKT_{it}$ is the natural logarithm of the market value of equity for firm i in year t.

LEV_{it} is total debt for firm i in year t divided by average total assets in year t.

ROA_{it} is net income for firm i in year t divided by average total assets in year t.

$ROAL_{it}$ is net income for firm i in year t–1 divided by average total assets in year t–1.

$LOSS_{it}$ takes on the value of 1 if net income is negative for firm i in year t, and is 0 otherwise.

CFO_{it} is operating cash flow for firm i in year t divided by average total assets in year t.

BTM_{it} is the book value of equity divided by the market value of equity for firm i in year t.
$ABS\_AC\_LAG_{it}$ is the absolute value of total accruals for firm i in year t–1 divided by average total assets in year t–1.
$GROWTH_{it}$ is sales growth from the prior year for firm i in year t.
$ALTMAN_{it}$ is the Altman (1983) financial distress score, as clarified by Altman (2013), for firm i in year t.
$STDEARN_{it}$ is the standard deviation of income before extraordinary items for firm i over the prior four years (t–1 to t–4).
$TENURE_{it}$ takes on the value of 1 if the client has had the same auditor for five or more years, and is 0 otherwise. Our *TENURE* variable is defined slightly differently than in MM. MM sets his tenure variable to 1 if a client has the same auditor for more than two years, and reports that auditor tenure exceeds two years in 99.3 percent of his sample observations. Our definition allows for variation in the variable and is more consistent with prior literature (Davis, Soo, and Trompeter 2009).
$YEAR\_FE_t$ are year fixed effects.
UE_{it} is the actual year-end earnings minus the most recent median earnings forecast, scaled by stock price two days before the earnings announcement (Balsam et al. 2003).
$VOLATILITY_{it}$ is the standard deviation of daily stock returns over the 90-day window ending seven days prior to the earnings announcement (Lim and Tan 2008).

Propensity Score Matching Model

Observations are matched by propensity score, within common support, using a caliper distance of 0.03. In alternative specifications, we vary the number of treatment to control firms (one-to-one, one-to-two, one-to-three, one-to-five, and one-to-ten) and match with and without replacement. Following MM, the propensity of choosing a specialist auditor is predicted using a logistic regression of the auditor's specialist status on variables related to the client's level of earnings quality and year and industry fixed effects (Equation (A2)) as follows:

$$\begin{aligned} SPECIALIST_{it} = {} & \alpha_0 + \beta_1 LOGAT_{it} + \beta_2 LEV_{it} + \beta_3 ROAL_{it} + \beta_4 ROA_{it} + \beta_5 LOSS_{it} + \beta_6 CFO_{it} + \beta_7 BTM_{it} \\ & + \beta_8 ABS\_AC\_LAG_{it} + \beta_9 GROWTH_{it} + \beta_{10} ALTMAN_{it} + \beta_{11} STDEARN_{it} + \beta_{12} TENURE_{it} + \beta_t YEAR\_FE_t \\ & + \beta_r IND\_FE_i + v_{it} \end{aligned} \tag{A2}$$

$LOGAT_{it}$ is the natural logarithm of firm i's total assets in year t. Consistent with MM, we use the natural logarithm of total assets in the matching model instead of $LOGMKT_{it}$ because $LOGAT_{it}$ generally results in better matches.
IND_FE_i are two-digit SIC industry fixed effects.