Journal of Accounting Research

DOI: 10.1111/1475-679X.12105
Journal of Accounting Research
Vol. 54 No. 3 June 2016
Printed in U.S.A.

CHICAGO BOOTH

Auditor–Client Compatibility and Audit Firm Selection

STEPHEN V. BROWN* AND W. ROBERT KNECHEL†

Received 12 March 2013; accepted 4 January 2016

ABSTRACT

We examine auditor switching conditional on the compatibility of clients and their auditors using a unique text-based measure of similarity of financial disclosures. We find clustering of clients within an audit firm based on this measure. We find that clients with the lowest similarity scores are significantly more likely (9.4%–10.6%) to switch auditors, and will change to an audit firm to which they are more similar. Regarding the effect on audit quality, we find that discretionary accruals are lower when similarity is higher. However, accounting restatements are more likely when text disclosures that are unaudited—business description, and management discussion and analysis (MD&A)—are more similar. We find no such similarity effect for the audited footnotes. Finally, we find that firms that are more similar are less likely to receive a going concern opinion (GCO), but the GCO reporting decision is more accurate. It is unclear if this reflects higher or lower audit quality since firms that are candidates for a GCO are intrinsically different from the average firm in an auditor's portfolio due to their financial distress. One

*Unaffiliated; †University of Florida.

Accepted by Phil Berger. We are grateful for feedback from Stephen Asare, Ryan Cerf, Praveen Pathak, and Jennifer Wu Tucker. We would also like to thank workshop participants at the University of Texas–Austin, University of Toronto, Arizona State University, Florida International University, Southern Methodist University, the University of Florida, the University of Illinois at Urbana–Champaign, the University of Kentucky, the University of South Carolina, and the University of Virginia for their feedback on portions of this paper.

implication of these results is that auditors might have greater involvement in the quality of the text disclosures that are currently not audited.

JEL code: M42

Keywords: audit markets; auditor switching

1. Introduction

The process by which a client selects an auditor can be complex and can be influenced by a number of factors. The factors that might affect the degree of compatibility between an auditor and a client include pricing, expertise, location, interpersonal associations, and the extent of agency problems in the client (Johnson and Lys [1990], Chaney, Jeter, and Shaw [1997], Knechel, Niemi, and Sundgren [2008]). Some of these attributes are obviously more relevant than others for determining the overall quality of the resulting audit. A limited amount of research has examined alignment between clients and certain *types* of auditors based on factors such as the size of the audit firm (Shu [2000], Landsman, Nelson, and Rountree [2009]) or degree of industry specialization (Knechel, Naiker, and Pacheco [2007]). However, there is less research on the compatibility of *specific* auditors and *specific* clients. What literature exists has focused on operational and financing changes in a client (Johnson and Lys [1990]), the effects of a client hiring a former audit partner (e.g., Lennox and Park [2007]), or the effect of auditors and clients connected via the university education of management (Guan et al. [2016]) or the audit committee (He et al. [2014]). Further, there is evidence that the supply of audits shifts over time, both at the national level (DeFond and Lennox [2011]) and at the local level (DeKeyser et al. [2015]). A recent paper by Gerakos and Syverson [2015] analyzes the consumer surplus implicit in current auditor–client combinations by estimating the implicit cost of lost surplus that would occur with the introduction of mandatory audit firm rotation. The sizable loss of surplus suggests that the current alignment of clients and auditors is beneficial to clients, on average.[1]

In general, clients may have preferences about various aspects of the audit process, its outcomes, and the nature of the relationship with their auditor. While audit firm alumni, common educational experiences, and

[1] In related research, Francis, Pinnuck, and Watanabe [2014] report that clients of a single audit firm in the same industry have earnings that are more comparable to each other than to clients of other audit firms. Francis, Pinnuck, and Watanabe [2014] examine comparability of accruals across pairwise sets of clients of a single auditor in the same industry. They do not extend their analysis to include auditor switches, nor do they discuss the implications of the observed comparability for overall audit quality. We discuss the study by Francis, Pinnuck, and Watanabe [2014] in more detail later in this paper. However, if there is a gain in audit quality by breaking client–auditor links that are based on social ties (He et al. [2014]), there could be a net benefit to such rotation.

social ties may proxy for some of these characteristics, the alignment of auditors and clients is likely to be much more complex. In this paper, we define *auditor–client compatibility* as the ability of the auditor to satisfy a client's preferences, given the auditor's own preferences, abilities, and constraints. If client preferences vary from company to company, and the ability of an auditor to meet a client's needs varies from firm to firm, the degree of fit between any two entities will also vary.[2] A client's preferences for a certain audit firm can be represented conceptually as a vector of attributes that are desirable to the client (e.g., location, cost, expertise, social links, etc.), which can be mapped to a vector of attributes supplied by any given auditor.[3] Compatibility would then reflect the degree of commonality across the two vectors. Since many of these attributes are likely to be unobservable (other than very obvious attributes such as specialization), the actual degree to which a client and a specific auditor are compatible across the entire spectrum of potential attributes is also unobservable.

One way that compatibility may manifest is in the numbers included in the financial statements as examined in Francis, Pinnuck, and Watanabe [2014]. However, the narrative disclosures of the financial statements are also critical to evaluating a company's performance, and are much harder to analyze. Further, the auditor has different levels of responsibility—audited versus reviewed—for different elements of the narrative disclosures, potentially revealing either more (or less) influence of the auditor on some disclosures. In this paper, we examine the narrative disclosures included in the text-based parts of the financial statements that provide information about a company, its operations, and its accounting choices. Based on this perspective, we develop a unique measure of auditor–client compatibility for Big 4 firms based on the similarity of their financial *disclosures* (rather than their financial *results*). We then use the similarity score to evaluate a firm's decision to switch auditors and the impact of increasing compatibility (similarity) on audit quality.

The basic idea is that, on average, clients "cluster" in a specific audit firm for a reason, even though we may be unable to observe the exact

[2] In this paper, we use the term "fit" interchangeably with "compatibility." Johnson and Lys [1990] argue that a lack of fit between an auditor and a client reflects a loss of efficiency in the cost of conducting an audit. While that is certainly one possible explanation for an auditor–client mismatch, our approach does not require the restricting assumption that a change in auditor is due solely to cost considerations. More specifically, we assume that changes in service, expertise, and quality considerations that are reflected in the accounting disclosures can also be associated with an auditor–client misalignment.

[3] For purposes of this discussion, we do not specify whose desirable attributes drive the auditor choice decision, that is, management, the audit committee, others, or some combination of all three. A manager may want an auditor that is nondisruptive to the personnel of the company or "cooperative" on accounting issues, while an audit committee may be focused on reputation and ability to meet deadlines. The priorities assigned to these attributes depend on the internal interactions of the stakeholders of the firm. See Fiolleau et al. [2013] for a discussion of the process used, and the issues that arise, during an audit client tender and subsequent change in auditors.

attributes that cause a company to select a specific auditor. Ceteris paribus, it seems reasonable to expect that like-clients are drawn to like-auditors. This means that the extent to which companies have similar audit preferences will cause them to choose a similar auditor, subject to potential constraints on choice. In this study, we compare the similarity of an individual client to all of the current clients within an industry of a specific auditor to generate a proxy for how well that company fits into each auditor's client base. When a company appears similar to other clients of the same auditor, we presume that the auditor is likely to have developed expertise and cost advantages related to that "type" of client. Therefore, we consider a company and auditor to be compatible when the company is similar to other clients of the audit firm in the same industry, and we consider low compatibility to occur when there is low similarity between the company and the audit firm's existing clients. Further, we expect that the clients with the lowest compatibility are more likely to switch to auditors with which they are more compatible.

Our measure of intercompany similarity within an industry is based on the narrative disclosures in a client's annual report and is derived from the similarity score introduced in Brown and Tucker [2011]. In this paper, we consider three narrative disclosures separately and together: (1) the company's business description, (2) the accounting footnotes, and (3) management discussion and analysis (MD&A). This metric is based on the assumption that unobservable factors in a company's operations and accounting choices will manifest in their annual report, from which we can then derive a proxy for similarity. Using the textual disclosures in the financial statements, we calculate the similarity of each client to other clients of an audit firm in the same industry and year (the "reference group") using the Vector Space Model from the document retrieval literature (e.g., internet search algorithms). We then use this similarity score to predict auditor–client alignment and switches. The fact that the auditor audits only the footnotes but reviews the business description and MD&A may lead to different levels of similarity across the various disclosures.[4]

In general, we find that clients clustered within an industry at the audit firm level tend to have higher similarity scores when compared to clients of other auditors in the same industry and time period. This result suggests that our proxy for similarity is capturing information about the compatibility between an individual client and an individual audit firm. For example, we find that up to 58% of clients employ one of the two best-fitting auditors

[4] The PCAOB is increasingly concerned with the quality of audit procedures applied to the footnotes. In a recent Staff Audit Practice Alert (#12, 9/9/2014), the PCAOB specifically mentions a concern that auditors do not perform appropriate or adequate procedures "to identify and assess the risks of material misstatement of the financial statements, *including consideration of the risk of omitted, incomplete, or inaccurate disclosures*" [emphasis added]. This highlights the importance of footnotes in evaluating a company's financial report.

among the Big 4.[5] We also observe that, on average, similarity increases as a function of auditor tenure. However, because auditor–client relationships are "sticky," it is more effective to examine how similarity relates to the specific event of an auditor–client switch. We observe that the poorer the fit with an existing auditor, the greater the probability the client will choose to switch to a new auditor. Further, we find that the successor auditor is generally the nonincumbent firm that has the best relative fit. This pattern of results also occurs when we control for local market conditions.

Extending the analysis, we also examine the association between fit and audit quality. We find that discretionary accruals are lower when auditor–client compatibility is better, suggesting higher audit quality. However, we find a higher incidence of accounting restatements when the similarity of the unaudited MD&A and client business description is high, but not when the similarity of the audited footnotes is high. We also find that financially distressed firms that are more similar are less likely to receive a going concern opinion (GCO), but similarity is also associated with increased accuracy in GCO reporting. It is unclear if these results reflect a relative gain or loss of audit quality or the possibility that financially distressed firms are systematically different from the average client of an audit firm due to their financial distress, resulting in low similarity scores. In these cases, audit clients may be *appropriately dissimilar* from the average client because their disclosures, by necessity, include discussion of issues related to poor financial performance that do not apply to other—financially healthy—clients. On balance, the audit quality tests are mixed but suggest that reviewed—in contrast to audited—disclosures may be associated with quality problems as they become more similar across clients.

This paper makes a number of important contributions to the literature. First, we introduce a text-based metric to measure auditor–client compatibility (similarity) based on a reference group of other companies clustered by audit firm and industry. In contrast to Francis, Pinnuck, and Watanabe [2014], we measure compatibility using the rich text disclosures of the 10-K rather than the reported financial results (i.e., accruals). We demonstrate that there are significant differences across audit firms that cause a "clustering" of audit clients based on this similarity. Second, we add to the limited literature on auditor–client fit by considering the suitability of a specific auditor for a specific client. For example, Johnson and Lys [1990] use a small number of operating, investing, and financing attributes to predict voluntary auditor changes. Shu [2000] uses a number of client and auditor attributes such as size and profitability to model client–auditor mismatches between the choice of a Big N or non–Big N auditor. Guan et al. [2016] examine university ties between auditors and CEOs, while He et al. [2014]

[5] This is statistically higher than 50% (the expected rate in random assignment of clients to auditors among the Big 4). Further, as we will show later, the probability of hiring a specific auditor is strictly increasing in similarity.

examine school and accounting department ties between auditors and audit committees.[6] The narrow focus of these studies reveals a number of interesting phenomena, but our approach provides a more comprehensive measure of similarity based on the full financial statement disclosures that can then be used to predict auditor–client mismatches for *any* Big 4 auditor–client relationship.

Third, we show that similarity in stable auditor–client relationships may be associated with some loss of audit quality when similarity arises in *unaudited* (i.e., reviewed) portions of the narrative disclosures, but this is less likely in the audited footnotes. We speculate that high similarity in non-audited disclosures may reduce transparency of those disclosures. Taken together, these results may be of interest to policy makers for two important reasons: (1) regulatory discussions on mandatory audit firm rotation could have implications for the cost and quality of auditing if a client is forced to switch from a compatible auditor to one that is less compatible, and (2) proposals to expand the auditor's reporting responsibilities might mitigate the loss of audit quality when similarity arises in unaudited disclosures.

The rest of this paper proceeds as follows. The next section develops the hypotheses and discusses prior literature. Following that section is the rationale and foundation for the similarity measures, a demonstration of how they are calculated, and a discussion of observed trends. A description of the design and results of the empirical tests follows, while the next section examines these results for their sensitivity to alternative similarity measures. The final section contains the conclusion.

2. *Hypotheses and Prior Literature*

Predicting the choice of an auditor based on auditor–client compatibility requires two conditions: (1) variation in client preferences regarding an auditor, and (2) variation across auditors in their ability to satisfy those preferences. If auditors are all essentially equivalent (no auditor variation), clients could randomly choose amongst them, possibly based only on price.[7] On the other hand, if auditors vary but clients all have the same preferences, then all the clients would prefer the same auditor, subject to capacity constraints. The audit literature has documented substantial evidence of variation in client preferences and auditor capability. For example, a large, multinational client is more likely to choose a Big N auditor (Chaney, Jeter, and Shivakumar [2004]), at least in part because a smaller auditor does not have the resources and capability of auditing such a company (Carson 2009). Further, as audit firms enter or leave specific audit

[6] Landsman, Nelson, and Rountree [2009] used an approach similar to Shu [2000] to observe that auditor switching became more sensitive to auditor–client misalignment after Enron.

[7] This would reflect the audit as a commodity, something that many professionals feel has become more prevalent (Christensen et al. [2014]).

markets, the changing supply may influence auditor–client alignments (DeFond and Lennox [2011]).[8] Therefore, the U.S. audit market has enough variation to suggest that auditor–client compatibility is an issue worth examining (Numan and Willekens [2012]).

In a related vein, a number of studies have also focused on industry specialization as a differentiator of both demand and supply. Industry specialists are those auditors that have invested significantly in expertise in auditing a particular industry, typically reflected by larger market shares in that industry. While market share measures have been questioned as an appropriate proxy for specialization (Minutti-Meza [2013]), industry specialists are generally considered to deliver higher quality (Bell, Causholli, and Knechel [2015]), are better at detecting errors (Owhoso et al. [2002]), are associated with clients having higher earnings response coefficients (Balsam, Krishnan, and Yang [2003], Gul, Fung, and Jaggi [2009]), and report lower discretionary accruals (Krishnan [2003]). They are also better at improving audit quality through knowledge spillovers from non-audit services (Lim and Tan [2008]). Beyond variation in quality, there may also be cost differences among specialists. While most studies have found specialists charge higher audit fees (Gramling and Stone [2001]), there is also the possibility of cost savings through the leveraging of expertise (Craswell, Francis, and Taylor [1995], Willenborg [2002], Cahan et al. [2008]). Further, audits by industry specialists tend to be more efficient (Cairney and Young [2006]). As a result, client preferences for certain levels of cost and quality will lead them to choose an auditor with structural characteristics that best meet their needs.[9]

A handful of studies have examined auditor–client compatibility at a more granular level. For instance, Lennox and Park [2007] find that a company is more likely to engage a particular auditor when a former employee of that auditor is on the management team. Guan et al. [2016] show that audit quality is negatively affected when an auditor and CEO are connected through a mutual university education. In a similar vein, He et al. [2014] show that educational ties between the auditor and audit committee are associated with lower audit quality, especially if the tie is via the same area of study (e.g., accounting department) at a university. They also find that the school ties are associated with higher audit fees, suggesting that the compatibility of some clients and auditors can be attributable to factors other than cost.

[8] In this study, we only consider Big 4 auditors so changes in the supply side of audits are unlikely to have much impact on our results other than in some very small local audit markets where only two or three of the Big 4 may operate. This is discussed in more detail later in this paper.

[9] Investor preference can also play a role in the decision. Switches to larger auditors and specialist auditors are associated with positive market reactions (Fried and Schiff [1981], Nichols and Smith [1983], Knechel, Naiker, and Pacheco [2007]). Today's optimal choice in auditor could change in the future in a manner similar to how a firm's optimal capital structure changes over time.

Research on client–auditor disagreements finds that a client is more likely to switch from auditors that prefer conservative accounting, presumably in an effort to find an auditor who is more amenable to the company's preferences (Krishnan [1994], DeFond and Subramanyam [1998]). Bamber and Iyer [2007] find that auditors that strongly identify with the client will be more likely to allow a desired accounting treatment. While this "opinion shopping" may appear disreputable, Dye [1991] shows that the firm may simply be trying to better communicate its internal information to the market. The broad conclusion is that interpersonal relationships and opinion shopping are just two of many possible reasons (or attributes) related to why a client can find one auditor to be more compatible than others. In sum, given sufficient variation among clients and auditors, each is likely to choose a counterparty that best matches a broad range of preferences and needs.[10]

2.1 AUDITOR–CLIENT COMPATIBILITY AND AUDITOR SWITCHING

If clients prefer high auditor compatibility, then it follows that low compatibility with their current auditor will likely trigger an eventual change in auditor. Johnson and Lys [1990] demonstrate that companies are more likely to switch auditors as the client's operating, investing, and financing activities change over time. They interpret the increased likelihood of switching as an efficient response to temporal changes in the company's audit preferences. In effect, the auditor–client compatibility that was utility-maximizing in the past has shifted such that another auditor may now be a better fit. However, given the "stickiness" of auditor–client relationships and the potentially high cost of changing auditors, the shift in circumstances will have to be large enough to overcome the inertia present in an existing auditor–client relationship. Shu [2000] examines auditor–client fit based on whether the client has a Big N auditor when an empirical model would predict a non–Big N auditor, or vice versa, finding clients are more likely to change auditors when there is a mismatch. We extend the concept of auditor–client compatibility by examining whether a comprehensive measure of accounting similarity that reflects mismatches with a specific auditor is likely to lead to client–auditor realignment. Based on the degree of fit between the auditor and client, we expect that poor compatibility is more likely to be associated with an auditor switch, leading to our first hypothesis:

> *H1:* Clients having relatively poor compatibility (similarity) with their current auditor are more likely to change auditors than those with high compatibility (similarity).

[10] The decision is also subject to various constraints. We do not address constraints on choice in this paper, but the effect will be to shift a client away from its apparent best fit, thus working against our findings. Further, we are unable to incorporate the effect of specific partners who might service a client due to lack of partner data in the United States.

Since we cannot directly measure compatibility,[11] we use as a proxy the similarity between a given client and the cluster of other clients in an industry that are audited by the audit firm. That is, compatibility is inherently a 1:1 relationship between a single company and a single auditor, but the various attributes of that relationship are not observable. What is observable is the relationship between the audit firm and its clients as reflected in pooled attributes of the client base. If we assume that the 1:M relationship between an audit firm (1) and its client base (M) reflects, on average, a certain level of compatibility, the M:1 relationship (similarity) between the client base (M) and a single client (1) can proxy for compatibility in the auditor–client relationship. A low degree of similarity between a client and the cluster of other clients in the firm may cause the client to look for a new auditor.[12]

Once a client has made the decision to change auditors, it will need to choose a new one from the set of available auditors. We assume that most public companies currently using a Big 4 auditor will limit themselves to choosing a new Big 4 auditor given those auditors' dominant market share and the potential for the market to penalize a firm that downgrades from a Big 4 to a non–Big 4 auditor (Knechel, Naiker, and Pacheco [2007]).[13] That means a company changing auditors must select from the three remaining nonincumbent firms (subject to possible constraints).[14] If compatibility is low with the present auditor who is to be replaced, it then follows that a company will generally prefer a new auditor that has better compatibility from among the remaining options. This perspective leads to our second hypothesis:

H2: A client switching auditors will choose a new auditor that has a relatively higher degree of compatibility (similarity) from among the nonincumbent auditors.

While H1 and H2 are written in terms of the overall (national) market for auditors, it is not clear whether compatibility should be considered at the national or local level. There are reasons to consider both. An audit firm's reputation, brand value, and audit methodology and training are

[11] Francis, Pinnuck, and Watanabe [2014] examine the covariance of accruals between individual companies and auditor fixed effects ("style") to test financial statement comparability across clients within an audit firm.

[12] The decision to actually change auditors is complex and would take into account considerations beyond that of fit or similarity. For example, the transaction costs of an auditor change can be quite high so the degree of incompatibility would have to be relatively large to justify an auditor change on a cost/benefit basis.

[13] Excluding non–Big 4 auditors from our analysis is a potential limitation but is consistent with other research that examines audit markets (Casterella et al. [2004], Numan and Willekens [2012]).

[14] Some of the constraints that might influence the choice of a new auditor include firms having offices in the local area or the fact that the potential alternative auditors are already engaged with the client as suppliers of non-audit services.

generally established at the national level. Also, many of the policies of a firm that impact the structure and content of financial statements are set at the national level, for example, implementation of new accounting standards and disclosure checklists that could influence the textual disclosures that are the focus of our analysis. All of the Big 4 firms have professional practice groups and internal quality review processes (Bell, Causholli, and Knechel [2015]) that exist at the national level to ensure consistency across audits, independent of the client's location. Therefore, the characteristics of the national level practice are likely to be relevant for assessing compatibility. These same forces could induce higher levels of similarity over time as a new auditor becomes more experienced with a client and better understands the client's financial and reporting structure.[15] However, we also know from extensive research that the local office that conducts an audit can have a direct impact on the quality of audits (Francis and Yu [2009]). Further, client personnel interact primarily with local office personnel, and any problems or complaints are likely to emanate from local office relationships, especially if there is turnover in the local client service team. Finally, a client seeking to hire an auditor will look primarily at the options that are locally available. Taken together, these factors lead to our third hypothesis:

H3: Client–auditor switching will be influenced by similarity between a client and auditor at the local level.

2.2 EFFECT OF AUDITOR–CLIENT SIMILARITY ON AUDIT QUALITY

While it is relatively intuitive to grasp how auditor–client compatibility can influence the decision to change or retain an auditor, it is less clear how changes in compatibility would influence audit quality. In experimental studies, Hammersley [2006] shows that "matched" specialists—defined as those operating within their industry of expertise—are more likely to correctly process experimental cues regarding misstatements than mismatched specialists, while Low [2004] shows that an industry-based mismatch negatively affects audit planning and risk assessments. Using archival data, Bell, Causholli, and Knechel [2015] show that partners with expertise in a client's industry perform better as measured by internal quality assessments. These studies imply that audit quality will be higher when there is

[15] While we do not hypothesize increasing similarity as a function of tenure, our results suggest that similarity increases for stable client–auditor relationships. Some of the conditions that might cause increasing similarity include firm-specific disclosure checklists, reporting templates, internal quality review, and individual partner styles. Bell, Causholli, and Knechel [2015] show that first-year audits have significantly lower quality than other audits and that quality improves for many years as the auditor tenure increases. Also, research in audit production has shown that audits are often not efficient, suggesting a less-than-perfect fit between an auditor and a client (Dopuch et al. [2003], Knechel, Rouse, and Schelleman [2009]). This also suggests that there is a learning curve for both the auditor and client over time (Causholli [2015]). Increased similarity may be a manifestation of such a learning curve.

a better fit between the auditor and client. Johnstone, Li, and Luo [2014] find that clients within the same supply chain—a measure of relatedness, if not similarity—tend to have higher audit quality because they can share information across audits. Further, an extensive body of empirical research on auditor industry specialization, another potential dimension of compatibility, suggests that audit quality is higher in engagements conducted by an auditor that specializes in the industry of a client (Balsam, Krishnan, and Yang [2003], Kwon, Lim, and Tan [2007], Gul, Fung, and Jaggi [2009], Reichelt and Wang [2010]).

On the other hand, arguments are often made—typically in the context of either extended auditor tenure or auditor-provided non-audit services (NAS)—that audit quality can be undermined by economic and social bonding between an auditor and a client. Compatibility may actually reflect a form of bonding that could undermine audit quality. While the evidence related to auditor tenure and NAS is decidedly mixed, some research finds decreased audit quality when the auditor closely identifies with the client. For example, Cahan and Zhang [2006] find a negative association between long tenure and audit quality, while Frankel, Johnson, and Nelson [2002] find a negative association between NAS and audit quality.[16] In essence, the possibility exists that auditor–client compatibility could increase to the point that the two parties become so close that an auditor's independence (objectivity) is compromised. For example, Lennox [2005] shows that companies are more likely to receive a clean audit opinion if the client hires personnel from the audit firm into accounting positions. Menon and Williams [2004] find that clients employing former audit partners in executive positions tend to report higher levels of discretionary accruals. Guan et al. [2016] report that an auditor who is connected to a client's CEO or Chairman through university education is less likely to issue a modified audit opinion. He et al. [2014] find that school ties between an auditor and audit committee are also associated with lower audit quality. These papers illustrate the potential negative effect of bonding between an auditor and a client and suggest that increased compatibility could be associated with decreased audit quality.

The dynamics of audit quality are also influenced by a change of auditors. DeFond and Subramanyam [1998] observe that an auditor change is associated with income-decreasing accruals in the last year of an incumbent's tenure but are mostly insignificant in the first year of the successor's tenure. Krishnan [1994] reports that the threshold for receiving qualified audit opinions is lower for clients that switch audit firms than those that do not switch, that is, the auditor is more conservative when accepting a new client. Geiger and Raghunandan [2002] report that audit reporting failures

[16] In both the cases of auditor tenure and NAS, there is also a great deal of research that either fails to find an association with audit quality or finds a positive association. For example, for auditor tenure, see Myers, Myers, and Omer [2003], and for NAS, see Ashbaugh, LaFond, and Mayhew [2003].

related to GCOs are highest in the early years of an engagement. These results suggest that increased auditor–client similarity at the time of a switch, or over time, may not necessarily lead to reduced audit quality. Given the competing nature of the various lines of argument, and the potentially offsetting effects that would follow, our fourth hypothesis is nondirectional:

H4: Audit quality is associated with auditor–client compatibility.

To test the effect of similarity on audit quality, we consider measures of discretionary accruals, the occurrence of accounting restatements, the propensity of the auditor to issue GCOs, and the accuracy of GCO reporting.

3. Measurement of Auditor–Client Similarity (Compatibility)

When clients are similar to each other in terms of industry, operations, risks, or transactions, auditors have the opportunity to specialize in those companies for both reputational (quality) and audit production (cost) reasons. Prior literature typically uses an all-or-nothing industry membership test to organize clients into similar groups.[17] While the accounting systems and underlying economics of multiple organizations are not separately observable, their joint effect is presumably reflected in the financial statements (De Franco, Kothari, and Verdi [2011], Francis, Pinnuck, and Watanabe [2014]) and related narrative disclosures. Consequently, we develop a continuous-measure proxy for the degree of compatibility between a company and an auditor's existing clientele based on text retrieval using narrative-based documents as the underlying data. The narrative disclosures are especially flexible, giving management the opportunity to communicate firm-specific information or to influence the market's view of the company. Our proxies indicate how similar a specific client is to all the other clients audited by a specific accounting firm in a given year and industry (i.e., each similarity score is specific to an auditor–industry–year nexus).

3.1 THE SAMPLE AND COMPUTATION OF SIMILARITY SCORES

A similarity metric is calculated for a single company by comparing it to some well-defined group of other companies called the *reference group*. In this paper, we define the reference group to be all of the other clients of a specific Big 4 accounting firm in the same industry and same year having assets greater than $1 million and no change in the fiscal year end.[18] We use industries to define our reference groups because of potential systematic differences in the structure of financial statements and narrative disclosures across different industries. The sample begins in 1997, when EDGAR

[17] Gramling and Stone [2001] summarize the industry specialization literature, which generally finds differences in both quality and audit fees for industry specialists versus nonspecialists.

[18] We further consider the effect of client size on similarity in supplemental analysis.

data were first widely available, and ends in 2009. Former Arthur Andersen clients are included, but only after they have not engaged Andersen for at least one year to limit the potential confounding effect of the collapse of the firm.[19] Financial data were collected from Compustat for each observation with at least one similarity score available. The variables used in our testing are summarized in table 1. Auditor tenure is calculated based on the auditor information in Compustat beginning in 1974. We exclude utilities and financial services industries due to inherent differences in the structure of their financial statements and related disclosures (Blay and Geiger [2013]). We also exclude any reference groups that do not have at least five observations because the similarity score is unreliable in small groups.[20] Finally, a company is omitted from the reference group in the year that it switches auditors.

The information retrieval literature has developed numerous methods for measuring the similarity of two documents, often in the context of matching a user's Internet search query to the closest applicable Web pages (Singhal [2001]). A common method for analyzing text documents is the Vector Space Model (VSM). The VSM maps a document into a numerical vector based on the unique words used in the document (Salton, Wong, and Yang [1975]). The frequency of specific words in a document defines a ray that starts at the origin in n-dimensional space where "n" is the number of unique words in the document. The most common approach to calculate the similarity of two documents is to take the cosine of the angle between any two text-based vectors, that is, rays from the origin (Singhal [2001]). This approach extends the method of Brown and Tucker [2011] from a simple pairwise comparison of companies (based only on MD&A disclosures) to an aggregate measure of multiple comparisons within the same auditor-industry-year reference group based on the significant text disclosures in a 10-K.

The VSM-based similarity score for narrative disclosures is calculated based on three different types of nonvoluntary text disclosures in the 10-K. It is preferable to use more than one disclosure item because of considerable variation in terms of subject matter, time-horizon, and audit requirements. Specifically, we examine:

- *Company business description.* Item 101 of Regulation S-K requires a detailed description of the business, including industrial and geographic

---

[19] Note that we also repeated the auditor switching and audit quality tests after completely excluding former Andersen clients from the sample and found that all of our results were qualitatively unchanged.

[20] Exclusion of non–Big 4 firms is consistent with prior literature (Casterella et al. [2004], Numan and Willekens [2012]). Non–Big 4 firms rarely have reference groups large enough for meaningful analysis. We consider non–Big 4 firms in more detail in our supplemental analysis for clients that upgrade from the non–Big 4 to the Big 4. Adding the non–Big 4 firms would increase the overall sample size by only 6.9% but would result in reference groups that are quite small.

TABLE 1
Descriptive Statistics

Panel A: Client descriptive statistics

Variable	Mean	Std Dev	25%	Median	75%	*N*
SIM_{BUS}	0.000	0.078	−0.055	−0.019	0.038	33,293
$SIM_{MD\&A}$	0.000	0.088	−0.059	−0.025	0.036	31,213
SIM_{NOTES}	0.000	0.048	−0.029	−0.013	0.012	22,026
SIZE	5.804	1.919	4.466	5.730	7.063	35,593
IRISK	0.256	0.195	0.096	0.222	0.370	35,055
TACC	−0.056	0.222	−0.092	−0.046	−0.004	33,955
$\|DACC\|$	0.082	0.164	0.020	0.045	0.093	33,946
CASH	0.224	0.251	0.028	0.115	0.352	35,589
ROA	−0.080	0.348	−0.083	0.022	0.068	35,550
LOSS	0.397					35,593
GROWTH	0.273	1.178	−0.063	0.052	0.215	35,284
ACQUIS	0.127					35,593
CFEARLY	0.452					35,593
CFMATURE	0.357					35,400
TENURE	9.570	8.206	4.000	7.000	13.000	35,593
MODOPIN	0.402					35,593
EXPERT	0.153					35,593
OFFICESIZE	43.488	68.592	11.000	26.000	44.000	29,910
MKTSIZE	3.204	1.389	3.000	4.000	4.000	33,704
SWITCH	0.064					32,597

Panel B: Similarity by client size

Variable	(Smallest) Q1	Q2	Q3	(Largest) Q4	*t*-stat	Prob > *t*
SIM_{BUS}	−0.026	−0.011	0.008	0.025	38.65	0.000***
$SIM_{MD\&A}$	−0.014	−0.002	0.006	0.008	14.91	0.000***
SIM_{NOTES}	−0.010	−0.006	0.002	0.012	20.70	0.000***
SIM_{COMB}	8.305	8.936	9.470	9.950	21.43	0.000***

Panel A: *SIM* = similarity to other clients in the auditor–industry–year reference group (subscripts: *BUS* = business description, *MD&A* = management's discussion & analysis, *NOTES* = footnotes to financial statements). *SIZE* = log of total assets. *IRISK* = receivables plus inventory, scaled by assets. *TACC* = total accruals. $|DACC|$ = absolute discretionary accruals from the cross-sectional modified Jones model. *CASH* = cash and equivalents, scaled by assets. *ROA* = income before extraordinary items, scaled by assets. *LOSS* = 1 if $ROA < 0$. *GROWTH* = change in assets, scaled by prior year assets. *ACQUIS* = 1 if the acquisition activity in the current year exceeds 10% of assets. *CFEARLY* = 1 if cash flows indicate the client is in the introduction or growth stage of life cycle (Dickinson [2011]). *CFMATURE* = 1 if the client is in the mature stage. *TENURE* = number of years with the current auditor. *MODOPIN* = 1 for a nonstandard audit opinion. *EXPERT* = 1 if the auditor has at least 5% more clients than the next largest auditor in both the industry and the MSA. *OFFICESIZE* = number of clients in the auditor's local office. *MKTSIZE* = number of Big 4 auditors having five or more clients in the same MSA as the observation. *SWITCH* = 1 if an auditor change in the following year.

Panel B: Presents the average similarity scores for each quartile of client size. *t*-stat is for a test of whether the largest clients have a different similarity score from the smallest clients.

***, **, * indicate the significance at 1%, 5%, and 10%, respectively.

segments, principal products and services, R&D spending, and competitive conditions. This is the disclosure that is likely to have the least formal structure but may be relatively unchanged over time, unless the nature of the company's business changes significantly (e.g., a merger). This item is *reviewed* by the auditor and serves as a baseline for the other two metrics.[21]

- *Management discussion and analysis* (*MD&A*). Item 303(a) requires that the MD&A contain a discussion of liquidity, capital resources, results of operations, off-balance sheet arrangements, and contractual obligations. MD&A is intended to be an interpretation of past and future operations "through the eyes of management" (SEC [2003]). This item is considered to have significant information content for users but is only *reviewed* by the auditor.
- *Financial footnotes*. The footnote content is the only disclosure specifically determined by GAAP. This is also the only disclosure that is formally *audited* by the auditor (AU Sections 550; 551).[22]

These are the three longest disclosures in a typical 10-K. Excluding exhibits, there are an average of 6,334 words in the business description, 7,044 in the MD&A, and 8,598 in the footnotes (see table B1), comprising 17%, 18%, and 21%, respectively, of the length of the typical 10-K.

The data for the narrative disclosure score are taken from 10-Ks filed electronically via the SEC's EDGAR system for fiscal years 1997 through 2009 for the clients of Big 4 audit firms having at least five other observations available for comparison within the same auditor–industry–year reference group. Appendix A describes the selection and extraction process, which yields 33,293 observations for the business description, 31,213 for MD&A, and 22,026 for the footnotes. Treating the three narrative disclosure items of the annual report as separate data sets, the similarity score is calculated using the process summarized in appendix B. This yields three variables—SIM_{BUS}, $SIM_{MD\&A}$, and SIM_{NOTES}—that proxy for the degree of similarity between one client and all other clients in the same auditor–industry–year. We also create a composite measure described below, SIM_{COMB}. Our analysis is based on the assumption that higher similarity scores correspond to greater auditor–client compatibility arising from a self-selected clustering of like-clients and like-auditors.

3.2 DESCRIPTIVE RESULTS

Table 1, panel A, contains descriptive statistics for the disclosure similarity measures. The sample size is largest for SIM_{BUS} (33,293), followed

[21] The fact that the business descriptions (and MD&A) are only reviewed by the auditor does not mean that the auditor has no influence over its content. While our results vary, we show later in this paper that there is, in fact, some clustering of clients within audit firms based on these disclosures. However, our research method does not facilitate a direct test of clustering for the audited footnotes relative to the business description and MD&A.

[22] The footnotes and MD&A are particularly important to stakeholders, given the large number of accounting standards requiring or encouraging specific footnote disclosures and the relatively frequent guidance by the SEC on MD&A (SEC [1987, 1989, 2003]). Prior studies have demonstrated the usefulness of footnotes (Shevlin [1991], Amir [1993], Wahlen [1994], Riedl and Srinivasan [2010]). Other research has shown some of the potential information contained in MD&A (Feldman et al. [2010], Li [2010], Sun [2010]).

by SIM_{MDA} (31,213). There are fewer observations for SIM_{NOTES} (22,026) because some companies include the financial statements and related footnotes in locations other than the main body of the 10-K (see appendix A). The average auditor–industry–year reference group size is about 22 for the business description, 22 for MD&A, and 18 for the footnotes. SIM_{BUS}, $SIM_{MD\&A}$, and SIM_{NOTES} are approximately centered around zero. Similarity scores are significantly higher for companies in the top quartile of size than for those in the bottom quartile, indicating that larger clients tend to be at the "core" of the auditor's portfolio in terms of their similarity (table 1, panel B). The three similarity scores are all positively correlated with one another, indicating they measure related constructs, with correlations ranging from 0.59 to 0.69.

To provide some background on the audit market, table 2, panel A, contains the average portfolio size of each of the Big 4 auditors and all non–Big 4 auditors combined ("Other"). While the non–Big 4 firms audit a substantial *number* of clients, the Big 4 audit the vast majority of client assets of public firms and receive the preponderance of the available audit fees. Deloitte and KPMG are smaller than E&Y and PwC along most dimensions. Panel B shows that a number of client features vary among the Big 4, including client size, various proxies for audit risk and complexity (*IRISK*, *TACC*, *CASH*, *ROA*, *LOSS*, and *GROWTH*), and length of the auditor–client relationship (*TENURE*). These results suggest that differences among the four firms exist that could be associated with the compatibility (similarity) clustering we test in this paper. Finally, panel C presents the number of auditor changes per year. For the Big 4, an average of 6.7% of clients switch auditors each year, although there is substantial annual variation in the frequency of switches, from a low of 4.5% in 2009 to a peak of 9.3% in 2005. When the clients of non–Big 4 auditors are included, the switching frequency increases slightly to 8.3%.

The similarity scores are not directly comparable to one another because of variations in how they are calculated (e.g., a score of 0.20 for MD&A is not necessarily higher or even different from a score of 0.15 for footnotes). To compare across measures, we standardize the scores to have a mean of zero and standard deviation of 1 in the year prior to a change. Figure 1 plots the similarity scores against auditor tenure. Tenure of 1 indicates the first year of an auditor's tenure (i.e., the year after an auditor switch). For visual comparability, all scores are adjusted to begin at zero in panel A. Panel B contains the same graph before making this adjustment and shows that the average similarity one year prior to a change in auditor is negative for all scores (i.e., similarity with the predecessor auditor was negative just prior to a switch on average).[23] All similarity measures increase in year 1 indicating that an auditor switch has an immediate impact on auditor–client compatibility. Similarity then generally increases over the length

[23] The graphs only present tenure up to year 10, after which the decreasing number of observations leads to heightened volatility in the scores.

TABLE 2

Auditor Descriptive Statistics

Panel A: Auditor portfolio size (annual average)

	Number of Clients		Audit Fees		Client Assets Audited	
	Frequency	%	$ (millions)	%	$ (billions)	%
Deloitte	1,062	15	1,245	19	3,153	18
E&Y	1,489	21	1,632	25	4,002	23
KPMG	1,150	16	1,307	20	3,978	23
PwC	1,442	20	1,963	30	5,680	33
Other	1,936	27	392	6	440	3
Total	7,079		6,539		17,253	

Panel B: Client descriptive statistics by auditor

Variable	Deloitte	E&Y	KPMG	PwC	*F*	Prob > *F*
SIZE	5.860	5.790	5.630	5.901	21.67	0.000***
IRISK	0.276	0.246	0.253	0.255	35.71	0.000***
TACC	–0.046	–0.055	–0.064	–0.059	7.63	0.000***
|*DACC*|	0.079	0.083	0.086	0.081	2.44	0.062*
CASH	0.182	0.250	0.218	0.228	111.98	0.000***
ROA	–0.051	–0.102	–0.085	–0.072	35.36	0.000***
LOSS	0.360	0.425	0.402	0.389	28.15	0.000***
GROWTH	0.210	0.273	0.272	0.319	11.76	0.000***
ACQUIS	0.129	0.122	0.135	0.127	2.11	0.096*
TENURE	9.125	9.393	8.991	10.499	63.06	0.000***
MODOPIN	0.413	0.396	0.407	0.398	2.27	0.078*

Panel C: Auditor switches per year

	Clients of Big 4 Auditors			Clients of All Auditors		
Year	Audits	Switches	%	Audits	Switches	%
1997	5,582	269	4.8	6,899	436	6.3
1998	5,325	337	6.3	6,563	606	9.2
1999	5,467	368	6.7	6,863	639	9.3
2000	5,186	399	7.7	6,722	714	10.6
2001	4,919	359	7.3	6,397	628	9.8
2002	4,700	290	6.2	6,116	462	7.6
2003	5,323	366	6.9	6,883	545	7.9
2004	5,099	445	8.7	6,816	624	9.2
2005	4,786	445	9.3	6,720	653	9.7
2006	4,484	341	7.6	6,649	558	8.4
2007	4,254	239	5.6	6,491	475	7.3
2008	4,064	213	5.2	6,216	412	6.6
2009	3,988	180	4.5	6,060	371	6.1
Total	63,177	4,251		85,395	7,123	

Panel A: Frequency and dollar amounts are annual averages for each of the Big 4 auditors and non–Big N auditors ("Other"). Audit fees are available from fiscal year 2000 onward.

Panel B: Average value of each client descriptive per auditor. *F*-statistic is from a one-way ANOVA testing whether that descriptive is equal across all auditors. *SIZE* = log of total assets. *IRISK* = receivables plus inventory, scaled by assets. *TACC* = total accruals. |*DACC*| = absolute discretionary accruals from the cross-sectional modified Jones model. *CASH* = cash and equivalents, scaled by assets. *ROA* = income before extraordinary items, scaled by assets. *LOSS* = 1 if *ROA* < 0. *GROWTH* = change in assets, scaled by prior year assets. *ACQUIS* = 1 if the acquisition activity in the current year exceeds 10% of assets. *TENURE* = number of years with the current auditor. *MODOPIN* = 1 for a nonstandard audit opinion. ***, **, * indicate the significance at 1%, 5%, and 10%, respectively.

Panel C: Number of auditor changes per year for just Big 4 auditors and for Big 4/non–Big N auditors. "Audits" is the number of audits performed that year, "Switches" is the number of auditor changes, and "%" is the number of switches divided by the total number of audits.

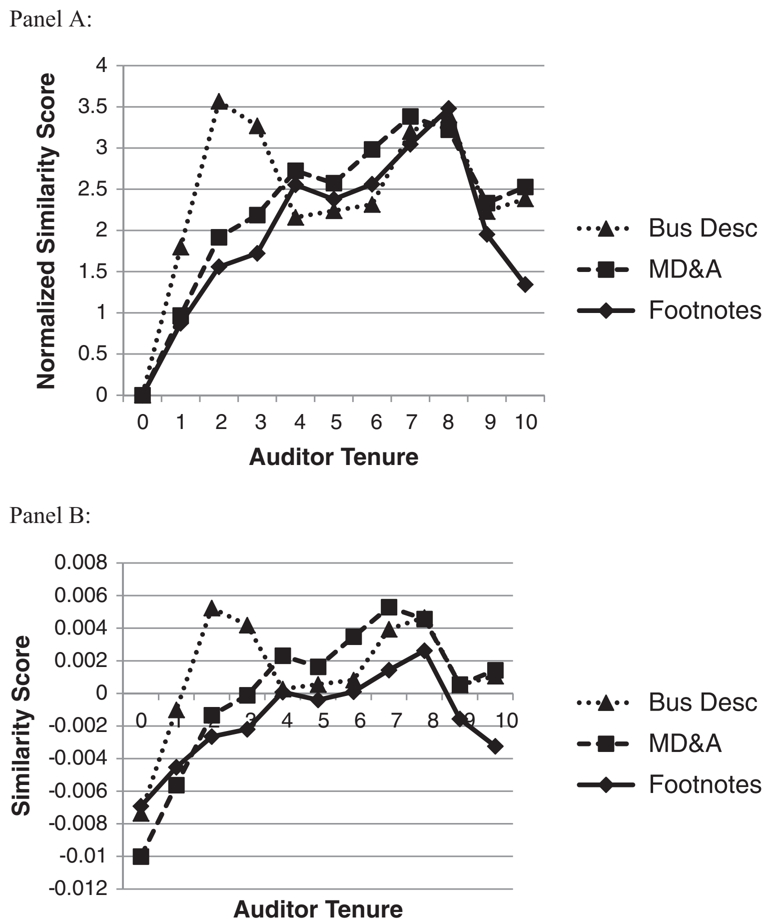

FIG. 1.—Trends in similarity scores over auditor tenure. These graphs contain plots of similarity over the first 10 years of the auditor–client engagement. Panel A plots a normalized similarity score, used only in this graph, which is standardized to have a mean of zero and standard deviation of one, and then adjusted to begin at zero in the year before an auditor switch. Without the normalization, the different similarity scores are not on the same scale. For reference, panel B contains the unadjusted scores used in all statistical tests. The auditor switch occurs when tenure equals one. Auditor tenure of zero indicates the similarity to the new auditor in the year before the switch, thus serving as a baseline before that new auditor has had any effect on client disclosures.

of the auditor–client relationship, indicating that auditor–client compatibility improves over time (on average).[24] Auditors might find this trend beneficial if it results in higher audit quality or reduces the effort involved

[24] We do not hypothesize this increase in our study. An argument could be made that comparability would be the highest at the time of an auditor switch when the fit between a client and an auditor was greatest. Such an argument would be analogous to a company choosing

in the audit engagement. Likewise, clients might benefit from adopting best practices arising from the auditor's expertise developed in similar client engagements.

SIM_{BUS} experiences a rapid increase in similarity during the first two years of the engagement, then declines slightly but generally becomes more stable. This result is not a surprise given the relative lack of year-on-year changes in the business model of most companies absent a major strategic shift in the company. $SIM_{MD\&A}$ also increases quickly in the first two years and then rises more slowly until approximately year 7. SIM_{NOTES} increases gradually over time, reaching a peak in year 8. To test for the statistical significance of these trends, we compare the similarity scores for short-tenure (one to three years) and long-tenure (8 to 10 years) engagements. In all cases, the similarity scores are significantly higher for longer-tenure clients than for shorter ones. We perform a related test using year-on-year changes in similarity and find that the annual changes in SIM_{BUS} and $SIM_{MD\&A}$ for long-tenure clients are not as large as the changes seen in newer clients, implying levels of these measures are increasing at a decreasing rate.[25] Changes in SIM_{NOTES} scores are not significantly different between short- and long-tenure engagements. In general, these patterns demonstrate that auditor–client compatibility is not static and is a function of auditor tenure.

Even though similarity scores increase over time on average, we expect certain subgroups of clients to display a different pattern. In figure 2, we compare similarity score trends for clients that switch a single time during our sample period to those that switch twice. We limit the tenure period on the graph to five years to have sufficient observations for our analysis (i.e., the second switch must occur in 2005 or before given our sample period ends in 2009). Clients that changed auditors only once during the sample presumably had a more successful switch than those that changed more than once. Figure 2 presents diverging trends for these two groups. Single-switchers have an increase in compatibility with their new auditor over the first five years of the engagement, while double-switchers experience a much smaller increase initially that leads to decreases after the first few years.[26] Therefore, while similarity appears strongly related to tenure following an auditor change, not all clients may experience this increase in compatibility, and those firms are more likely to switch a second time.

---

an optimal capital structure that becomes less optimal over time. If that was the case, then we would expect to see similarity scores decline in general after the change in auditors. Given that we do not observe this, we believe that both the auditor and client experience a learning curve effect through their repeated interactions because the assessment of "fit" at the time of an auditor switch is, by necessity, based on incomplete or imperfect information given the wide range of attributes that would affect the decision (Causholli [2015]).

[25] The t-statistics for the difference in means between long- and short-tenure clients of the business description, MD&A, and footnotes are 1.90, 3.00, and 2.74, respectively. The corresponding t-statistics for the difference in annual changes are 4.32, 2.85, and 0.69.

[26] Some volatility is evident in the footnote trend, most likely due to a relatively small number of observations (an average of 69 double-switchers per year).

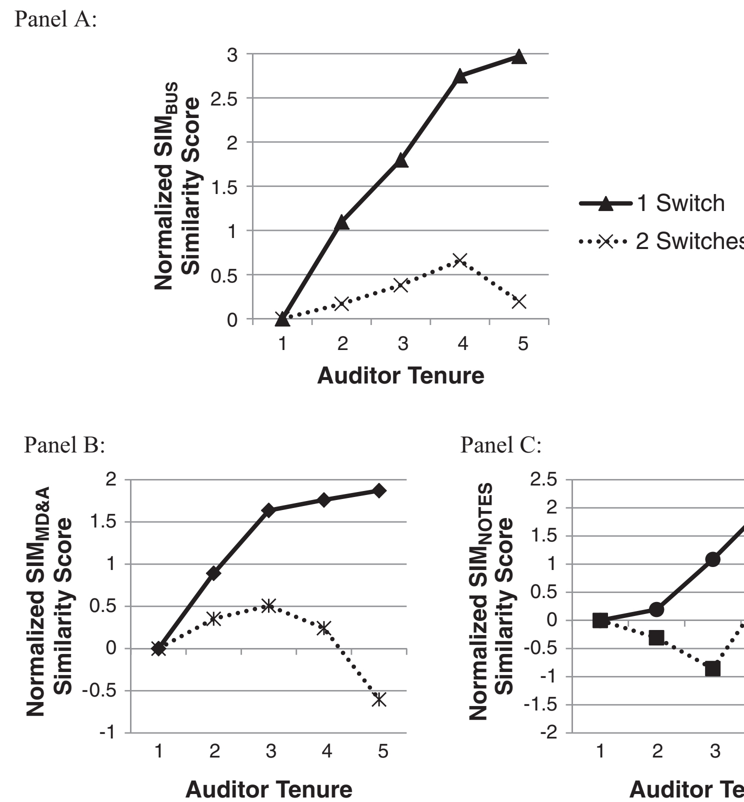

FIG. 2.—Similarity score trends for single and double auditor switches. This graph presents trends in similarity scores to a newly selected auditor for clients switching auditors during our sample. The solid lines represent clients that remain with this new auditor for at least five years while the dashed lines are for clients switching a second time. Switches must occur before 2005 to have enough data for five years before the end of our sample period.

Finally, we examine the stability of the similarity scores since we expect that auditor–client compatibility will not change dramatically over short time periods. In untabulated analysis, we find that the autocorrelation coefficients for SIM_{BUS} (0.93), $SIM_{MD\&A}$ (0.92), and SIM_{NOTES} (0.91) are quite high, demonstrating a high degree of time-series stability in all measures.

If clients and auditors randomly choose to enter into an audit engagement without regard to auditor–client alignment, one would expect a relatively equal distribution of clients across firms when sorting on the compatibility score (i.e., approximately a 25% market share to each audit firm after omitting clients of non–Big 4 firms). Table 3, panel A, summarizes auditor–client alignment using each of the similarity scores. For all three measures, the probability of using a specific auditor monotonically increases as auditor–client alignment increases. For SIM_{BUS}, 27% of clients use the auditor with which they are best aligned, while only 23% use the

TABLE 3
Auditor Selection Based on Auditor–Client Compatibility

Panel A: Rank of incumbent auditor based on similarity to auditor's client base

	Bus Desc			MD&A			Footnotes			Combined		
	Freq	%	Cml	Freq	%	Cml	Freq	%	Cml	Freq	%	Cml
1 (similar)	7,979	27	27	8,304	30	30	5,726	31	31	4,684	28	28
2	7,675	26	53	7,038	25	55	4,875	27	58	4,129	25	53
3	7,341	25	77	6,712	24	79	4,136	23	81	4,212	25	79
4 (different)	6,720	23	100	5,761	21	100	3,567	19	100	3,515	21	100
Total observations	29,715			27,815			18,304			16,540		

Panel B: Average rank of incumbent auditor

	Avg Rank	*t*-stat
Bus desc	2.43***	10.73
MD&A	2.36***	21.41
Footnotes	2.30***	24.08
Combined	2.40***	12.00

Panel C: Rank of new auditor following a Big 4-to-Big 4 auditor change

	Bus Desc			MD&A			Footnotes			Combined		
	Freq	%	Cml	Freq	%	Cml	Freq	%	Cml	Freq	%	Cml
1 (similar)	288	37	37	317	42	42	188	41	41	148	35	35
2.5	269	34	71	239	31	73	139	30	71	143	34	69
4 (different)	232	29	100	205	27	100	135	29	100	130	31	100
Total observations	789			761			462			421		

Panel D: Average rank of new auditor following a Big 4-to-Big 4 auditor change

	Avg Rank	*t*-stat
Bus desc	2.39***	2.69
MD&A	2.28***	5.03
Footnotes	2.30***	3.50
Combined	2.44***	0.96

Panel A: Freq = number of times an auditor of a given rank is engaged by a client. Rank 1 corresponds to the most compatible auditor, while rank 4 indicates the most incompatible auditor. % = percentage of client-years engaging that rank. Cml = cumulative frequency. Only markets with four auditors are included; including markets with only two or three auditors yields qualitatively identical results.

Panel B: Avg Rank = average rank of the auditor engaged by a client. Random choice (null hypothesis) is 2.5.

Panel C: Similar to panel A, but only for clients changing from one Big 4 auditor to another Big 4 auditor in the year of the change. In contrast to panel A, the possible ranks are 1, 2.5, and 4 because only three auditors are available.

Panel D: Similar to panel B, but only for clients changing auditors.

***, **, * indicate the significance at 1%, 5%, and 10%, respectively.

auditor with which they are least aligned.[27] The MD&A pattern is stronger, with 30% of clients using the most aligned auditor and just 21% using the

[27] While the variance from 25% is small, we do not expect to see extremely large differences for stable client–auditor combinations because of the cost of auditor switching and the duration of auditor tenure in the sample. Many factors can influence auditor–client alignment when the relationship is stable, that is, fit as measured by our compatibility score is just one dimension to alignment. As a result, we feel that the best test of auditor–client compatibility relates to auditor switches (to be discussed) since that is the point in time when the client's alignment with the old and new auditors would be most salient.

least aligned auditor. The clearest pattern occurs based on SIM_{NOTES}, where 31% of clients use the most similar auditor and only 19% use the least similar auditor. Since the auditor is likely to have the most influence on a client's audited footnotes, this result supports our contention that the auditor can influence a client's narrative disclosures. In combination, these results suggest that our proxy for similarity is reflective of underlying differences in the clientele of firms and compatibility across clients and firms.

To evaluate the statistical significance of these patterns, we use the similarity scores to assign a rank to each auditor available in a client's geographical area, based on the Metropolitan Statistical Area (MSA) of the corporate headquarters. The rank for each auditor ranges from 1 to 4, where 1 corresponds to the most compatible auditor and 4 to the least compatible. Not all Big 4 auditors are active in a given location, so we require that at least two auditors be available with at least five clients each in an MSA before calculating the ranks. If all four auditors are available (87.0% of the audit market), ranks of 1, 2, 3, and 4 are assigned. For MSAs with only three auditors (7.7% of the market), ranks of 1, 2.5 (the range midpoint), and 4 are used. Finally, if two auditors are available (5.3% of the market), only ranks 1 and 4 are assigned.[28]

The average auditor–client alignment rank pooled across all markets is then compared to the expected rank (i.e., 2.5) under the null hypothesis of a random distribution. Note that the test focuses on whether a client is associated with a *more similar* auditor rather than the *most* similar auditor. Table 3, panel B, shows the average rank of the incumbent auditor based on auditor–client alignment. The average ranks are 2.43 ($t = 10.73$) for SIM_{BUS}, 2.36 ($t = 21.41$) for $SIM_{MD\&A}$, and 2.30 ($t = 24.08$) for SIM_{NOTES}, all of which are significantly less than 2.5. These results suggest that our similarity metric is consistent with client clustering on compatibility with an auditor. Consequently, we feel we can use the similarity scores to test for compatibility effects surrounding auditor switches.

4. *Hypothesis Tests*

4.1 HYPOTHESIS 1: SIMILARITY AND THE LIKELIHOOD OF AUDITOR CHANGE

In a stable client–auditor relationship, we see that clients are more likely to be aligned with auditors whose portfolio consists of similar clients (on average). However, such alignment at any point in time may reflect firm–client history and inertia as much as it does a desire for "fit." This section presents several analyses surrounding the decision to switch auditors. First, we develop the following model of auditor change based on variables from the existing literature on auditor switching, augmented with our similarity measures (firm and year subscripts are suppressed):

[28] Scaling so that the interval between ranks is consistent without changing the mean rank (e.g., 1.5, 2.5, and 3.5 in an MSA with three firms) does not change our results.

$$\begin{aligned} SWITCH = {} & a_0 + a_1 SIM + a_2 SIZE + a_3 IRISK + a_4 \left|DACC\right| + a_5 CASH \\ & + a_6 ROA + a_7 LOSS + a_8 GROWTH + a_9 ACQUIS \\ & + a_{10} CFEARLY + a_{11} CFMATURE + a_{12} \ln\left(TENURE\right) \\ & + a_{13} MODOPIN + a_{14} EXPERT + industry + year + \varepsilon. \qquad (1) \end{aligned}$$

The dependent variable, *SWITCH*, is an indicator set to 1 if the client changes auditors in the *subsequent* year; all other variables are measured in the current year. *SIM* is the placeholder for the various similarity scores (SIM_{BUS}, $SIM_{MD\&A}$, SIM_{NOTES}, and SIM_{COMB}) that we test in this paper.

The model includes a variety of control variables. Because larger firms tend to change auditors less frequently, the natural log of total assets (*SIZE*) in the year before the switch is included in the model. The inherent audit risk (*IRISK*) is defined as receivables plus inventories, scaled by total assets.[29] We include the client's absolute discretionary accruals (*DACC*) as defined by DeFond and Subramanyam [1998].[30] Higher cash and equivalents scaled by total assets (*CASH*) proxy for the financial risk of a client. We further control for financial risk with return on assets (*ROA*) and a dummy variable when a company incurs a loss (*LOSS*). Both of these variables are expected to be positively associated with auditor switches. *GROWTH* in assets is associated with greater litigation risk for the auditor, which we expect to be associated with more auditor changes. Because M&A activity can lead to an increased likelihood of changing auditors when the previously separate entities engaged different auditors, a dummy variable is included to indicate that a company has engaged in acquisition activity that exceeds 10% of total assets (*ACQUIS*).[31] Using firm life cycle proxies from Dickinson [2011], we include dummies for firms in the start-up or growth phase of their life cycle (*CFEARLY*) and for mature firms (*CFMATURE*).[32]

We also control for various features of the auditor and auditor–client relationship. The natural log of the number of years the client has engaged its current auditor (*TENURE*) is expected to be negatively associated with auditor changes due to stability in long-term audit–client relationships. *MODOPIN* is a dummy set to 1 for anything other than a standard unqualified opinion. We follow Landsman, Nelson, and Rountree [2009] and include a

[29] All continuous control variables, other than *SIZE* and *TENURE*, are winsorized at the 1st and 99th percentiles.

[30] Absolute discretionary accruals also proxy for the "unusualness" of the client relative to the industry, which is related to our similarity construct. Excluding this variable does not change our results.

[31] We omit the leverage variable in Landsman, Nelson, and Rountree [2009] since it is not significant in either their study or this one.

[32] In an alternative test, we drop the two life cycle variables and include company age instead. The test of probability of auditor change gives qualitatively similar results, except the *p*-value for the footnotes coefficient is now 0.013 instead of the prior 0.010. Our later tests on audit quality all yield qualitatively identical conclusions using age.

dummy set to 1 for an expert auditor (*EXPERT*), assessed as having at least 5% more clients in the same industry and metropolitan location as the next largest firm.[33] Finally, the model contains industry and year fixed effects.[34] The correlations presented in table 4, panel A, indicate that larger (*SIZE*), more profitable (*ROA*), and less risky companies (*IRISK*) are more likely to have a better fit with their auditor.

Model (1) is augmented with each of the proxies for auditor–client fit. The results are presented in table 4, panel B. The controls are generally consistent with prior literature. Two similarity measures are negatively related to auditor switches, supporting H1, which states that clients having poorer auditor fit are more likely to change auditors: $SIM_{MD\&A}$ ($z = 3.07$) and SIM_{NOTES} ($z = 2.59$). SIM_{BUS}, the metric that is likely to be the most generic and stable across firms, is not significant. To perform a test of "composite" similarity, the three similarity scores are converted to quintile rankings and then the quintile rankings are summed across the three similarity scores. SIM_{COMB} ranges from 3 (minimum similarity) to 15 (maximum similarity). When using this variable in the auditor switch model, the coefficient is also negative ($z = 2.59$). Holding all the other variables at their means, an interquartile decrease in MD&A similarity is associated with a 10.6% higher probability of switching, while the increase for the footnotes is 9.4%. These results support H1, which states that clients are more likely to leave an incumbent auditor with which they are less compatible (i.e., are dissimilar to other clients of the same firm).[35]

4.2 HYPOTHESIS 2: SIMILARITY AND CHOICE OF NEW AUDITOR

Hypothesis 2 predicts that a company that hires a new auditor will select the firm that is most compatible from the set of nonincumbents available. Since the company has already decided that the net benefits of a change outweigh the switching costs—possibly due to fee, service, or compatibility considerations—it is expected that the new auditor, on average, will be the more compatible of the nonincumbent firms.

Table 3, panel C, summarizes the average compatibility rank of the *new* auditor (with 1 being most similar, 4 the least similar, and 2.5 as the middle

---

[33] On average, 7.0% of clients of nonexpert auditors switch each year versus 6.3% for clients of experts.

[34] In separate tests, we also include the standard deviation of client size within an auditor–industry–year as a control for diversity of the client base. We do not include this variable in the primary test because the standard deviation is calculated based on a relatively small number of clients in a number of cases. The control is significantly positive for the business description and MD&A, but not for the footnotes, but does not change our inferences.

[35] We also examine whether the similarity scores are related to the *reason* for the auditor change—resignation or dismissal—but find no evidence of a relationship between the two. So, while client similarity predicts an auditor change, it does not seem to predict whether the change is initiated by the auditor or the client, implying the lack of fit may be a concern to either party.

TABLE 4

Probability of Auditor Change

Panel A: Pairwise Pearson correlations for continuous variables

	SIM_{BUS}	$SIM_{MD\&A}$	SIM_{NOTES}	SIM_{COMB}	*SIZE*	*IRISK*	$\lvert DACC \rvert$	*CASH*	*ROA*	*GROWTH*
SIZE	**0.234**	**0.075**	**0.190**	**0.176**						
IRISK	**–0.178**	**–0.116**	**–0.113**	**–0.153**	**–0.074**					
$\lvert DACC \rvert$	**–0.036**	**–0.046**	**–0.063**	**–0.058**	**–0.125**	0.003				
CASH	**–0.052**	**0.089**	**–0.085**	**0.062**	**–0.390**	**–0.400**	**0.080**			
ROA	**0.056**	**0.039**	**0.106**	**0.060**	**0.406**	**0.166**	**–0.109**	**–0.295**		
GROWTH	**–0.024**	0.001	**–0.030**	–0.012	0.000	**–0.121**	**0.378**	**0.176**	**0.088**	
TENURE	0.016	–0.007	**0.095**	**0.048**	**0.283**	**0.115**	**–0.076**	**–0.157**	**0.126**	**–0.111**

Panel B: Logit model of auditor switch in subsequent year

		Bus Desc		MD&A		Footnotes		Combined	
	Exp	Coef	*z*-stat	Coef	*z*-stat	Coef	*z*-stat	Coef	*z*-stat
(Intercept)		–0.580	–1.98**	–0.300	–1.04	–0.121	–0.35	0.158	0.43
SIZE	–	–0.396	–17.95***	–0.411	–17.69***	–0.463	–16.65***	–0.452	–16.25***
IRISK	+	0.095	0.52	0.045	0.24	0.045	0.20	–0.063	–0.27
$\lvert DACC \rvert$	?	0.316	2.70***	0.314	2.76***	0.143	0.69	0.186	0.95
CASH	–	–0.875	–5.60***	–0.951	–5.96***	–1.091	–5.73***	–1.092	–5.56***
ROA	–	–0.091	–1.26	–0.096	–1.30	–0.074	–0.79	–0.100	–1.07
LOSS	+	0.382	6.21***	0.362	5.76***	0.424	5.65***	0.412	5.29***
GROWTH	?	–0.077	–2.19**	–0.054	–1.61	–0.062	–1.41	–0.065	–1.46
ACQUIS	+	0.041	0.48	0.111	1.30	0.089	0.85	0.065	0.59
CFEARLY	?	–0.157	–2.42**	–0.179	–2.70***	–0.200	–2.56**	–0.192	–2.41**
CFMATURE	–	–0.201	–2.69***	–0.218	–2.85***	–0.226	–2.49**	–0.231	–2.45**
ln(TENURE)	–	–0.130	–3.77***	–0.122	–3.44***	–0.109	–2.64***	–0.114	–2.69***
MODOPIN	+	0.261	4.29***	0.269	4.31***	0.272	3.75***	0.218	2.89***
EXPERT	–	–0.122	–1.66*	–0.048	–0.65	0.034	0.41	0.003	0.04
SIM_{BUS}	–	–0.439	–0.99						
$SIM_{MD\&A}$	–			–1.247	–3.07***				
SIM_{NOTES}	–					–2.575	–2.59***		
SIM_{COMB}	–							–0.027	–2.59***
Year FE		Yes		Yes		Yes		Yes	
Industry FE		Yes		Yes		Yes		Yes	
Obs		28,741		26,830		19,105		17,313	
Pseudo-R^2		0.09		0.09		0.11		0.10	

Panel A: Correlations in bold are significant at the 1% level. *SIM* = similarity to other clients in the auditor–industry–year reference group (subscripts: *BUS* = business description, *MD&A* = management's discussion and analysis, *NOTES* = footnotes to financial statements, *COMB* = combined disclosures). *SIZE* = log of total assets. *IRISK* = receivables plus inventory, scaled by assets. $\lvert DACC \rvert$ = absolute discretionary accruals from the cross-sectional modified Jones model. *CASH* = cash and equivalents, scaled by assets. *ROA* = income before extraordinary items, scaled by assets. *GROWTH* = change in assets, scaled by prior year assets. *ln(TENURE)* = log of number of years with the current auditor.

Panel B: The results of an auditor–industry–year-level logistic regression with the dependent variable of *SWITCH,* indicating an auditor change in the following year. *LOSS* = 1 if *ROA* < 0. *ACQUIS* = 1 if the acquisition activity in the current year exceeds 10% of assets. *CFEARLY* = 1 if cash flows indicate the client is in the introduction or growth stage of life cycle (Dickinson, [2011]). *CFMATURE* = 1 if the client is in the mature stage. *MODOPIN* = 1 for a nonstandard audit opinion. *EXPERT* = 1 if the auditor has at least 5% more clients than the next largest auditor in both the industry and the MSA. Standard errors clustered on firm and corrected for serial correlation. ***, **, * indicate the significance at 1%, 5%, and 10%, respectively.

case).[36] Consistent with the earlier results, we see that the likelihood that a firm is selected is monotonically increasing in the similarity scores. To

[36] The tests of the new auditor are only applied in markets where there are four audit firms present. Markets with only two firms do not represent a choice since there is only one alternative for the new auditor. We drop markets for three firms due to sample size considerations.

test the statistical significance of this pattern, the average rank of the new auditor is compared to what the average rank would be if a new auditor was randomly selected from the remaining three firms (i.e., 2.5). In table 3, panel D, the average rank for SIM_{BUS} is 2.39 ($t = 2.69$), for $SIM_{MD\&A}$ it is 2.28 ($t = 5.03$), and for SIM_{NOTES} it is 2.30 ($t = 3.50$). All three are significantly less than the rank expected for random selection—SIM_{COMB} is not significant—implying that clients are more likely to choose a better-fitting auditor when switching among the Big 4.

We next test the choice of new auditor using a multivariate model. Such an analysis is complicated by the fact that the new auditor must be selected from the three nonincumbent firms (i.e., the new auditor cannot be the same as the old auditor). For example, if Deloitte was the incumbent, then the new auditor must be selected from E&Y, KPMG, and PwC. This requires construction of four subsamples based on the identity of the incumbent firm. For the subsample of clients switching *from* Deloitte ($n = 406$), the dependent variable now has three possible values, so we use multinomial logit to analyze the firm's choice of the new auditor and include the similarity scores of the three options in the model. For ex-Deloitte clients, a positive coefficient on one of the other firm's similarity score corresponds to a higher probability of choosing that firm, while a negative coefficient indicates a lower probability.[37] For example, we expect that clients switching from Deloitte *to* KPMG will have a positive coefficient for SIM_{KPMG} but smaller or even negative coefficients for the other two firms. The results of this analysis are presented in table 5. Even though sample sizes are relatively small (ranging from an average of 91 observations for the footnotes to 148 for the business description), 18 of the significant coefficients are consistent with our expectations, while only two are marginally inconsistent (both at $p < 0.10$).[38]

Taken together, the univariate and multivariate tests of H1 and H2 indicate that clients will change auditors if they have low compatibility (similarity) with the other clients serviced by that auditor and will choose a new auditor from the nonincumbents that is generally more compatible (similar). That is, the results support the argument that auditor–client compatibility is partially the result of the auditor selection process, that is, when a company decides to change its auditor, it tends to gravitate to one that is relatively compatible. The pattern in figure 1 suggests that, on average, compatibility continues to grow over the tenure of the auditor. However, for the clients that fall in the lower portion of the similarity distribution

[37] The multinomial logit model requires one of the three options be considered the "base case." Following convention, we set the base case to the auditor with the largest number of observations, that is, all coefficients are relative to this case. Coefficients for the controls are not tabulated.

[38] The binomial probability of obtaining 18 positive results out of 20 tries if two states are equally likely is less than 0.001.

TABLE 5

Multinomial Logit Model of Auditor Choice

		Bus Desc		MD&A		Footnotes		Combined	
		Coef	*z*-stat	Coef	*z*-stat	Coef	*z*-stat	Coef	*z*-stat
Prior auditor: Deloitte									
New auditor: KPMG	$SIM_{E\&Y}$	5.954	0.58	–21.424	–1.66*	–8.293	–0.41	–0.638	–1.93*
	SIM_{KPMG}	12.078	1.61	16.121	1.91*	57.069	2.57**	0.482	1.89*
	SIM_{PwC}	–14.096	–1.31	8.275	0.90	–44.866	–1.99**	0.391	1.16
New auditor: PwC	$SIM_{E\&Y}$	5.471	0.63	–21.999	–1.79*	–33.777	–1.20	–0.160	–0.56
	SIM_{KPMG}	10.075	1.41	11.420	1.33	42.643	1.82*	0.386	1.60
	SIM_{PwC}	–13.010	–1.43	9.774	1.04	–12.899	–0.56	–0.235	–0.75
Prior auditor: E&Y									
New auditor: Deloitte	SIM_{DT}	8.112	1.02	2.995	0.41	21.033	1.15	0.421	1.56
	SIM_{KPMG}	–2.030	–0.30	2.015	0.23	–0.470	–0.03	0.139	0.66
	SIM_{PwC}	–4.162	–0.68	–1.238	–0.19	–7.139	–0.43	–0.446	–1.62
New auditor: KPMG	SIM_{DT}	14.529	1.78*	–3.059	–0.36	4.892	0.25	0.290	1.14
	SIM_{KPMG}	3.269	0.48	18.430	2.03**	12.162	0.94	0.200	0.99
	SIM_{PwC}	–16.661	–2.26**	–11.051	–1.36	–11.820	–0.66	–0.325	–1.21
Prior auditor: KPMG									
New auditor: Deloitte	SIM_{DT}	14.712	1.72*	14.039	1.36	20.189	1.23	0.251	1.47
	$SIM_{E\&Y}$	–2.555	–0.43	1.444	0.20	16.569	0.64	–0.313	–0.92
	SIM_{PwC}	–8.994	–1.10	–13.926	–1.45	–21.363	–1.07	0.188	0.63
New auditor: PwC	SIM_{DT}	3.993	0.40	4.653	0.40	5.695	0.33	0.023	0.12
	$SIM_{E\&Y}$	–19.674	–1.95*	1.979	0.23	6.378	0.25	–0.482	–1.56
	SIM_{PwC}	16.259	1.60	–9.879	–0.91	–7.408	–0.35	0.408	1.39
Prior auditor: PwC									
New auditor: Deloitte	SIM_{DT}	7.102	1.29	13.626	2.03**	–6.287	–0.40	0.294	1.50
	$SIM_{E\&Y}$	–7.125	–1.22	–18.607	–2.45**	–14.375	–0.76	–0.689	–2.86***
	SIM_{KPMG}	–3.112	–0.69	5.867	1.28	13.265	0.93	0.411	2.31**
New auditor: KPMG	SIM_{DT}	–5.233	–0.89	0.768	0.14	9.439	0.53	0.152	0.67
	$SIM_{E\&Y}$	3.755	0.70	–4.043	–0.64	–38.673	–1.69*	–0.458	–1.84*
	SIM_{KPMG}	0.208	0.05	5.659	1.32	33.700	2.36**	0.400	1.90*
Obs		148		137		91		87	
Pseudo-R^2		0.13		0.14		0.25		0.04	

Results of a multinomial logit model with the new auditor for the next year (one of the Big 4) as the dependent variable. The model is executed four times for each similarity score, once for each auditor the client is switching *from* since that auditor is not an available choice. The base case for each model is the auditor with the greatest number of clients, typically E&Y. Coefficients in boxes are significant and consistent with our exceptions; other significant coefficients are inconsistent. *SIM* = similarity to other clients in the current auditor–industry–year reference group. *SIM* subscripts indicate the current-year auditor used for comparison. The following control variables were included in the model, but omitted from the table: *SIZE, IRISK,* |*DACC*|, *CASH, ROA, LOSS, GROWTH, ACQUIS, CFEARLY, CFMATURE, ln(TENURE), MODOPIN,* and *EXPERT*. ***, **, * indicate the significance at 1%, 5%, and 10%, respectively.

(figure 2), it is likely that they will eventually change to a more compatible auditor.

4.3 HYPOTHESIS 3: SIMILARITY AND AUDITOR CHOICE AT THE LOCAL OFFICE LEVEL

Hypotheses 1 and 2 predict auditor choice at the national level. We now repeat those tests at the office level. To do so, we recalculate the three similarity scores by comparing each client to its auditor–MSA–year reference group (in contrast to the national level test that used an auditor–industry–year reference group): $SIMMSA_{BUS}$, $SIMMSA_{MD\&A}$, and $SIMMSA_{NOTES}$.[39] These scores proxy for how similar the client is to other clients of the same auditor in the same local office, rather than requiring the same industry. Consistent with the national level test, table 6, panel A, shows a monotonically increasing preference for higher-ranked auditors at the office level across all three disclosures. The average auditor ranks are all significantly less than 2.5 ($SIMMSA_{BUS} = 2.34$; $SIMMSA_{MD\&A} = 2.33$; $SIMMSA_{NOTES} = 2.27$, $SIMMSA_{COMB} = 2.34$). Although the evidence is somewhat weaker when considering changes, the same general pattern exists in choosing a new auditor. $SIMMSA_{BUS}$ (2.42) and $SIMMSA_{NOTES}$ (2.38) have average ranks that are significantly less than 2.5 (panel D), while $SIMMSA_{MD\&A}$ and $SIMMSA_{COMB}$ are not significant. This variation in results could be due to differences in the industry makeup of each audit firm in a given MSA, which we cannot control for, and may reduce the comparability of individual financial statements. However, there may also be additional preferences and constraints that influence choices at the office level such as existing relationships with audit partners, intensified competitive concerns due to geographical proximity, and auditor capacity constraints.

To test the relation between client similarity and auditor switching at the office level, we expand model (1) with two additional controls. *OFFICESIZE* is the number of clients in the auditor's local office scaled by the total number of clients in the geographical MSA. *MKTSIZE* is the number of Big 4 auditors having five or more clients in the same metropolitan area. Table 7 contains the results of this logit model. While neither office variable is significant, $SIMMSA_{BUS}$ ($p = 0.052$), $SIMMSA_{MD\&A}$ ($p = 0.030$), and $SIMMSA_{COMB}$ ($p = 0.01$) are significantly negative, supporting the hypothesis that clients having poorer auditor fit with the local auditor office are more likely to change auditors. These results suggest that both national and local attributes of the audit firm influence auditor switching decisions, consistent with H3.

[39] Due to sample size considerations, we are unable to calculate an auditor–industry–year–MSA similarity because very few markets have an adequate number of clients (i.e., the calculation requires a minimum of five clients for each auditor–industry–year–MSA nexus) and inclusion of both industry and MSA would lead to a loss of 93% of the sample. This dramatic fall in sample size is consistent with a GAO [2008] report that found that 82% of Fortune 1000 firms feel that they have three or fewer options when selecting an auditor. Variations across industries at the office level reduce the likelihood of finding significant results.

TABLE 6

Auditor Selection Based on Auditor–Client Compatibility Within a Geographical Area

Panel A: Rank of incumbent auditor based on similarity to auditor's client base

	Bus Desc			MD&A			Footnotes			Combined		
	Freq	%	Cml	Freq	%	Cml	Freq	%	Cml	Freq	%	Cml
1 (similar)	6,759	31	31	6,322	31	31	4,628	33	33	4,021	32	32
2	5,648	26	56	5,248	25	56	3,630	26	59	3,013	24	55
3	5,063	23	79	4,870	24	80	3,203	23	82	3,100	24	80
4 (different)	4,605	21	100	4,175	20	100	2,599	18	100	2,610	20	100
Total observations	22,075			20,615			14,060			12,744		

Panel B: Average rank of incumbent auditor

	Avg Rank	t-stat
Bus Desc	2.34***	21.18
MD&A	2.33***	21.32
Footnotes	2.27***	24.82
Combined	2.34***	16.33

Panel C: Rank of new auditor following a Big 4-to-Big 4 auditor change

	Bus Desc			MD&A			Footnotes			Combined		
	Freq	%	Cml	Freq	%	Cml	Freq	%	Cml	Freq	%	Cml
1 (Similar)	201	36	36	195	36	36	129	38	38	94	30	30
2.5	182	32	68	174	32	69	100	29	67	105	34	64
4 (Different)	178	32	100	167	31	100	111	33	100	110	36	100
Total Obs	561			536			340			309		

Panel D: Average rank of new auditor following a Big 4-to-Big 4 auditor change

	Avg Rank	t-stat
Bus Desc	2.42*	1.54
MD&A	2.46	0.81
Footnotes	2.38**	1.79
Combined	2.54	0.64

Similar to table 3, but for an auditor–MSA–year instead of an auditor–industry–year.

Panel A: Freq = number of times an auditor of a given rank is engaged by a client. Rank 1 corresponds to the most compatible auditor, while rank 4 indicates the most incompatible auditor. % = percentage of client-years engaging that rank. Cml = cumulative frequency. Only markets with four auditors are included; including markets with only two or three auditors yields qualitatively identical results.

Panel B: Avg Rank = average rank of the auditor engaged by a client. Random choice (null hypothesis) is 2.5.

Panel C: Similar to panel A, but only for clients changing from one Big 4 auditor to another Big 4 auditor in the year of the change. In contrast to panel A, the possible ranks are 1, 2.5, and 4 because only three auditors are available.

Panel D: Similar to panel B, but only for clients changing auditors.

***, **, * indicate significance at 1%, 5%, and 10%, respectively.

4.4 HYPOTHESIS 4: SIMILARITY AND AUDIT QUALITY

We test the possible link between similarity and audit quality using financial restatements, discretionary accruals, and GCOs.

4.4.1. Restatements. For our first measure of audit quality, we examine restatements of a prior year's earnings using the following model of Dechow,

TABLE 7
Probability of Auditor Change Within a Geographical Area

		Bus Desc		MD&A		Footnotes		Combined	
	Exp	Coef	*z*-stat	Coef	*z*-stat	Coef	*z*-stat	Coef	*z*-stat
(Intercept)		–0.728	–1.89*	–0.696	–1.75*	0.134	0.27	1.060	1.71*
SIZE	–	–0.394	–16.00***	–0.408	–15.71***	–0.509	–16.10***	–0.495	–15.45***
IRISK	+	0.154	0.75	0.065	0.31	0.046	0.18	–0.002	–0.01
|*DACC*|	?	0.250	1.69*	0.239	1.69*	–0.154	–0.52	–0.075	–0.26
CASH	–	–0.845	–4.91***	–0.875	–4.94***	–0.989	–4.64***	–0.925	–4.19***
ROA	–	–0.128	–1.57	–0.125	–1.52	–0.111	–1.01	–0.123	–1.10
LOSS	+	0.440	6.35***	0.445	6.26***	0.503	5.81***	0.467	5.16***
GROWTH	?	–0.079	–2.11**	–0.057	–1.63	–0.048	–0.91	–0.046	–0.86
ACQUIS	+	0.043	0.45	0.144	1.53	0.146	1.19	0.117	0.92
CFEARLY	?	–0.095	–1.30	–0.111	–1.51	–0.149	–1.68*	–0.167	–1.83*
CFMATURE	–	–0.102	–1.20	–0.116	–1.33	–0.120	–1.14	–0.157	–1.43
ln(TENURE)	–	–0.142	–3.65***	–0.132	–3.26***	–0.103	–2.19**	–0.109	–2.20**
MODOPIN	+	0.279	4.06***	0.262	3.73***	0.315	3.79***	0.247	2.84***
EXPERT	–	–0.147	–1.66*	–0.131	–1.45	–0.005	–0.05	0.132	1.10
OFFICESIZE	?	0.001	0.49	0.002	1.15	–0.002	–0.72	–1.201	–2.75***
MKTSIZE	+	0.003	0.05	0.035	0.57	–0.021	–0.27	–0.093	–0.96
$SIMMSA_{BUS}$	–	–1.882	–2.09**						
$SIMMSA_{MD\&A}$	–			–1.591	–2.31**				
$SIMMSA_{NOTES}$	–					–1.576	–0.96		
$SIMMSA_{COMB}$	–							–0.032	–2.60***
Year FE		Yes		Yes		Yes		Yes	
Industry FE		Yes		Yes		Yes		Yes	
Obs		22,654		20,918		14,742		13,350	
Pseudo-R^2		0.09		0.09		0.12		0.11	

The results of an auditor–MSA–year-level logistic regression with the dependent variable of *SWITCH*, indicating an auditor change in the following year. *SIMMSA* = similarity to other clients in the auditor–MSA–year reference group (subscripts: *BUS* = business description, *MD&A* = management's discussion and analysis, *NOTES* = footnotes to financial statements, *COMB* = combined disclosures). *SIZE* = log of total assets. *IRISK* = receivables plus inventory, scaled by assets. |*DACC*| = absolute discretionary accruals from the cross-sectional modified Jones model. *CASH* = cash and equivalents, scaled by assets. *ROA* = income before extraordinary items, scaled by assets. *LOSS* = 1 if *ROA* < 0. *GROWTH* = change in assets, scaled by prior year assets. *ACQUIS* = 1 if the acquisition activity in the current year exceeds 10% of assets. *CFEARLY* = 1 if cash flows indicate the client is in the introduction or growth stage of life cycle (Dickinson, [2011]). *CFMATURE* = 1 if the client is in the mature stage. *ln(TENURE)* = log of number of years with the current auditor. *MODOPIN* = 1 for a nonstandard audit opinion. *EXPERT* = 1 if the auditor has at least 5% more clients than the next largest auditor in both the industry and the MSA. *OFFICESIZE* = number of clients in the auditor's local office. *MKTSIZE* = number of Big 4 auditors having five or more clients in the same MSA as the observation. Standard errors clustered on firm and corrected for serial correlation. ***, **, * indicate the significance at 1%, 5%, and 10%, respectively.

Ge, and Schrand [2010]:

$$\begin{aligned} RESTATE = {} & a_0 + a_1 SIM + a_2 SIZE + a_3 LEV + a_4 ROA + a_5 LOSS \\ & + a_6 IRISK + a_7 CFEARLY + a_8 CFMATURE + a_9 M2B \\ & + a_{10} RETVOL + \varepsilon. \end{aligned} \tag{2}$$

SIM is again a placeholder for the various similarity scores tested in this paper. We use restatement data from Audit Analytics, setting *RESTATE* to 1 if there is a subsequent restatement of the current year's

TABLE 8
Probability of Restating Financials

Panel A: Descriptive statistics for new variables

Variable	Mean	Std Dev	25%	Median	75%	*N*
RESTATE	0.102	0.303	0.000	0.000	0.000	35,593
M2B	1.919	4.935	0.537	1.049	2.015	33,345

Panel B: Logit model of current year financial statements being restated in the future

		Bus Desc		MD&A		Footnotes		Combined	
	Exp	Coef	*z*-stat	Coef	*z*-stat	Coef	*z*-stat	Coef	*z*-stat
(Intercept)		–2.913	–28.4***	–3.028	–28.22***	–2.737	–20.73***	–2.922	–19.94***
SIZE	+	0.094	7.95***	0.114	8.85***	0.087	5.61***	0.085	5.10***
LEV	+	0.055	0.57	0.129	1.30	0.188	1.66*	0.177	1.46
ROA	?	0.359	3.4***	0.297	2.89***	0.494	3.66***	0.466	3.34***
LOSS	+	0.085	1.53	0.092	1.64	0.141	2.17**	0.136	1.98**
IRISK	+	0.187	1.72*	0.244	2.21**	0.099	0.75	0.190	1.39
CFEARLY	+	0.087	1.56	0.051	0.90	0.068	1.06	0.033	0.50
CFMATURE	–	–0.108	–1.73*	–0.130	–2.02**	–0.110	–1.53	–0.136	–1.82*
M2B	+	–0.002	–0.26	0.001	0.09	–0.031	–2.08**	–0.024	–1.64
RETVOL	+	0.675	4.52***	0.559	3.77***	0.819	3.98***	0.792	3.80***
SIM_{BUS}	?	1.669	6.62***						
$SIM_{MD\&A}$	?			2.079	9.81***				
SIM_{NOTES}	?					–0.835	–1.54		
SIM_{COMB}	?							0.021	2.90***
Obs		28,678		26,578		19,019		17,239	
Pseudo-R^2		0.01		0.01		0.01		0.01	

Panel A: *RESTATE* = 1 if the current year financials are restated in a later period. *M2B* = market-to-book ratio.

Panel B: The results of a logistic regression with the dependent variable of *RESTATE*, indicating the client will restate the current year's financials at a later time. *SIZE* = log of total assets. *LEV* = long-term debt, scaled by assets. *ROA* = income before extraordinary items, scaled by assets. *LOSS* = 1 if *ROA* < 0. *IRISK* = receivables plus inventory, scaled by assets. *CFEARLY* = 1 if cash flows indicate the client is in the introduction or growth stage of life cycle (Dickinson [2011]). *CFMATURE* = 1 if the client is in the mature stage. *RETVOL* = standard deviation of monthly excess returns for the prior 12 months. ***, **, * indicate the significance at 1%, 5%, and 10%, respectively.

financials. Larger firms (*SIZE*) have more opportunities for misstatements due to greater complexity and may receive more scrutiny from regulators and market participants. Long-term debt scaled by total assets (*LEV*) proxies for reporting complexity and potential financial pressure on the client. We include a variety of risk proxies that parallel those in the auditor change model. *LOSS* and *IRISK* capture client risk from an auditor perspective. Younger companies (*CFEARLY*) are more likely to restate because of new and rapidly changing firm conditions, while mature clients (*CFMATURE*) have had more time to stabilize their financial situation and reporting requirements. The market-to-book ratio (*M2B*) controls for market-based pressure on clients to report positive performance. Finally, we include the standard deviation of monthly returns over the prior 12 months (*RETVOL*) to capture company risk as reflected in stock return volatility.

Panel A of table 8 contains descriptive results for the new variables, indicating that restatements occur for approximately 10% of firm-years in our sample.[40] The results of the logistic regressions, including each similarity score, are reported in panel B. The signs on the coefficients are significantly positive for SIM_{BUS} (z = 6.62), $SIM_{MD\&A}$ (z = 9.81), and SIM_{COMB} (z = 2.90), but insignificant for SIM_{NOTES}.[41] These results could be due to bonding arising from increased auditor–client compatibility. Alternatively, they could also be due to a "contagion" effect across clients if an accounting or auditing position of the firm is challenged by the SEC. Interestingly, the similarity score (SIM_{NOTES}) most related to client accounting decisions, and the only one directly influenced by the audit itself, is not significant.[42] Thus, we find some evidence that similarity is associated with a loss of audit quality but only in regards to unaudited disclosures.

4.4.2. Discretionary Accruals. Another common proxy for audit quality is discretionary accruals, with income-decreasing accruals representing conservatism and income-increasing accruals corresponding to aggressive accounting practices. Because conservatism is unlikely to be viewed as negatively by an auditor as income-increasing accruals, we focus on the latter. For these tests, we use the same measure of discretionary accruals (*DACC*) as used in testing H1 and H3, but set all income-decreasing discretionary accruals to zero. Our measure of income-increasing accruals (*IIDACC*) is negatively correlated (untabulated) with all three similarity measures (–0.066 for SIM_{BUS}; –0.071 for $SIM_{MD\&A}$; –0.086 for SIM_{NOTES}). We use the following OLS model to test whether similarity is related to income-increasing accruals:

$$\begin{aligned} IIDACC = {} & \mathrm{a}_0 + a_1 SIM + a_2 GROWTH + a_3 CFO + a_4 LITIG + a_5 LOSS \\ & + a_6 LEV + a_7 M2B + a_8 SIZE + a_9 FIN + a_{10} TACC\_LAG + \varepsilon. \end{aligned} \tag{3}$$

CFO is cash flow from operations, scaled by total assets, and *LITIG* is a dummy set to 1 for litigious industries.[43] *FIN* is a dummy set to 1 if debt issuances exceed 20% of client assets or equity issuances exceed 10%. Finally,

---

[40] This relatively high restatement rate of 10% implies that some of the restatements are likely minor or technical in nature, despite Audit Analytics' attempts to filter out these types of restatements. Some research (Hennes, Leone, and Miller [2008]) has found differences between restatements due to "errors" and those due to "irregularities." Our results are not sensitive to excluding various types of less serious issues.

[41] In an untabulated test, we include variables for whether the CEO/chairman roles are combined and a proxy for board independence (from the Investor Responsibility Research Center). Including these variables results in an average sample size reduction of 64%. Their presence does not change the qualitative conclusions.

[42] As an alternative test, we proxy for audit failure with the presence of SEC Accounting and Auditing Enforcement Releases (AAERs). After expanding the set of controls to include those used in the Dechow, Larson, and Sloan [2011] AAER model, and including each of the similarity scores in turn, only $SIM_{MD\&A}$ is significant (and positive).

[43] Litigious industries are the following SIC codes: 2833–2836, 3570–3577, 3600–3674, 5200–5961, or 7370–7374 (see Francis, Philbrick, and Schipper [1994]).

TABLE 9
Accruals-Based Audit Quality

	Bus Desc		MD&A		Footnotes		Combined	
	Coef	*z*-stat	Coef	*z*-stat	Coef	*z*-stat	Coef	*z*-stat
(Intercept)	0.13	23.22***	0.132	22.61***	0.124	16.60***	0.141	17.03***
GROWTH	0.039	13.96***	0.039	13.14***	0.043	12.42***	0.043	12.37***
CFO	–0.04	–4.77***	–0.034	–4.30***	–0.032	–3.14***	–0.028	–2.88***
LITIG	0.009	3.36***	0.010	3.58***	0.010	3.09***	0.010	2.83***
LOSS	–0.012	–3.71***	–0.011	–3.41***	–0.009	–1.98**	–0.009	–1.84*
LEV	0.035	3.85***	0.029	3.59***	0.043	3.74***	0.044	3.74***
M2B	0.001	1.75*	0.001	1.21	0.001	0.77	0.001	1.01
SIZE	–0.012	–12.7***	–0.012	–12.32***	–0.011	–9.26***	–0.012	–9.31***
FIN	0.003	0.75	0.003	0.85	0.000	0.00	0.002	0.39
TACC_LAG	–0.018	–2.66***	–0.028	–2.30**	–0.025	–2.22**	–0.022	–2.22**
SIM_{BUS}	–0.028	–1.95*						
$SIM_{MD\&A}$			–0.079	–6.33***				
SIM_{NOTES}					–0.168	–6.82***		
SIM_{COMB}							–0.001	–3.38***
Obs	12,642		11,899		8,736		7,877	
Adjusted-R^2	0.12		0.13		0.12		0.12	

The results of an OLS regression with the dependent variable of *IIDACC,* indicating income-increasing discretionary accruals. *IIDACC* = calculated the same as |*DACC*|, except that negative values of *DACC* are replaced with zero. *GROWTH* = change in assets, scaled by prior year assets. *CFO* = cash flow from operations, scaled by assets. *LITIG* = 1 for the following high-litigation SIC codes: 2833–2836, 3570–3577, 3600–3674, 5200–5961, or 7370–7374. *LOSS* = 1 if *ROA* < 0. *LEV* = long-term debt, scaled by assets. *M2B* = market-to-book ratio. *SIZE* = log of total assets. *FIN* = 1 if debt issuance in the current year exceeds 20% of assets or equity issuance exceeds 10% of assets. *TACC_LAG* = total accruals from the prior year. Standard errors clustered on firm and corrected for serial correlation. ***, **, * indicate the significance at 1%, 5%, and 10%, respectively.

we include lagged total accruals (*TACC_LAG*). *GROWTH*, *LOSS*, *LEV*, *M2B*, and *SIZE* retain their same definitions from earlier tests.

We execute the model separately for each of the similarity scores, reporting the results in table 9. All four similarity scores are significantly negative.[44] This pattern suggests that higher similarity scores are associated with lower income-increasing discretionary accruals. To the extent that accruals proxy for audit quality, this result implies that high similarity is associated with *higher* audit quality. Taken together, the accrual and restatement results may suggest that similarity has a positive effect on the portions of the financial statements for which the auditor is most responsible (the footnotes and financial numbers) but a potentially negative effect on the unaudited portion of the text disclosures in the financial statements. This result is consistent with current discussions among regulators as to whether the auditor should assume more responsibility related to MD&A disclosures (see Goezler [2011]).

[44] Repeating the test with absolute discretionary accruals (i.e., both income-increasing and income-decreasing accruals are considered signs of poor audit quality) gives similar results with the exception that SIM_{BUS} is no longer significant.

4.4.3. Going Concern Opinions. The relative propensity of an auditor to issue GCOs to clients is often considered to be a proxy for higher audit quality (Knechel et al. [2013]) based on the logic that, ceteris paribus, an auditor whose independence is impaired is less likely to issue a GCO. However, when considered in the context of "similarity," the interpretation of a GCO as a proxy for audit quality may be more nuanced because a client that is a candidate for a GCO will likely exhibit signs of financial distress that may make it appear to be less similar to financially healthy firms in the auditor's portfolio. While a reduced incidence of GCOs may indicate a loss of audit quality in many settings, the opposite may be true in our study if more highly distressed clients have more distinct and informative text disclosures due to their distress. Therefore, auditors may be more likely to issue a GCO to clients with the greatest dissimilarity if that lack of similarity can be traced to financial distress.[45]

To test the link between going concern reports and our proxy for similarity, we develop the following logit model based on results in prior literature (DeFond, Raghunandan, and Subramanyam [2002], Callaghan, Parkash, and Singhal [2009], Blay and Geiger [2013]):

$$\begin{aligned} GCO = {} & a_0 + a_1 SIM + a_2 SIZE + a_3 LEV + a_4 ROA + a_5 CASH + a_6 CFO \\ & + a_7 LOSS + a_8 DEBTISS + a_9 EQISS + a_{10} CFEARLY \\ & + a_{11} ZMIJPROB + a_{12} RET + a_{13} RETVOL + a_{14} DELAY + \varepsilon. \end{aligned} \tag{4}$$

To implement the model, we use a subsample of financially distressed manufacturing firms (SIC 20–39). We consider a firm to be financially distressed if it has negative net income *and* negative cash flows from operations.[46]

We use Audit Analytics to determine if the company received a going concern opinion (*GCO*) from the auditor. We control for various client characteristics that are generally associated with potential financial distress. *LEV* is the firm's long-term debt, scaled by total assets. Higher cash flow from operations (*CFO*) is generally associated with a lower propensity to receive a GCO. Debt (*DEBTISS*) and equity (*EQISS*) issuances are common in smaller, less stable companies, as well as those that are experiencing insufficient cash flow to sustain the business. Based on Dickinson (2011), *CFEARLY* is a dummy variable equal to 1 for companies in an introductory or growth phase of their life cycle. *ZMIJPROB* is the probability of an impending bankruptcy using the traditional Zmijewski [1984] score. We also control for the excess equity market return for the fiscal year (*RET*) and the standard deviation of monthly returns over the same period (*RETVOL*)

---

[45] Alternatively, a firm under financial pressure that the auditor concludes is not a going concern risk may not need particularly unique disclosures.

[46] Results do not qualitatively change if we consider either condition of financial distress separately.

as proxies for risk. Finally, we include the number of days between the end of the fiscal year and the filing of the 10-K (*DELAY*) to control for the complexity related to the audit and the issuance of the GCO itself.[47]

We expand the base model with each of the similarity scores with the results presented in table 10. $SIM_{MD\&A}$, SIM_{NOTES}, and SIM_{COMB} are significantly negative. These results suggest that clients that are *more* similar to others in a firm's portfolio are *less* likely to receive a GCO. The business description coefficient is not significant. Using a traditional interpretation of the incidence of GCOs, these results suggest that similarity is associated with lower quality. However, this interpretation needs to be tempered by the fact that firms in financial distress, by their very nature, are also less likely to be similar to other firms audited by the same auditor (i.e., less financially healthy). In fact, financial distress may be a condition where a specific client could be *expected to be dissimilar* from an auditor's overall portfolio.[48]

Panel C of table 10 reports descriptive results for three subsamples in our analysis: (1) nondistressed firms, (2) distressed firms that do *not* receive a GCO, and (3) distressed firms that receive a GCO. First, we can see that the subsamples are significantly different across virtually all variables used in the analysis. For example, the GCO firms are smaller, have higher levels of accruals, and exhibit higher *Z*-scores than the other two groups. Of relevance to our analysis is that nondistressed firms are significantly more similar to each other than to either set of distressed firms.[49] The distressed firms actually exhibit *negative* similarity for five of the six raw similarity scores. Also, firms receiving a GCO have a significantly lower composite similarity score (7.62) than the other two groups (9.33, 9.21) and exhibit significantly different business descriptions, footnotes, and MD&A from firms that do not receive a GCO (all $p < 0.001$). This result suggests that auditors issue going concern reports to clients that are unlike others in their portfolio in terms of financial performance and prospects, resulting in differences in their text disclosures also.

Finally, we expand our tests of GCO reporting by considering if similarity is associated with more accurate reporting of going concerns (Bruynseels, Knechel, and Willekens [2011]). For this test, we replace the dependent variable in equation (4) with *GCO_FAILURE*, which takes a value of 1 if a GCO is *not* followed by a bankruptcy or a subsequently bankrupt company did *not* receive a GCO (i.e., "inaccurate" opinions). In contrast,

---

[47] All continuous variables are winsorized at the 1st and 99th percentiles.

[48] We also reran our analysis using only firms that received a first time GCO. SIM_{BUS} continues to be insignificant, $SIM_{MD\&A}$ is still significant and negative at $p < 0.001$, but SIM_{NOTES} is no longer significant. SIM_{COMB} is significant and negative at the 10% level. Full results are included in supplemental table R1.

[49] In an alternative test, we rerun our primary analysis of the likelihood of changing auditors after dropping all firms that receive a GCO from the sample. Our results are very similar to the primary results: SIM_{BUS} remains insignificant, SIM_{NOTES} is now significant at 5% rather than 1%, $SIM_{MD\&A}$ remains significant at 1%, and SIM_{COMB} is now significant at 10% rather than 1% (although the coefficient is nearly identical to the primary analysis).

TABLE 10

Probability of Distressed Firms Receiving a Going Concern Report

Panel A: Descriptive statistics for new variables

Variable	Mean	Std Dev	25%	Median	75%	*N*
GCO	0.050					28,397
LEV	0.206	0.254	0.001	0.119	0.321	35,457
CFO	0.013	0.243	−0.008	0.068	0.125	35,400
DEBTISS	0.111	0.223	0.000	0.002	0.119	33,955
EQISS	0.079	0.204	0.001	0.007	0.029	34,827
CFEARLY	0.452					35,593
ZMIJPROB	0.271	0.350	0.005	0.080	0.463	35,235
RET	0.023	0.665	−0.320	−0.014	0.302	31,552
RETVOL	0.158	0.117	0.089	0.130	0.190	31,374
DELAY	79.047	19.905	69.000	80.000	90.000	35,593

Panel B: Logit model of going concern report in year of financial distress

	Bus Desc		MD&A		Footnotes		Combined	
	Coef	*z*-stat	Coef	*z*-stat	Coef	*z*-stat	Coef	*z*-stat
(Intercept)	−3.84	−5.85***	−3.915	−6.48***	−4.385	−5.70***	−4.458	−4.71***
SIZE	−0.168	−2.48**	−0.143	−2.07**	−0.286	−3.48***	−0.247	−2.91***
LEV	−0.928	−2.65***	−1.094	−3.07***	−0.570	−1.45	−0.520	−1.27
ROA	0.135	0.62	0.101	0.47	0.144	0.57	0.175	0.67
CASH	−3.043	−9.04***	−2.839	−8.55***	−2.917	−7.32***	−2.644	−6.30***
CFO	−2.076	−5.72***	−2.162	−6.02***	−2.302	−5.49***	−2.337	−5.39***
LOSS	0.081	0.34	0.105	0.43	0.258	0.78	0.283	0.82
DEBTISS	0.143	0.42	0.284	0.83	0.292	0.73	0.193	0.46
EQISS	−0.064	−0.26	−0.115	−0.47	0.137	0.51	0.146	0.52
CFEARLY	−0.153	−0.99	−0.106	−0.68	−0.054	−0.30	−0.065	−0.34
ZMIJPROB	1.893	7.54***	1.837	7.38***	1.703	5.72***	1.699	5.57***
RET	−0.497	−5.29***	−0.503	−5.37***	−0.398	−3.57***	−0.379	−3.30***
RETVOL	2.454	4.48***	2.366	4.36***	2.729	3.99***	2.756	3.74***
DELAY	0.016	3.07***	0.014	3.20***	0.024	4.16***	0.032	4.12***
SIM_{BUS}	−1.695	−1.34						
$SIM_{MD\&A}$			−4.637	−3.89***				
SIM_{NOTES}					−7.930	−2.12**		
SIM_{COMB}							−0.083	−3.01***
Obs	3,027		3,064		2,377		2,188	
Pseudo-R^2	0.35		0.36		0.39		0.39	

Panel C: Client descriptive statistics by financial condition

Variable	Nondistressed	Distressed No GCO	Distressed GCO	*F*	Prob > *F*
SIZE	6.221	4.605	3.820	3,754.00	0.000***
IRISK	0.273	0.172	0.235	629.79	0.000***
TACC	−0.047	−0.081	−0.101	107.65	0.000***
|*DACC*|	0.072	0.109	0.132	234.63	0.000***
CASH	0.168	0.466	0.331	4,168.84	0.000***
ROA	0.030	−0.389	−0.612	10,372.94	0.000***
LOSS	0.228	0.992	0.992	12,949.75	0.000***
GROWTH	0.216	0.535	0.344	167.01	0.000***
ACQUIS	0.145	0.062	0.077	170.41	0.000***
ZMIJPROB	0.213	0.405	0.612	2,271.31	0.000***
TENURE	10.347	6.964	6.609	579.90	0.000***

(Continued)

TABLE 10—*Continued*

Panel C: Client descriptive statistics by financial condition

Variable	Nondistressed	Distressed No GCO	Distressed GCO	*F*	Prob > *F*
MODOPIN	0.409	0.330	0.470	85.94	0.000***
SIM_{BUS}	0.0037	–0.0034	–0.0314	243.97	0.000***
$SIM_{MD\&A}$	0.0023	0.0030	–0.0281	139.85	0.000***
SIM_{NOTES}	0.0033	–0.0106	–0.0153	195.65	0.000***
SIM_{COMB}	9.3286	9.2050	7.6229	125.36	0.000***
Obs	15,525	3,302	1,143		

Panel D: Probability of receiving an *inaccurate* GCO report

	Bus Desc		MD&A		Footnotes		Combined	
	Coef	*z*-stat	Coef	*z*-stat	Coef	*z*-stat	Coef	*z*-stat
(Intercept)	–5.792	–9.32***	–5.861	–10.17***	–6.043	–8.40***	–6.366	–7.07***
SIZE	–0.198	–2.69***	–0.181	–2.43**	–0.278	–3.34***	–0.246	–2.76***
LEV	–0.525	–1.52	–0.677	–1.91*	–0.440	–1.14	–0.367	–0.92
ROA	0.212	1.05	0.187	0.94	0.150	0.66	0.181	0.76
CASH	–2.674	–9.07***	–2.544	–8.62***	–2.538	–7.06***	–2.424	–6.43***
CFO	–1.878	–5.58***	–1.917	–5.76***	–2.078	–5.42***	–2.108	–5.25***
LOSS	2.58	9.31***	2.646	9.25***	2.341	6.23***	2.355	6.20***
DEBTISS	–0.097	–0.28	0.018	0.05	0.303	0.78	0.218	0.54
EQISS	–0.06	–0.27	–0.077	–0.36	0.033	0.14	0.047	0.18
CFEARLY	–0.073	–0.49	–0.044	–0.29	0.013	0.08	0.015	0.08
ZMIJPROB	1.456	5.83***	1.480	5.96***	1.438	4.98***	1.475	4.96***
RET	–0.394	–4.46***	–0.408	–4.62***	–0.308	–2.90***	–0.275	–2.48**
RETVOL	2.234	4.27***	2.136	4.12***	2.435	3.78***	2.435	3.55***
DELAY	0.014	2.88***	0.012	2.95***	0.020	3.94***	0.028	3.83***
SIM_{BUS}	–0.683	–0.56						
$SIM_{MD\&A}$			–2.997	–2.57**				
SIM_{NOTES}					–7.441	–2.03**		
SIM_{COMB}							–0.052	–1.91*
Obs	3,844		3,837		2,675		2,469	
Pseudo-R^2	0.32		0.33		0.35		0.36	

Panel A: *GCO* = 1 if the auditor issues a going concern qualification in the current year. *LEV* = long-term debt, scaled by assets. *CFO* = cash flow from operations, scaled by assets. *DEBTISS* = additional debt issued, scaled by assets. *EQISS* = additional equity issued, scaled by assets. *CFEARLY* = 1 if cash flows indicate the client is in the introduction or growth stage of life cycle (Dickinson [2011]). *ZMIJPROB* = probability of bankruptcy using the Zmijewski score. *RET* = excess return over a market model for the fiscal year. *RETVOL* = standard deviation of monthly excess returns for the prior 12 months. *DELAY* = days between the fiscal year end and filing of a 10-K.

Panel B: The results of a logistic regression with the dependent variable of *GCO*, indicating the presence of a modified GCO in the auditor's report. The sample contains distressed manufacturing firms (SIC 20–39), defined as those having both negative net income and negative cash flows from operations. *SIZE* = log of total assets. *ROA* = income before extraordinary items, scaled by assets. *CASH* = cash and equivalents, scaled by assets. *LOSS* = 1 if *ROA* < 0. *SIMQ1* = 1 if the corresponding similarity score is in the first (lowest) quintile.

Panel C: Test of joint difference in means between the three classes of firms.

Panel D: The results of a logistic regression with the dependent variable of *GC_FAILURE*, indicating the issuance of a GCO but no bankruptcy within 12 months or the lack of a GCO but a bankruptcy within 12 months. The sample contains distressed manufacturing firms (SIC 20–39), defined as those having both negative net income and negative cash flows from operations. *SIZE* = log of total assets. *LEV* = long-term debt, scaled by assets. *ROA* = income before extraordinary items, scaled by assets. *CASH* = cash and equivalents, scaled by assets. *CFO* = cash flow from operations, scaled by assets. *LOSS* = 1 if *ROA* < 0. *DEBTISS* = additional debt issued, scaled by assets. *EQISS* = additional equity issued, scaled by assets. *CFEARLY* = 1 if cash flows indicate the client is in the introduction or growth stage of life cycle (Dickinson [2011]). *ZMIJPROB* = probability of bankruptcy using the Zmijewski score. *RET* = excess return over a market model for the fiscal year. *RETVOL* = standard deviation of monthly excess returns for the prior 12 months. *DELAY* = days between the fiscal year end and filing of a 10-K.

***, **, * indicate the significance at 1%, 5%, and 10%, respectively.

GCO_FAILURE takes a value of zero if a bankrupt company received a GCO or a nonbankrupt company received an unmodified opinion (i.e., "accurate" opinions). The results are reported in table 10, panel D. Three of the similarity scores are negative and significant ($SIM_{MD\&A}$, SIM_{NOTES}, and SIM_{COMB}). These results suggest that, while clients with higher similarity are less likely to receive a GCO, the accuracy of the GCOs issued (or not issued) is more accurate.

4.4.4. Discussion of Audit Quality Results. The results across the tests of audit quality are somewhat mixed. Accruals show that audit quality is positively associated with similarity, while restatements show the opposite, but only for the disclosures that are reviewed by an auditor and not audited (MD&A, business descriptions). The results for GCOs are more complicated because firms that are in financial trouble are not necessarily representative of, and thus less similar to, the healthy clients in an auditor's portfolio at that point in time. We find that MD&A disclosures and the footnotes are associated with fewer modifications (suggesting lower quality), but the same disclosures are associated with more accurate reporting (suggesting higher audit quality).[50] Further, the footnote results become insignificant in the case of *first time* GCOs. On balance, there is some evidence that audit quality improves with similarity. More specifically, there is evidence that audit quality is not undermined when there is similarity related to the aspects of the financial statements that are directly under the control of the auditor, that is, the footnotes and the accruals. However, in circumstances that might be deemed to be less frequent—restatements and GCOs—we find that similarity can be associated with some loss of audit quality. Of specific concern is that much of the evidence of a loss of audit quality is associated with similarity in the unaudited text disclosures (MD&A, business description).

5. *Supplemental and Robustness Analyses*

5.1 EFFECT OF SIZE

To verify whether the results hold for the largest or smallest companies, we repeat our tests on a subsample containing only the largest and smallest companies (based on assets by terciles). For the largest firms, this approximates the companies in the S&P 1500. The results for the smallest tercile are consistent with our primary analysis using the full sample, but, for the largest firms, the MD&A and footnote coefficients are no longer significant in our national level tests, although they are similar in magnitude to the original results. No coefficients are significant when repeating the test on the office-level auditor changes. Finally, for both the going concern and restatement tests of audit quality, the coefficients are similar in magnitude to

[50] The test of the propensity to issue a GCO can be considered a test of auditor independence, while the test of the accuracy of GCO reporting can be considered a test of auditor competence per the traditional definition of audit quality.

the original tests but somewhat weaker. Overall, the results are consistent with our original results in terms of coefficient magnitude, but not for statistical significance, which might indicate a loss of power arising from using a sample size only one-third as large as our original sample.

5.2 AUDITOR SWITCHING INVOLVING NON–BIG 4 FIRMS

Hypothesis 2 predicts that clients will prefer a better-fitting auditor when they decide to change auditors. The primary test only examined switches between Big 4 auditors and explicitly excludes clients switching from or to a non–Big 4 auditor. In a separate analysis (untabulated), we consider clients switching from a non–Big N auditor to a Big 4 auditor (i.e., "upward switches").[51] These clients have four auditors to choose from. Only SIM_{NOTES} implies upward switchers are more likely to choose a better-fitting auditor (rank = 2.31, $p = 0.009$). The other similarity scores are insignificant. This result suggests that the decision to move to a Big 4 auditor may be more dependent on other factors than compatibility but, all things equal, similarity can influence which Big 4 firm to choose once the decision is made to move to a Big 4 auditor.

5.3 EFFECT OF AUDITOR-PROVIDED NON-AUDIT SERVICES

One reason why a client may or may not switch auditors is because of existing contracts to provide NAS between an accounting firm and a potential client. We cannot observe NAS provided by firms that are not the auditor, but we can analyze whether auditor-provided NAS affect the likelihood of switching auditors conditional on the similarity scores. We do not consider NAS in our primary analysis because inclusion of NAS results in a significant reduction in our sample size since such fees were not disclosed in the early years of our sample. In untabulated analysis, we find that auditor-provided NAS have a negative effect on the likelihood of an auditor switch ($p < 0.001$ for all four measures).[52] However, the results for our similarity scores are not influenced by the inclusion of NAS in the analysis.

5.4 AUDITOR SWITCHING INVOLVING EX-ANDERSEN CLIENTS

Another form of switching observed during the time period of the sample was former Arthur Andersen clients switching to new auditors after the firm's collapse. Examining the new auditors of these firms, we observe that only $SIM_{MD\&A}$ is significant in predicting the new auditor (rank mean of 2.40, $p = 0.02$). This result may be due to relatively small sample sizes or can be caused by capacity constraints induced by the rapid auditor turnover affecting so many large clients at once. While the sample sizes are even smaller, we also look at ex-Andersen clients that switched a second time

[51] Note that we cannot test auditor switches from Big 4 to non–Big 4 firms because we do not know the similarity scores for the non–Big 4 firms that could have been selected as the new auditor. See supplemental table R2.

[52] See supplemental table R3.

in the sample period after the initial capacity constraints have presumably been resolved. We find that the rank of the new auditor is significantly less than 2.5 for SIM_{BUS} ($p = 0.05$). $SIM_{MD\&A}$ and SIM_{NOTES} are insignificant. Given the relative lack of statistical power in the test,[53] this provides some evidence that the initial capacity limitations affecting former Andersen clients' auditor choices were resolved at the time of later switches.

5.5 SIMILARITY BASED ON THE ENTIRE 10-K

In the primary analysis, we use three narrowly defined similarity scores, along with a composite measure, to maximize comparability across different companies within an industry. Mathematically, it is also possible to run our analysis on the entire 10-K. We do not include this as our primary analysis because of the likelihood that variations in auditor responsibility and presentation of such a wide variety of information may disguise the similarity that occurs in specific disclosure items. The full 10-K narrative yields a similarity score that is highly correlated with the disaggregated scores, ranging from 0.71 to 0.79. In untabulated analysis, we find results that are fully consistent with our primary analyses.[54] Specifically, we continue to find that companies are associated with the auditor that is most compatible, with an average rank of 2.34 ($t = 26.02$, $p < 0.001$). Further, they are significantly more likely to switch auditors when compatibility is low ($p < 0.001$). After switching auditors, they also choose a new auditor with relatively high compatibility (rank = 2.36, $p < 0.001$). Higher 10-K similarity is associated with more restatements ($t = 2.22$; $p = 0.03$), lower accruals ($t = 7.78$; $p < 0.001$), and fewer GCOs ($t = 1.72$, $p = 0.086$), consistent with our primary analysis.

5.6 COMPARABILITY OF ACCRUALS

Francis, Pinnuck, and Watanabe [2014] show that clients of a single auditor in the same industry have accruals that are more comparable to each other than to clients of other auditors. In their analysis, Francis, Pinnuck, and Watanabe [2014] examine measures of the comparability of total accruals and the comparability of discretionary accruals. We substitute their metrics in some of our models in place of our similarity scores. We find that neither comparability measure based on accruals is significantly related to changes in auditors. For the incidence of GCOs, we find that the comparability of total accruals is not significant while the comparability of discretionary accruals is positive ($p < 0.05$). For restatements, we find that both measures of comparability are marginally significant and negative ($p < 0.10$).[55] Taken together, these results suggest that our measures of similarity based on text disclosures are capturing something different in

---

[53] The number of observations ranges from 65 to 76, depending on the similarity score.

[54] See supplemental table R4.

[55] We do not repeat our accruals tests because the test variables are accrual measures themselves.

the client organizations from the Francis, Pinnuck, and Watanabe [2014] measure of accrual comparability.

5.7 ACCOUNTING SYSTEM COMPARABILITY

In a related line of research, De Franco, Kothari, and Verdi [2011] examine the comparability of accounting systems between companies. However, their metric is not based on financial disclosures or narrative text. For each company in their sample, they regress 16 quarters of earnings (considered to be an accounting system output) on returns (a measure of the net effect of economic events) to estimate the "accounting function" for that company. To determine the similarity between any two observations, they use the fitted accounting function to predict earnings for each observation using actual returns. They interpret the difference between the two predicted earnings values as a measure of the difference in accounting systems. Aggregating these differences for all pairs of observations gives a measure of accounting system similarity for each company within an industry-year (*COMPACCT-IND*). They construct an alternative measure using only earnings by regressing 16 quarters of earnings of one company on the earnings of another. Aggregating the R^2 from each regression also gives a proxy for accounting system similarity (*COMPACCT-R2*).

The *COMPACCT-IND* variable is uncorrelated with the three individual similarity scores but positively correlated with SIM_{COMB} (correlation of 0.02, significant at the 5% level). *COMPACCT-R2* has small but significantly positive correlations of 0.06, 0.07, 0.04, and 0.09 with SIM_{BUS}, $SIM_{MD\&A}$, SIM_{NOTES}, and SIM_{COMB}, respectively. As an alternative test of auditor changes, we separately include the two accounting comparability measures in the model for auditor switches. They are both insignificant. For the test of GCOs, only the coefficient on *COMPACCT-IND* is significantly positive at the 1% level. For restatements, both comparability scores are significantly negative at the 10% level. Finally, for the accrual test of audit quality, only the *COMPACCT-R2* measure is significantly negative at the 1% level.

6. *Conclusion*

The purpose of this paper is to examine the effect of client compatibility on auditor switching and audit quality. To measure compatibility, we compute the similarity of various narrative disclosures in the financial statements for a given auditor–industry–year nexus. We find evidence that companies are associated with auditors where they are similar to other clients of the firm in the same industry. Further, similarity increases over time for most clients of an audit firm. However, we also find evidence that clients that are dissimilar (low compatibility) with other clients of the audit firm are more likely to change auditors and will choose a new auditor with which they are more compatible. An interquartile shift in similarity with the current auditor's client base can change the probability of switching auditors by 9.4%–10.6%. These results apply if we also consider local market conditions. These results supplement the earlier analysis of Johnson and Lys [1990] because they suggest that the rationale for an accounting change can arise from both demand (quality) and supply (cost) reasons that are reflected in the comprehensive text disclosures of the financial statements. They also supplement papers such as Lennox and Park [2007] and He et al. [2014] that examine specific social links between auditors and clients. The advantage of our approach is that we adopt a measure that captures the multiple—but generally unobservable—dimensions of auditor–client compatibility.

We also address the concern that increased compatibility (similarity) may undermine audit quality if the observed similarity is a proxy for social or economic bonding. Here the results are somewhat mixed. Using accruals as a measure of audit quality, we find that similarity is associated with higher audit quality. In contrast, using accounting restatements, we find that similarity in terms of the unaudited (but auditor-reviewed) business description or MD&A is associated with reduced audit quality. However, similarity in the footnotes—the only disclosure in our analysis that is audited—is not associated with accounting restatements, suggesting that it is the unaudited portion of the financial statement that may exhibit a link between similarity and a potential loss of audit quality. Finally, we find that the issuance of a GCO is less likely when similarity is higher but the implied accuracy of GCO reporting is better. Since financially distressed firms are different from the average client in an auditor's portfolio, it is unclear whether these results suggest lower audit quality, or whether they arise because of the unique disclosures that may be needed when a company is having financial difficulties.

Our paper has a number of potential limitations that should be mentioned. First, our similarity score is an indirect measure of compatibility based on text-based financial statement disclosures and does not directly measure the specific attributes that might influence the compatibility of a client and auditor. In any given client–auditor relationship, some attributes of the auditor may be more important than others, and clients may value different attributes in different ways. For example, if the choice of an auditor is indeed optimal at the time of an auditor switch, our metric may be capturing something other than compatibility since it increases after the switch, while other research (Johnson and Lys [1990]) suggests it would likely decrease. Second, we assume that all nonincumbent Big 4 firms in an audit market are potential replacements for an incumbent auditor. However, some firms may be precluded from auditing a client due to existing non-audit service contracts, the relative size of a client, or competitive pressures among potential clients. These constraints would reduce the role of our similarity score in auditor–client alignment and suggest some counterexplanations for specific auditor changes. Further, there may be non–Big 4 firms in some markets that represent a viable alternative to Big 4 firms. Third, we exclude nonmanufacturing firms due to the high variability of financial reporting and disclosure issues. Fourth, since the estimation of similarity requires the use of a firm's entire client base, we cannot construct a hold-out sample to test the auditor choice models. The use of the same data for computing similarity and auditor choice could result in some overfitting of the choice model. Fifth, our sample selection process resulted in a slight overweighting of firms with modified opinions (40.2% in our sample versus 32% in Audit Analytics in general). Finally, our measures of audit quality have well-known flaws that may undermine the interpretation of our results (Knechel et al. [2013]).

In spite of these limitations, our findings have several implications for both regulators and researchers. First, the PCAOB and European Union continue to consider mandatory auditor rotation (PCAOB [2011]). Second, other commentators have raised concerns about the implications of another accounting firm collapse (Gerakos and Syverson [2015]) or regulators deciding to break up the Big 4 into smaller audit firms (U.K. House of Lords [2011]). Requiring an auditor change under such circumstances could force some auditors and clients into less compatible engagements. If like-clients naturally cluster with like-auditors for legitimate economic reasons, forcing clients that have a good fit to move to a less compatible auditor may raise audit costs and could influence audit quality. However, if similarity results in a loss of quality, the effect of mandatory audit firm rotation could be positive. Third, our results for audit quality are somewhat mixed but generally suggest that similarity only undermines audit quality when the disclosure is unaudited. This raises the possibility that standard setters might consider increasing the auditor's responsibility for verifying disclosures subject to review but not audit. Fourth, for researchers, prior literature has directly examined the effect of auditor type, such as auditor size and specialization, implicitly assuming auditors are indistinguishable within these groups (i.e., all Big N auditors are essentially the same). Our findings indicate significant heterogeneity among the Big 4 that could significantly influence auditor–client alignment. This finding is consistent with recent research on quality differences across audit partners (Gul, Wu, and Yang [2013], Knechel, Vanstraelen, and Zerni [2015]), and suggests that traditional measures of audit firm size and specialization may not capture the more subtle differences across firms that broader measures of comparability (Francis et al. [2014]) or compatibility/similarity (this paper) capture. Therefore, depending on the nature of the research question, it may be worthwhile to consider the differential effects of specific audit firms rather than examining them in such broad categories.

APPENDIX A

Extraction of Annual Report Items

To gather the business description, MD&A, and footnotes sample, we begin by downloading all 10-Ks and 10-K405s available on the SEC's EDGAR system that meet the following requirements: (1) fiscal years between 1997

TABLE A1
Narrative Disclosure Sample Selection Process

	Reports		
10-K available on EDGAR; fiscal years 1997–2009; Compustat assets > $1million; no FYE change; excluding financials and utilities; Big 4 auditor	41,622		
Less: Short reports (<50,000 characters)	(1,620)		
Total annual reports available	40,002		
	Bus Desc	MD&A	Footnotes
Less: Item not succesfully extracted	(1,890)	(905)	(7,692)
Less: Item specifically omitted	(10)	(42)	–
Less: Item included by reference	(23)	(2,836)	–
Less: Short items (<150 characters)	(1,275)	(1,168)	(2,077)
Less: <500 or >20,000 words	(888)	(1,247)	(5,213)
Less: Fewer than five other clients in auditor–industry–year	(2,623)	(2,591)	(2,994)
Total items available	33,293	31,213	22,026
At least five other clients in auditor–office–year	26,477	24,540	17,116

and 2009, (2) assets greater than $1 million, (3) no change in fiscal year-end, (4) not in the utilities or financial services industries, and (5) engaging a Big 4 auditor. As described in table A1, this initial screen leaves 41,622 annual reports. We next screen out any unusually short annual reports since these typically belong to holding companies, firms that are winding down, and other atypical observations. We use a cutoff of 50,000 characters for this purpose (approximately the fourth percentile of 10-K length). This filters out most of the unwanted observations without losing a substantial number of reports. We use characters instead of words because the tables and numbers contained in the report make it difficult to split the document into "words" at this point in the process. This leaves 40,002 annual reports.

Next, we strip all HTML formatting and data tables as in Li ([2008, 2010]) and split each annual report into its component items, keeping only the business description, MD&A, and footnotes (the financial statements are removed when data tables are discarded). We remove any narrative disclosures that contain language indicating that the relevant section has been omitted as permitted by regulation and skip disclosures that are included by reference, either to an external document or an attached exhibit, since the variety of alternate locations dramatically increases the difficulty in obtaining that data. The footnotes, in particular, are frequently included by reference. We drop any remaining items that do not contain at least 150 characters. Items shorter than this cutoff have typically been omitted or included by reference, but do so using somewhat unusual wording that the initial string search may not have recognized.

We then split each item into words, keeping disclosures with at least 500 words. Items shorter than this length are relatively unusual and are unlikely to provide a meaningful comparison to disclosures by peers in the auditor–industry–year reference group. Finally, we exclude items exceeding 20,000 words because these frequently indicate problems splitting the 10-K into separate items. For example, the extraction process might erroneously treat the entire annual report as the business description due to misspellings and other idiosyncratic document features. Archival studies frequently handle outliers such as these through deletion, winsorization, or robust techniques during the empirical analysis. However, doing so in the current study would allow these outliers to be in reference groups and therefore have an undesirable influence on the calculation of the similarity scores.

The overall sample selection is summarized in table A1. There are fewer observations in the narrative disclosure samples than in Compustat because reports may not be available on EDGAR, there may be textual idiosyncrasies that lead to problems extracting the 10-K items of interest, and some information is included by reference but stored in other data locations. Data stored in other locations often include exhibits at the end of the 10-K and disclosures contained in non–10-K documents. Footnotes are often incorporated by reference, making their automated extraction very difficult because non-EDGAR presentations tend to have less predictable formatting and headings. The overall effect is a smaller footnote sample size when compared to the business description and MD&A samples.

APPENDIX B

Calculation of Narrative Disclosure Similarity Score

As described in Brown and Tucker [2011], the VSM maps a document into a vector, v, with each vector element, w_i, representing the weighted frequency of a word in that document. The weighted frequency is zero if the word does not occur in that document and the length of the vector is n, the number of unique words in all documents of the sample:

$$v = (w_1, w_2, \ldots, w_n) .$$

TABLE B1
Calculation of Narrative Disclosure Similarity Measures

Variable	Mean	Std Dev	25%	Median	75%	Obs
SIM_{BUS}	0.000	0.078	–0.055	–0.019	0.038	33,293
$SIM_{MD\&A}$	0.000	0.088	–0.059	–0.025	0.036	31,213
SIM_{NOTES}	0.000	0.048	–0.029	–0.013	0.012	22,026
$RAWSIM_{BUS}$	0.108	0.086	0.045	0.087	0.148	35,916
$RAWSIM_{MD\&A}$	0.113	0.095	0.044	0.086	0.155	33,804
$RAWSIM_{NOTES}$	0.056	0.056	0.023	0.040	0.068	25,020
LEN_{BUS}	6,334	3,736	3,601	5,469	8,243	36,137
$LEN_{MD\&A}$	7,044	3,866	3,980	6,426	9,393	34,024
LEN_{NOTES}	8,598	4,121	5,409	8,017	11,208	25,282

SIM = similarity to other clients in the auditor–industry–year reference group, adjusted for *LEN* (subscripts: *BUS* = business description, *MD&A* = management's discussion and analysis, *NOTES* = footnotes to financial statements). Higher values of *SIM* indicate greater similarity to other clients of the same auditor. *RAWSIM* = *SIM* before the length adjustment. *LEN* = number of words in the observation's text.

For example, assume there are only two documents in the sample: (1) "Earnings have increased" and (2) "Earnings have decreased." The length of each document vector is four, since there are four unique words in the sample: w_1 corresponds to "earnings," w_2 to "have," w_3 to "increased," and w_4 to "decreased." The two documents are then represented as:

$$v_1 = (1, 1, 1, 0) \text{ "Earnings have increased"}$$
$$v_2 = (1, 1, 0, 1) \text{ "Earnings have decreased"}$$

The vectors allow for comparisons between documents in the sample (Manning and Schütze [1999]). The cosine of the angle, θ, between any two vectors, v_i and v_j, is a proxy for the similarity of any two underlying documents, $SIM_{DOC,i,j}$:

$$SIM_{DOC,i,j} = \cos(\theta) = \frac{v_i}{\|v_i\|} \cdot \frac{v_j}{\|v_j\|} = \frac{v_i \cdot v_j}{\|v_i\| \, \|v_j\|},$$

where $(\cdot)$ is the vector dot product operator, $\|v_i\|$ is the length of v_i, and $\|v_j\|$ is the length of v_j. SIM_{DOC} ranges from 0 (completely dissimilar documents) to 1 (identical documents). All words are "stemmed" using the Porter stemming algorithm to reduce the dimensionality of the data, which in turn reduces the computing time and resources required (e.g., "earnings," "earned," and "earn" are all converted to "earn").[56] Consistent with Brown and Tucker [2011], the term frequency-inverse document frequency (TF-IDF) algorithm is used to decrease the weight on frequently used words and increase the weight on uncommon words.[57] Therefore, instead of a raw frequency count, each document vector element is the frequency count of the word multiplied by a weight based on the relative prominence of that word in the entire sample.

Because Brown and Tucker [2011] were interested in differences between just two documents at a time, they only calculated pairwise similarity scores. In contrast, in this paper the pairwise scores are aggregated to get a measure of the similarity between one narrative disclosure and the disclosures issued by the client reference group. As with the financial statement similarities, the reference group contains other clients of the same auditor, within the same GICS industry and year. To combine the pairwise $SIM_{DOC,i,j}$ scores between client i and all other clients j in the same auditor–industry–year, the pairwise similarities are averaged to get $SIM_{DOC,i}$ for each observation in the sample.

[56] Even with the reduced dimensions, the calculations take over one week to run on a 2.66-GHz, quad-core machine, while occupying most of the six gigabytes of working memory.

[57] We do not use a "stop word" list to remove extremely common (i.e., unimportant) words, such as "the" and "a," from the sample as in Li [2010]. These words will receive a weight of zero, or very close to it, via the TF-IDF weighting procedure. Brown and Tucker [2011] find no substantial difference in their conclusions between using the TF-IDF approach and a simple frequency count combined with a stop word list. The TF-IDF weights are generated independently for each type of narrative disclosure.

The $SIM_{DOC,i}$ similarity measure is calculated for each observation in the business description ($RAWSIM_{BUS}$), MD&A ($RAWSIM_{MD\&A}$), and footnote ($RAWSIM_{NOTES}$) samples. Brown and Tucker [2011] show that these raw scores are positively related to document length because of the mechanics of the calculation. They control for this relationship by regressing the raw similarity on the first five powers of the number of words in the observation i document (LEN_{BUS}, $LEN_{MD\&A}$, and LEN_{NOTES} in this study). In this study we use the first three powers because the magnitudes of the coefficients rapidly approach zero after this point.[58]

Regressing the raw similarity scores on the first three powers of the document length yields a residual that represents the variation in the raw similarity scores that cannot be explained by these factors. These residuals are labeled SIM_{BUS}, $SIM_{MD\&A}$, and SIM_{NOTES}, producing the similarity scores used in the analysis. Descriptive data for these measure components are shown in table B1.

REFERENCES

AMIR, E. "The Market Valuation of Accounting Information: The Case of Postretirement Benefits Other than Pensions." *The Accounting Review* 68 (1993): 703–24.

ASHBAUGH, H.; R. LAFOND; AND B. W. MAYHEW. "Do Nonaudit Services Compromise Auditor Independence? Further Evidence." *The Accounting Review* 78 (2003): 611–39.

BALSAM, S.; J. KRISHNAN; AND J. S. YANG. "Auditor Industry Specialization and Earnings Quality." *Auditing: A Journal of Practice & Theory* 22 (2003): 71–97.

BAMBER, E. M., AND V. M. IYER. "Auditors' Identification with Their Clients and Its Effect on Auditors' Objectivity." *Auditing: A Journal of Practice & Theory* 26 (2007): 1–24.

BELL, T.; M. CAUSHOLLI; AND W. R. KNECHEL. "Audit Firm Tenure, Non-Audit Services and Internal Assessments of Audit Quality." *Journal of Accounting Research* 53 (2015): 461–509.

BLAY, A., AND M. GEIGER. "Auditor Fees and Auditor Independence: Evidence from Going Concern Reporting Decisions." *Contemporary Accounting Research* 30 (2013): 579–606.

BROWN, S. V., AND J. W. TUCKER. "Large-Sample Evidence on Firms' Year-over-Year MD&A Modifications." *Journal of Accounting Research* 49 (2011): 309–46.

BRUYNSEELS, L.; W. R. KNECHEL; AND M. WILLEKENS. "Auditor Differentiation, Mitigating Management Actions and Audit Reporting Accuracy for Distressed Firms." *Auditing: A Journal of Practice & Theory* 30 (2011): 1–20.

CAHAN, S., AND W. ZHANG. "After Enron: Auditor Conservatism and Ex-Andersen Clients." *The Accounting Review* 81 (2006): 49–82.

CAHAN, S.; J. GODFREY; J. HAMILTON; AND D. JETER. "Auditor Specialization, Auditor Dominance, and Audit Fees: The Role of Investment Opportunities." *The Accounting Review* 83 (2008): 1393–423.

CAIRNEY, T. D., AND G. R. YOUNG. "Homogenous Industries and Auditor Specialization: An Indication of Production Economies." *Auditing: A Journal of Practice & Theory* 25 (2006): 49–67.

CALLAGHAN, J.; M. PARKASH; AND R. SINGHAL. "Going-Concern Audit Opinions and the Provision of Nonaudit Services: Implications for Auditor Independence of Bankrupt Firms." *Auditing: A Journal of Practice & Theory* 28 (2009): 153–69.

---

[58] Hanley and Hoberg [2012] use the VSM to measure the similarity of an IPO prospectus to all the recent IPOs experiencing litigation problems. However, they do not control for document length, making it difficult to ascertain the validity of their measure.

CARSON, E. "Industry Specialization by Global Firm Networks." *The Accounting Review* 84 (2009): 355–82.

CASTERELLA, J. R.; J. R. FRANCIS; B. L. LEWIS; AND P. L. WALKER. "Auditor Industry Specialization, Client Bargaining Power, and Audit Pricing." *Auditing: A Journal of Practice & Theory*: 23 (2004): 123–40.

CAUSHOLLI, M. "Evidence of Organizational Learning and Organizational Forgetting from Financial Statement Audits." *Auditing: A Journal of Practice and Theory* (2015): Forthcoming.

CHANEY, P. K.; D. C. JETER; AND P. E. SHAW. "Client-Auditor Realignment and Restrictions on Auditor Solicitation." *The Accounting Review* 72 (1997): 433–53.

CHANEY, P. K.; D. C. JETER; AND L. SHIVAKUMAR. "Self-Selection of Auditors and Audit Pricing in Private Firms." *The Accounting Review* 79 (2004): 51–72.

CHRISTENSEN, B. E.; T. C. OMER; N. Y. SHARP; AND M. K. SHELLEY. "Pork Bellies and Public Company Audits: Have Audits Once Again Become Just Another Commodity?" Working paper, SSRN, 2014. Available at http://ssrn.com/abstract=2184413 or http://dx.doi.org/10.2139/ssrn.2184413.

CRASWELL, A. T.; J. R. FRANCIS; AND S. L. TAYLOR. "Auditor Brand Name Reputations and Industry Specializations." *Journal of Accounting and Economics* 20 (1995): 297–322.

DE FRANCO, G.; S. P. KOTHARI; AND R. S. VERDI. "The Benefits of Financial Statement Comparability." *Journal of Accounting Research* 49 (2011): 895–931.

DECHOW, P.; W. GE; C. LARSON; AND R. SLOAN. "Predicting Material Accounting Misstatements." *Contemporary Accounting Research* 28 (2011): 17–82.

DECHOW, P.; W. GE; AND C. SCHRAND. "Understanding Earnings Quality: A Review of the Proxies, Their Determinants and Their Consequences." *Journal of Accounting and Economics* 50 (2010): 344–401.

DEFOND, M. L., AND C.S. LENNOX. "The Effect of SOX on Small Auditor Exits and Audit Quality." *Journal of Accounting and Economics* 52 (2011): 21–40.

DEFOND, M. L.; K. RAGHUNANDAN; AND K. R. SUBRAMANYAM. "Do Non-Audit Service Fees Impair Auditor Independence? Evidence from Going Concern Audit Opinions." *Journal of Accounting Research* 40 (2002): 1247–74.

DEFOND, M. L., AND K. R. SUBRAMANYAM. "Auditor Changes and Discretionary Accruals." *Journal of Accounting and Economics* 25 (1998): 35–67.

DEKEYSER, S.; A. GAEREMYNCK; W. R. KNECHEL; AND M. WILLEKENS. "Strategic Competition by Audit Firms." Working paper, KU Leuven, 2015.

DICKINSON, V. "Cash Flow Patterns as a Proxy for Firm Life Cycle." *The Accounting Review* 86 (2011): 1969–94.

DOPUCH, N.; M. GUPTA; D. SIMUNIC; AND M. STEIN. "Production Efficiency and the Pricing of Audit Services." *Contemporary Accounting Research* 20 (2003): 79–115.

DYE, R. A. "Informationally Motivated Auditor Replacement." *Journal of Accounting and Economics* 14 (1991): 347–74.

FELDMAN, R.; S. GOVINDARAJ; J. LIVNAT; AND B. SEGAL. "Management's Tone Change, Post Earnings Announcement Drift and Accruals." *Review of Accounting Studies* 15 (2010): 915–53.

FIOLLEAU, K.; K. HOANG; K. JAMAL; AND S. SUNDER. "How Do Regulatory Reforms to Enhance Auditor Independence Work in Practice?" *Contemporary Accounting Research* 30 (2013): 864–90.

FRANCIS, J.; PHILBRICK, D.; AND SCHIPPER, K. "Shareholder Litigation and Corporate Disclosures." *Journal of Accounting Research* 32 (1994): 137–64.

FRANCIS, J. R.; M. L. PINNUCK; AND O. WATANABE. "Auditor Style and Financial Statement Comparability." *The Accounting Review* 89 (2014): 605–33.

FRANCIS, J. R., AND M. D. YU. "Big 4 Office Size and Audit Quality." *The Accounting Review* 84 (2009): 1521–52.

FRANKEL, R. M.; M. F. JOHNSON; AND K. K. NELSON. "The Relation Between Auditors' Fees for Nonaudit Services and Earnings Management." *The Accounting Review* 7 (2002): 71–105.

FRIED, D., AND A. SCHIFF. "CPA Switches and Associated Market Reactions." *The Accounting Review* 56 (1981): 326–41.

GEIGER, M. A., AND K. RAGHUNANDAN. "Auditor Tenure and Audit Reporting Failures." *Auditing: A Journal of Practice & Theory* 21 (2002): 67–78.

GENERAL ACCOUNTING OFFICE. *Audits of Public Companies: Continued Concentration in Audit Market for Large Public Companies Does Not Call for Immediate Action.* Report 08–163.Washington, D.C.: Government Printing Office, 2008.

GERAKOS, J., AND C. SYVERSON. "Competition in the Audit Market: Policy Implications." Working paper, University of Chicago, 2015.

GOELZER, D. L. *Statement on Auditor's Reporting Model.* Public Company Accounting Oversight Board, Speech by Board Member Daniel L. Goelzer at the PCAOB Board Meeting. 2011. Available at http://pcaobus.org/News/Speech/Pages/03222011_GoelzerStatement.aspx.

GRAMLING, A. A., AND D. N. STONE. "Audit Firm Industry Expertise: A Review and Synthesis of the Archival Literature." *Journal of Accounting Literature* 20 (2001): 1–29.

GUAN, Y.; L.N. SU; D. WU; AND Z. YANG. "Do School Ties Between Auditors and Client Executives Influence Audit Quality?" *Journal of Accounting and Economics* (2016): Forthcoming.

GUL, F. A.; S. Y. K. FUNG; AND B. JAGGI. "Earnings Quality: Some Evidence on the Role of Auditor Tenure and Auditors' Industry Expertise." *Journal of Accounting and Economics* 47 (2009): 265–87.

GUL, F.A.; D. WU; AND Z. YANG. "Do Individual Auditors Affect Audit Quality? Evidence from Archival Data." *The Accounting Review* 88 (2013): 1993–2023.

HAMMERSLEY, J. S. "Pattern Identification and Industry-Specialist Auditors." *The Accounting Review* 81 (2006): 309–66.

HANLEY, K. W., AND G. HOBERG. "Litigation Risk, Strategic Disclosure and the Underpricing of Initial Public Offerings." *Journal of Financial Economics* 103 (2012): 235–54.

HE, X.; J. PITTMAN; O. M. RUI; AND D. WU. "Do Social Ties Between External Auditors and Audit Committee Members Affect Audit Quality?" Working paper, The Chinese University of Hong Kong, 2014.

HENNES, K. M.; A. J. LEONE; AND B. P. MILLER. "The Importance of Distinguishing Errors from Irregularities in Restatement Research: The Case of Restatements and CEO/CFO Turnover." *The Accounting Review* 83 (2008): 1487–519.

JOHNSON, W. B., AND T. LYS. "The Market for Audit Services: Evidence from Voluntary Auditor Changes." *Journal of Accounting and Economics* 12 (1990): 281–308.

JOHNSTONE, K. M.; C. LI; AND S. LUO. "Client-Auditor Supply Chain Relationships, Audit Quality, and Audit Pricing." *Auditing: A Journal of Practice & Theory* 33 (2014): 119–66.

KNECHEL, W. R.; G. V. KRISHNAN; M. PEVZNER; L. B. SHEFCHIK; AND U. K. VELURY. "Audit Quality: Insights from the Academic Literature." *Auditing: A Journal of Practice & Theory* 32 (2013): 385–421.

KNECHEL, W. R.; V. NAIKER; AND G. PACHECO. "Does Auditor Industry Specialization Matter? Evidence from Market Reaction to Auditor Switches." *Auditing: A Journal of Practice & Theory* 26 (2007): 19–45.

KNECHEL, W. R.; L. NIEMI; AND S. SUNDGREN. "Determinants of Auditor Choice: Evidence from a Small Client Market." *International Journal of Auditing* 12 (2008): 65–88.

KNECHEL, W. R.; P. ROUSE; AND C. SCHELLEMAN. "A Modified Audit Production Framework: Evaluating the Relative Efficiency of Audit Engagements." *The Accounting Review* 84 (2009): 1607–38.

KNECHEL, W. R.; A. VANSTRAELEN; AND M. ZERNI. "Does the Identity of Engagement Partners Matter? An Analysis of Audit Partner Reporting Decisions." *Contemporary Accounting Research* 32 (2015): 1443–78.

KRISHNAN, G. V. "Does Big 6 Auditor Industry Expertise Constrain Earnings Management?" *Accounting Horizons* 17 (2003): 1–16.

KRISHNAN, J. "Auditor Switching and Conservatism." *The Accounting Review* 69 (1994): 200–15.

KWON, S. Y.; C. Y. LIM; AND P. TAN. "Legal Systems and Earnings Quality: The Role of Auditor Industry Specialization." *Auditing: A Journal of Practice & Theory* 26 (2007): 25–55.

LANDSMAN, W.; K. NELSON; AND B. ROUNTREE. "Auditor Switches in the Pre- and Post-Enron Eras: Risk or Realignment?" *The Accounting Review* 84 (2009): 531–58.

LENNOX, C. "Audit Quality and Executive Officers' Affiliations with CPA Firms." *Journal of Accounting and Economics* 39 (2005): 201–31.

LENNOX, C., AND C. PARK. "Audit Firm Appointments, Audit Firm Alumni, and Audit Committee Independence." *Contemporary Accounting Research* 24 (2007): 235–58.

LI, F. "Annual Report Readability, Current Earnings, and Earnings Persistence." *Journal of Accounting and Economics* 45 (2008): 221–47.

LI, F. "The Information Content of Forward-Looking Statements in Corporate Filings—A Naïve Bayesian Machine Learning Approach." *Journal of Accounting Research* 48 (2010): 1049–102.

LIM, C. Y., AND H. T. TAN. "Non-Audit Service Fees and Audit Quality: The Impact of Auditor Specialization." *Journal of Accounting Research* 46 (2008): 199–246.

LOW, K. Y. "The Effects of Industry Specialization on Audit Risk Assessments and Audit-Planning Decisions." *The Accounting Review* 79 (2004): 201–19.

MANNING, C. D., AND H. SCHÜTZE. *Foundations of Statistical Natural Language Processing*. Cambridge, MA: MIT Press, 1999.

MENON, K., AND D. D. WILLIAMS. "Former Audit Partners and Abnormal Accruals." *The Accounting Review* 79 (2004): 1095–118.

MINUTTI-MEZA, M. "Does Auditor Industry Specialization Improve Audit Quality?" *Journal of Accounting Research* 51 (2013): 779–817.

MYERS, J. N.; L. A. MYERS; AND T. C. OMER. "Exploring the Term of the Auditor-Client Relationship and the Quality of Earnings: A Case for Mandatory Auditor Rotation?" *The Accounting Review* 78 (2003): 779–99.

NICHOLS, D. R., AND D. B. SMITH. "Auditor Credibility and Auditor Changes." *Journal of Accounting Research* 21 (1983): 534–44.

NUMAN, W., AND M. WILLEKENS. "An Empirical Test of Spatial Competition in the Audit Market." *Journal of Accounting and Economics* 53 (2012): 450–65.

OWHOSO, V. E.; W. F. MESSIER JR.; AND J. G. LYNCH JR. "Error Detection by Industry-Specialized Teams During Sequential Audit Review." *Journal of Accounting Research* 40 (2002): 883–900.

PUBLIC COMPANY ACCOUNTING OVERSIGHT BOARD (PCAOB). *Concept Release on Auditor Independence and Audit Firm Rotation*. PCAOB Release No. 2011-006. Washington, DC: PCAOB, 2011.

REICHELT, K. J., AND D. WANG. "National and Office-Specific Measures of Auditor Industry Expertise and Effects on Audit Quality." *Journal of Accounting Research* 48 (2010): 647–86.

RIEDL, E., AND S. SRINIVASAN. "Signaling Firm Performance Through Financial Statement Presentation: An Analysis Using Special Items." *Contemporary Accounting Research* 27 (2010): 289–332.

SALTON, G.; A. WONG; AND C. S. YANG. "A Vector Space Model for Automatic Indexing." *Communications of the ACM* 18 (1975): 613–20.

SECURITIES AND EXCHANGE COMMISSION (SEC). *Concept Release on Management's Discussion and Analysis of Financial Condition and Operations*. Release No. 33–6711. Washington, DC: SEC, 1987.

SECURITIES AND EXCHANGE COMMISSION (SEC). *Management's Discussion and Analysis of Financial Condition and Results of Operations; Certain Investment Company Disclosures*. Release No. 33–6835. Washington, DC: SEC, 1989.

SECURITIES AND EXCHANGE COMMISSION (SEC). *Commission Guidance Regarding Management's Discussion and Analysis of Financial Condition and Results of Operations*. Release No. 33–8350. Washington, DC: SEC, 2003.

SHEVLIN, T. "The Valuation of R&D Firms with R&D Limited Partnerships." *The Accounting Review* 66 (1991): 1–21.

SHU, S. Z. "Auditor Resignations: Clientele Effects and Legal Liability." *Journal of Accounting and Economics* 29 (2000): 173–205.

SINGHAL, A. "Modern Information Retrieval: A Brief Overview." *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering* 24 (2001): 35–43.

SUN, Y. "Do MD&A Disclosures Help Users Interpret Disproportionate Inventory Increases?" *The Accounting Review* 85 (2010): 1411–40.

U.K. HOUSE OF LORDS. "Auditors: Market Concentration and Their Role," 2011. Available at http://www.publications.parliament.uk/pa/ld201011/ldselect/ldeconaf/119/119.pdf.

WAHLEN, J. M. "The Nature of Information in Commercial Bank Loan Loss Disclosures." *The Accounting Review* 69 (1994): 455–78.

WILLENBORG, M. "Discussion of 'Brand Name Audit Pricing, Industry Specialization, and Leadership Premiums Post-Big 8 and Big 6 Mergers'." *Contemporary Accounting Research* 19 (2002): 111–16.

ZMIJEWSKI, M. E. (1984) "Methodological Issues Related to the Estimation of Financial Distress Prediction Models." *Journal of Accounting Research*, 22, 59–82.