McGladrey

Distinguished Lecture Series
School of Accountancy
W. P. Carey School of Business
Arizona State University

Stephen Brown
of
Fisher School of Accounting
University of Florida
will present

"Specialization Through Client Commonality and Its Effect on Audit Production Cost"

on

January 27, 2012

1:30pm in BA 258

Specialization Through Client Commonality and Its Effect on Audit Production Costs

Stephen V. Brown
Fisher School of Accounting
University of Florida
stephen.brown@ufl.edu

January 2012

I am especially grateful to my dissertation committee for their assistance on this paper: W. Robert Knechel (chair), Stephen Asare, Gary McGill, Praveen Pathak, and Jennifer Wu Tucker. I would also like to thank Will Ciconte, Julie Coffey, Johannes Impink, Carlos Jiménez, Lan Su, Ying Zhou, and the workshop participants at the University of Florida for their insightful comments.

Abstract

Similarities among clients create commonalities across audit processes that provide an auditor with an opportunity to reduce production costs through increased knowledge spillover, changes in the labor mix, and the implementation of additional audit technology. Auditors that take advantage of the similarities across clients can specialize in a particular set of companies, potentially allowing increased audit efficiency and lower audit risk. I introduce client-level measures of company similarity that are based on both financial statements and narrative disclosures from the annual report. I use these measures to proxy for the opportunities to specialize that arise from greater client commonality, finding strong evidence that higher client overlap is associated with lower audit fees. This relationship is incrementally stronger in industries for which the auditor has greater economic incentives. Because the financial statements and narrative disclosures are distinct disclosure channels, I explore the effect on audit fees when the two channels portray inconsistent messages about the degree of client commonality. When the financial statements are relatively unusual compared to peer clients, but the narrative disclosures do not reflect this dissimilarity, I expect the auditor to assess higher audit risk. Consistent with this prediction, the auditor charges higher audit fees under this condition. On the other hand, when the narrative disclosures are more unique than the financial statements reflect, audit fees are lower, which I argue is due to the greater, more useful firm-specific information contained in the text relative to the financial statements.

1. Introduction

Each engagement within an auditor's portfolio has both idiosyncratic and non-idiosyncratic features deriving from the extent to which the audits have elements in common. A portion of each accounting disclosure of a client is due to the economic and accounting choices of that company, while other portions are the result of common factors such as auditor preferences, industry norms, macroeconomic conditions, and accounting standards. In this paper, I argue the non-idiosyncratic, overlapping components represent opportunities for the auditor to reduce production costs by improving audit technology and reliance on common knowledge spillover, which I refer to collectively as specialization. I use the similarity of each client to other clients within the same auditor-industry-year as a proxy for potential opportunities to specialize in that group of companies.

I develop client-year measures of similarity for both financial statement and narrative textual disclosures since they represent different, but complimentary, signals. While the accounting systems and underlying economics are not separately observable, their joint effect is reflected in the financial statements. The narrative disclosures also provide economic and accounting system information, but contain additional information about management's interpretation of past events and expectations about the future. The narrative disclosures are especially flexible, giving the opportunity for management to communicate more firm-specific information or otherwise influence the market's view of the company. These two signals are each necessary to better understand the other and, taken together, provide a more comprehensive view of how the clients of an auditor relate to one another.

The financial statement similarity proxy relies on the Mahalanobis distance, used extensively in the cluster analysis literature to divide objects into groups based on sets of numbers associated with each object.[1] The set of numbers I use is motivated by financial components known to be important in an audit context, including proxies for effort, complexity, and risk. The narrative disclosure similarity measure extends the pairwise similarity score introduced in Brown and Tucker (2011) as a proxy for year-over-year changes in MD&A. As the source of narrative disclosures, I use the business description, MD&A, and footnotes contained in the mandatory annual report. Using each measure, I calculate the similarity of each client-year to other clients in the same auditor-industry-year (the "reference group").

In my first hypothesis, I argue the degree of commonality among clients of an auditor can affect audit production costs through its effect on both labor and audit technology. Commonality influences labor costs through knowledge spillovers between engagements and changes to the mix of more senior and less experienced labor. Production costs are also a function of audit technology, which is easier to implement and more effective when client overlap is greater. Given the opportunity for reduced production costs, I first predict that a client having more in common with its peer clients has lower audit fees. I find strong evidence of this association for both financial statement and narrative disclosure similarity. The effect is also economically significant: an increase in similarity from the 25th to the 75th percentile is associated with a decrease in audit fees of 4.3 to 8.3 percent.

[1] I typically use the term "similarity" in this paper, although the context occasionally calls for the term "distance." In the current context, distance is the conceptual inverse of similarity.

My second hypothesis is that the relationship between client similarity and fees is stronger when the auditor has greater financial incentives to take advantage of overlap in its portfolio. The auditor is unlikely to make the necessary investments solely because of the opportunity to do so, but will also consider how economically meaningful the investment might be for overall profitability. In support of this hypothesis, I document an incrementally negative effect for financial statement similarity when an industry provides a higher percentage of the auditor's revenues.

Finally, having two primary measures for client commonality allows me to examine situations in which the two proxies are inconsistent in their portrayal of similarity. I consider two types of inconsistency between the financial statements and corresponding narrative disclosures: pooled textual disclosures and differentiated textual disclosures. A *pooled* disclosure occurs when a company has unusual-looking financial statements relative to its peers, but the accompanying textual disclosures do not reflect those financial differences. The accompanying text should either reflect the atypical financial statements or explain why the differences are not a true representation of the company's situation. However, the text does not appear to do so, representing an incremental risk factor for the auditor and possibly eroding the production efficiencies predicted to be associated with greater client commonality. I proxy for each type of inconsistency by focusing on firms that are in opposing terciles of similarity for financial statements and narrative disclosures. As predicted, I find that pooled text disclosures are associated with higher audit fees than clients without such inconsistency.

The second type of inconsistency—a *differentiated* disclosure—occurs when a company has fairly typical financial statements relative to its peers, but the textual disclosures seem to contain more uncommon, possibly firm-specific, information. The prediction in this case is less clear than a pooled disclosure since differentiation can be the result of a client who is (1) unjustifiably trying to differentiate itself from its peers or (2) attempting to provide additional, firm-specific information that can be useful and risk-reducing to both auditors and investors. In contrast to pooled disclosures, I generally find that differentiated disclosures are associated with lower audit fees than clients not having this type of inconsistency.

My study makes several contributions to the literature. First, I provide empirical proxies of auditor specialization that have several advantages over existing measures. The proxies are at the client level, rather than the auditor-industry level, which allows a more direct mapping into client-level audit fees. This approach also allows for the existence of subgroups within an auditor-industry, since auditors do not necessarily orient their practices around the broad groups provided by third party industry classification systems. Because my proxies rely explicitly on client characteristics, I avoid the use of market share measures that are likely confounded by competitive pricing strategies and other audit market features, making my measures easier to interpret as proxies for the underlying specialization construct. Although interpreted in an audit context for the current study, the proxies are general purpose measures of overlap among companies, providing many potential applications outside of the audit setting.

My second contribution is to the limited literature on the relation between auditors and clients' narrative disclosures. Few audit-related studies consider the role of narrative disclosures in conveying information about the client,[2] even as the PCAOB has recently proposed substantially increasing the role of the auditor in reviewing these communications (PCAOB, 2011). In this study, I show the usefulness of narrative disclosures in examining the implications of how clients of an auditor relate to one another. In a third contribution, an extensive literature has looked at the relationship between specific client financial statement elements and the audit, without a higher-level understanding of what the broader financial data mean for the auditor's client portfolio. The measures I develop allow for a research design that simultaneously considers multiple dimensions of client commonality. I further combine the multiple disclosure channels used by the client to look for inconsistencies, which provides more nuanced insights than those provided by studies examining only one disclosure mechanism.

The rest of the paper proceeds as follows. Section 2 develops the hypotheses. Section 3 describes the rationale and foundation for the similarity measures, while Section 4 explains the sample and calculations of the measures. Section 5 describes the results of the empirical tests. Section 6 contains alternative similarity measures and other robustness tests, with a conclusion in Section 7.

2. Hypotheses and Prior Literature

2.1 Production Costs and Audit Fees

Simunic (1980) presents a widely-used model of audits in which fees charged to clients are a function of production costs ("effort") and any expected losses due to potential audit failure ("risk"). Production costs—primarily labor in an audit setting—are composed of the

[2] One exception is Dunn and Mayhew (2004), which finds that clients of industry specialists have higher quality narrative disclosures.

quantity and unit cost of resources consumed to provide a given level of audit quality. For example, the size of the client corresponds to a higher quantity of resources required; as such, client assets and sales are positively related to the quantity of labor hours expended (O'Keefe et al. 1994). Since there is a non-zero probability that an audit will fail by not detecting or reporting a material financial statement error, the auditor must either charge a higher fee to insure against the possible loss or expend greater effort to reduce the risk. For instance, Hackenbrack and Knechel (1997) show that the labor mix shifts towards more senior, costly auditor employees when audit risks are higher. Overall, prior literature has documented a very strong, positive relation between audit production costs (both effort and risk) and audit fees (Causholli, De Martinis, et al. 2010).[3]

2.2 Specialization and Audit Fees

An extensive literature has examined the effect an auditor's specialization in a group of clients has on audit fees. The "group" is typically implemented as some category of industry, leading to the customary term *industry specialization*. Studies in this area variously predict both a decrease and an increase in audit fees due to industry specialization, although the archival evidence has generally supported the latter.

Industry specialists are expected to charge *lower* fees when non-idiosyncratic audit components lead to knowledge sharing and investments in overlapping audit technology that are associated with lower production costs due to having a more efficient and less risky audit. Earlier studies have occasionally acknowledged the possibility of this negative relationship

[3] Lower production costs do not necessarily lead to lower audit fees if the auditor is retaining the entire increase in profit margin. However, as long as the audit market is sufficiently competitive, at least some portion of these lower costs will be passed along to the client.

(e.g., Craswell et al. 1995; Willenborg 2002) and some archival results support this prediction. For example, Mayhew and Wilkins (2003) find that auditors who have larger industry market share, but do not dominate the industry, charge lower fees to clients initially going public. Experimental evidence also lends credence to the potential for lower fees (e.g., Owhoso et al. 2002; Low 2004).

However, most studies in the area proxy for specialization using industry market share and typically find *higher* fees for specialists (Gramling and Stone 2001). The general interpretation is that the same knowledge sharing and audit technology described under the negative prediction improve audit quality or auditor reputation (e.g., Ward et al. 1994). Since clients are presumably willing to pay more for higher actual—or perceived—quality, specialization should be associated with higher audit fees.

Given the two divergent predictions, the choice of proxy for specialization is especially critical. For example, the offsetting effects could lead to no discernable relationship (e.g., Palmrose 1986). On the other hand, if the proxy better captures the quality and reputational effects of specialization, a positive relation will dominate, as appears to be the case when using industry market share. Market share-based proxies could also be measuring the competitive strategy of an auditor in the audit market for a particular industry rather than specialization per se (Numan and Willekens 2011). Minutti-Meza (2010) argues that studies documenting a positive relation between industry specialists and audit quality are the result of uncontrolled client characteristics, and finds no improvement in audit quality for specialist clients once fully matching on these attributes. Gramling and Stone (2001) note the link between market share and specialization is typically vague and that "existing research offers little justification for applying existing market share and market specialization measures as proxies for industry expertise" (p. 14). In the current study, I develop measures that more directly proxy for having a production process specialized for a subset of clients so that I can better address the negative relation between specialization and fees.

2.3 Opportunities to Lower Production Costs

Commonality and idiosyncrasies among clients of an auditor can affect audit production costs through their effect on both labor and audit technology. One effect on labor costs is that fewer idiosyncrasies will likely require less planning and oversight due to decreased risk and complexity, thus shifting the labor mix to lower-level, less expensive personnel (Hackenbrack and Knechel 1997). There is also the potential for knowledge overlap, which includes familiarity with certain "types" of clients, rules-of-thumb, and other relevant on-the-job experience (e.g., Beck and Wu 2006). Research in organizational behavior has found that knowledge gained by performing job tasks is transferred within an organization (e.g., Darr et al. 1995). Experimental evidence suggests that specialist auditors are better at detecting errors (Owhoso et al. 2002) and assessing audit risk (Low 2004). However, archival auditing studies have not found strong empirical evidence to support learning-by-doing or learning over time (Causholli, De Martinis, et al. 2010; Davis et al. 1993; O'Keefe, King, et al. 1994), possibly due to the specific proxies chosen.

Increased client overlap could also affect the auditor's ability to develop specialized audit technology. Audit technology is a set of fixed investments by an auditor in innovations such as customized workflow, employee training, specialized software, decision aids, and the formation of in-house consulting groups (Dowling 2009; Sirois and Simunic 2010).[4] Higher-quality audit technology is "better at identifying and directing effort to problem areas of individual clients" (Blokdijk et al. 2006, p. 29). A higher degree of client commonality could provide more input into the current audit. For example, analytical procedures have better predictive ability when based on similar peer firms (Minutti-Meza 2012). These techniques are likely to be more accurate when based on a larger number of more similar reference clients. Cahan et al (2008) argue that homogenous investment opportunity sets among clients are a specific type of client overlap that creates such an opportunity to invest in audit technology. I extend this line of reasoning to examine client overlap in a more general sense. If there is greater client overlap, there will be more common audit components to extract, and thus it will be less costly and more effective to develop common technologies based on those similarities. [5]

The lack of archival evidence notwithstanding, organizational theory and experimental studies suggest a greater ability to transfer knowledge within the audit firm will lead to lower audit risk and more efficient audits. Both of these outcomes will result in an audit with lower production costs, albeit with potentially higher audit quality. Therefore, I predict in alternative form:

> **H1: Clients having higher overlap with other clients of the auditor pay lower audit fees.**

---

[4] Note that audit technology is not necessarily implemented using computerized systems.

[5] While some technology and knowledge can be broadly applied, such as audit standards and firm-wide policies, I specifically focus on components that are relevant to subgroups of clients to provide adequate cross-sectional variation.

2.4 Incentives to Lower Production Costs

The first hypothesis derives from the opportunities inherent in client overlap, but audit firms and individuals will only invest in additional audit technology and develop common knowledge when there are incentives to do so. Economic incentives are likely to be highest for those clients that are relatively more important to the auditor's overall profitability. For example, an industry that provides audit fees that are higher than other industries might give the auditor greater incentives to develop audit technology appropriate for that industry. In contrast, if an industry represents a very small portion of the fee portfolio, the auditor is less likely to make investments in technological improvements for that group of clients, even in the presence of strong opportunities. I expect greater incentives to develop specialized audit technology and knowledge will accentuate the relation between opportunities and fees predicted in the first hypothesis. Supporting the significance of stronger portfolio incentives, Knechel, Niemi, and Zerni (2011) find that partner specialization is associated with higher compensation for economically important sectors. Therefore, the next alternative hypothesis is:

> **H2: The negative relation between client overlap and audit fees is stronger in industries that are economically important to the auditor.**

2.5 Inconsistent Signals of Commonality

Given multiple consistent signals of the true underlying client overlap, the prior hypotheses make predictions about the relationship between commonality and audit fees. For the purposes of this study, I use the financial statements and the narrative disclosures in the annual report as two broad disclosure channels. Bamber and Cheon (1998) show cross-sectional variation in management's choice of channels for disclosing earnings forecasts, along with differential investor reaction to those choices. Therefore, an incremental effect beyond the earlier predictions can arise if these disclosure channels are not in agreement with one another regarding the degree of underlying similarity.

One type of disclosure inconsistency occurs when the quantitative financial statements seem to represent a company that is relatively unusual for the industry, but the accompanying qualitative narrative disclosures make the client appear very typical. If the financials are dissimilar, one would expect that the accompanying text would either reflect these differences or explain why the differences are not a true representation of management's view of the company's position. In either case, the narrative disclosures should appear different from other clients of the auditor. Narrative disclosures give the company greater flexibility and discretion than is usually available in the financial statements. Under this flexible regime, the company is apparently choosing to downplay the differences in the underlying financials. I call this situation *pooled text* inconsistency. Inappropriately differentiated disclosures could cause additional risk for the auditor or require more effort to attain the same level of assurance. But even if the differences are justified, verifying the propriety of the claims will take additional effort by the auditor:

> **H3a: Clients with dissimilar financial statements but similar narrative disclosures ("pooled text") pay higher fees than other clients.**

A second type of inconsistency is when the financial statements indicate a client is relatively similar to other companies, but the narrative disclosures make the client appear more unusual. The client may be attempting to unjustifiably differentiate itself from other companies, as might occur before an upcoming equity offering. On the other hand, narrative differences could represent firm-specific disclosures that improve the quality of information available about the company. For example, Tasker (1998) shows that managers will use a more flexible disclosure channel when the financial statements are relatively less informative. This improvement in the information environment represents a potentially positive situation for the auditor. I call this type of inconsistency the *differentiated text* condition. Because there are both beneficial and problematic potential reasons for differentiated narrative disclosures, it is an empirical question as to the relation between this type of inconsistency and audit fees. Stated in alternative form, my final hypothesis is:

> **H3b: Clients with highly similar financial statements but dissimilar narrative disclosures ("differentiated text") pay different fees than other clients.**

3. Measurement of Similarity

There are at least two streams of academic research that study inter-entity similarity, the two being differentiated by the nature of the underlying data. The first stream is cluster analysis, along with related fields such as factor analysis, and is primarily concerned with grouping similar entities together based on a vector of numeric data. The second stream is the information retrieval literature, which uses documents as the underlying data. The documents must still be converted to a numerical representation before calculating similarities, but there are unique challenges in analyzing the similarity of documents that are not faced when using short vectors of fundamentally numeric data.

3.1 Financial Statement Similarity

Algorithms used in cluster analysis of numeric data include Euclidean distance, city-block distance, Chebychev distance, and Mahalanobis distance (Hair et al. 2006). Of these, the Mahalanobis distance-squared (D^2) is particularly sophisticated in its ability to weight each variable equally according to its individual scale, as well as account for covariances among the various components. Introduced in Mahalanobis (1936), the D^2 statistic imposes few restrictions on the underlying variables, only requiring non-degenerate distributions.[6] After scaling and accounting for covariance of the variables, an observation's distance from the group is larger when the variables for the observation are jointly more "unusual." The D^2 measure is the generally preferred algorithm in cluster analysis, when available (Hair et al. 2006).

Outside of the accounting literature, D^2 has been used in management to compare the distance between countries along multiple dimensions, including economic, financial, political, cultural, demographic, and geographic location (Berry et al. 2010). Climatologists have used the measure to look for boundaries between different regional climates (Mimmack et al. 2001). And chemists have used it for multivariate calibration, pattern recognition and process control (De Maesschalck et al. 2000).

Prior accounting literature has used the D^2 measure to a limited extent. Rege (1984) uses it to test the effectiveness of a discriminant function in classifying data into two groups of likely and unlikely takeover targets. If the distance between the two groups is significant then the discriminant function is considered effective. Iyer (1998) also employs the statistic in a discriminant function context to maximize the distance between subgroups. Guilding and McManus (2002) use the measure to test for potentially influential outlying observations. However, all these studies use the measure in a statistical context and do not examine the properties of the distance itself.

[6] The distance is generally left in squared form because D^2 has a known distribution when comparing one observation to a group and when comparing one group to another.

The abnormal accrual model used extensively in prior accounting research could also be conceptualized as a distance measure. It empirically models the relationship between total accruals and various explanatory variables to attempt detection of accrual levels that are unusual-looking relative to peer companies. However, due to the D^2 measure's power and flexibility, I use it as my primary measure of financial statement similarity.

3.2 Narrative Disclosure Similarity

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). Assuming the ability to map a document into a numeric representation, the D^2 measure is also conceptually possible in a document context. However, practical considerations limit its usefulness. When mapping a list of words from narrative text into variable vectors, the high dimensionality makes calculation of the D^2 measure impractical. For example, a covariance matrix using the 98,519 unique words for the MD&A in my sample would contain 9.7 billion elements. The matrix would then need to be inverted.[7]

Given the computational challenges, I instead use the Vector Space Model (VSM) from the document retrieval literature that can better accommodate large sets of long text. The VSM maps a document into a numeric vector (Salton et al. 1975). As in cluster analysis, there are a variety of comparison algorithms available, including the Dice coefficient, Jaccard coefficient, overlap coefficient, Euclidian distance, and cosine (Manning and Schütze 1999). There are

---

[7] The large size of the matrix arises not primarily from the number of documents, but from the unique words used in those documents. While the two are positively correlated, calculating the D^2 measure on a subset of documents would not generally address the computational difficulties.

known shortcomings with some of these measures, such as the Euclidian distance (Sohn 2001), while the cosine measure is the most commonly implemented (Singhal 2001).

Even though the VSM lacks some of the statistical sophistication of the D^2 measure, it has the advantage of being an inherently pairwise operation. The Mahalanobis distance is the distance of an object from the group, taken as a whole, so it is difficult to examine the relation between one object and other individual objects in the group. On the other hand, the VSM can easily be used to find the difference of an object from other specific objects in the group. This ability could be useful in a financial reporting context where similarity to a group of close competitors is potentially more relevant than similarity to the entire industry.

Within the accounting literature, Brown and Tucker (2011) have used the VSM cosine statistic to measure year-over-year dissimilarities in MD&A as a proxy for changes in narrative disclosure. Because they are interested in the differences between just two documents at a time, they only calculate pairwise similarity scores. In contrast, I aggregate these pairwise scores to get a measure of the similarity between one narrative disclosure and the disclosures issued by a reference group of clients within the same auditor-industry-year. Due to its relative ubiquity in information retrieval and its use in Brown and Tucker (2011), my primary measure of document similarity will be the VSM-based cosine similarity score.[8]

---

[8] While I cannot feasibly calculate the D^2 statistic for documents, I *can* calculate the VSM similarity for financial statements. Using such an approach, Jaffe (1986) uses vectors of different categories of patent applications to examine R&D spending overlap within industries. In the current study, only the H1 results hold when using the VSM for financial statements, probably because the VSM does not account for variances and covariances of the variable components, which reduces its statistical power.

4. Sample

4.1 Financial Statements

I gather the necessary financial statement variables from Compustat, only using observations with assets greater than $1 million, with no fiscal year end change, not in the financial or utility sectors, and having all data fields required to calculate the similarity scores. The sample begins in 2000, when audit fee data is first widely available, and ends in 2009.

The concept of similarity is with respect to some group of other objects and is undefined for a single observation on its own. I call these other observations the *reference group*, which I define as other clients in the same auditor-industry-year. I exclude any reference groups that do not have at least five observations; the similarity score is unlikely to be reliable if there are too few observations in the group. Because the reference groups are rarely large enough for non-Big4 auditors, I explicitly limit the sample to Big 4 clients. Finally, I do not allow companies in the reference group in the year that they switch auditors. These restrictions leave 32,412 observations in my financial statement sample.[9]

Because there is no theoretical guidance on which variables are appropriate for the financial statement similarity measure, I use financial statement variables having well-established relationships in an audit context. Based on the empirical audit fee model components described in Hay et al (2006), I include proxies for audit effort, audit complexity, and client risk. To focus on the client's financial statement similarity, I avoid engagement- or

[9] Hogan and Jeter (1999) document the increasing importance of auditor restructuring along industry lines at the national level to take better advantage of internal teams of experts. While client overlap can also occur at the office level within an auditor, there are few offices with enough clients to calculate my similarity measures within an industry. Calculating the scores at the office-industry level would lead to a 63% reduction in sample size (and a 30% reduction if calculated at the office-sector level).

auditor-specific variables and client-related variables that are not included in the financial statements. As a distance-based measure, using unscaled variables would cause the D^2 metric to be so heavily influenced by the size of the companies that it would effectively become a proxy for client size. Because client size typically explains a large portion of audit fees and large firms are more uncommon by definition than smaller firms, I do not directly include proxies for size and also scale all variables to remove a direct size effect. Correlations with size are normally observed in financial data (e.g., between size and profitability), which ensure client size has an indirect effect on the measure without overwhelming other patterns in the data.

I count the number of non-missing/non-zero financial statement variables in Compustat as a measure of audit effort and complexity (*CSITEMS*), since additional financial statement items are likely to increase the scope and intricacy of the audit.[10] I use long-term debt to proxy for the risk due to the client's leverage (*LEV*). The combination of inventory and receivables proxies for inherent audit risk (*IRISK*). Audit fee models usually include a measure of profitability, frequently some variant of income or a profit/loss dummy. Departing somewhat from prior literature, I include separate variables for revenues (*REV*) and operating expenses (*EXP*) to give the income statement roughly the same representation in the vector as the balance sheet.[11] All measures are scaled by total assets. I regress the natural log of audit fees on these scaled variables to verify they are all highly significant in the expected directions and consistent with prior literature.

---

[10] The number of reporting segments is frequently used to proxy for audit complexity, but is unavailable for many companies in Compustat. Counting the number of variables serves as a broadly available alternative. To my knowledge, this variable has not been used before in the audit fee literature, but is potentially superior to existing alternatives.

[11] In an untabulated robustness test, I use income before extraordinary items (*INC*) in place of *REV* and *EXP*, with no change in the qualitative conclusions.

If the $n = 5$ financial statement variables, x_i, are contained in vector, $x^T = (x_1, x_2, \ldots, x_n)$, the observation is contained in a group with mean values, $\mu^T = (\mu_1, \mu_2, \ldots, \mu_n)$, and the group has covariance matrix, Σ, then the Mahalanobis distance-squared is:[12]

$$D^2 = (x - \mu)^T \Sigma^{-1} (x - \mu)$$

Because D^2 is a measure of distance, I convert it to a financial statement similarity measure, SIM_{FS}, by taking the inverse, and then the natural log to reduce skewness and outliers:

$$SIM_{FS} = ln\left(\frac{1}{D^2}\right)$$

I calculate SIM_{FS} for each company-year in my sample, using other companies in the same auditor-industry-year as the reference group. The scores are calculated within a GICS industry due to the general lack of comparability across industries. I generate the $(x - \mu)$ portion of the measure by subtracting the auditor-industry-year mean for each variable from the variables for the company-year being analyzed. I compute the Σ covariance matrix at the industry-year level to account for different scales and covariances across industries and over time that are unlikely to vary significantly between auditors.[13]

4.2 Narrative Disclosures

I require the narrative disclosures in my sample to have limited self-selection bias. Therefore, any potential disclosures should be mandatory for a wide cross-section of firms; the

---

[12] The Euclidian distance between observations is a special case of the Mahalanobis distance. If the covariance matrix (Σ) is the identity matrix, the square root of D^2 simplifies to the familiar $[(x-\mu)^T(x-\mu)]^{1/2}$, which is the Pythagorean theorem if the vector has length two.

[13] Because the technique calculates distance, a relatively high degree of multicollinearity does not cause a problem as it would in a typical regression. However, as the variables approach nearly perfect multicollinearity, the covariance matrix will not be invertible, which can be a concern when using a small vector of variables in a small industry. While unusual in my sample, I exclude industry-years that do not have at least ten company-year observations (twice the number of variables). This restriction is almost always met given the earlier restriction of at least five observations with an *auditor*-industry-year.

annual report provides such a set of disclosures. Within the 10-K, the longest sections tend to be the business description, Management's Discussion & Analysis (MD&A), and the financial statement footnotes. Excluding exhibits, there are an average of 8,394 words in the footnotes, 6,803 in the MD&A, and 6,309 in the business description, comprising 23, 21, and 21 percent, respectively, of the length of the typical 10-K.

The footnotes and MD&A seem 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 (e.g., SEC 1987; 1989; 2003). Prior studies have demonstrated the usefulness of footnotes, at least to some extent (e.g., Shevlin 1991; Amir 1993; Wahlen 1994; Riedl and Srinivasan 2010). Other research has shown some of the potential information contained in MD&A (e.g., Feldman et al. 2010; Li 2010; Sun 2010). Of the three narrative disclosures, the business description is relatively unexplored except in studies of the full 10-K as a single document (e.g., Li 2008).

I use multiple types of narrative disclosures because there is considerable variation in the characteristics of each 10-K item. First, the topics discussed in each item are different. Item 101 of Regulation S-K requires Item 1 contain a detailed narrative description of the business, including industrial and geographic segments, principal products and services, R&D spending, and competitive conditions. 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. Even though there is some topical overlap, the footnote content is typically determined by GAAP. Second, the MD&A is intended to be an interpretation of past and future operations "through the eyes of management" (SEC 2003). Given certain conditions, any forward-looking statements receive Safe Harbor protection (Item 303(c) of Regulation S-K). In contrast, regulations do not broadly require footnotes to explicitly contain interpretive or forward-looking statements. Third, the footnotes are audited, while the business description and MD&A are not audited but are only reviewed for material misstatements and consistency with facts known to the auditor (AU Sections 550; 551). While all narrative disclosures are somewhat flexible, the lack of an explicit audit of the business description and MD&A gives management the greatest flexibility to choose the topics and quality of discussion. Because all three have distinct characteristics, I use each as a separate source of narrative disclosure.

For the narrative disclosure sample, I use 10-K's and 10-K405's filed electronically via the SEC's EDGAR system for fiscal years 2000 through 2009. As in the financial statement sample, the disclosures in the text samples are by Big4 clients 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 23,146 business description, 22,146 MD&A, and 10,666 footnote observations. There are fewer observations in the narrative disclosure samples than in the financial statement sample. This difference is primarily due to unavailable reports on EDGAR, items included by reference to other locations, and textual idiosyncrasies that lead to problems extracting the 10-K items of interest. The substantial drop in the number of footnote observations, as compared to the business description and MD&A samples, is because many companies attach financial statements and footnotes as an exhibit to the report in a variety of unpredictable ways, making their automated extraction difficult.

Treating the three narrative disclosure items of the annual report as separate data sets, I calculate the similarity score for each using an extension of the approach in Brown and Tucker (2011) that allows for a comparison of firm *i* to its peers. The process, summarized in Appendix B, produces three variables—SIM_{BUS}, $SIM_{MD\&A}$, and SIM_{NOTES}—that proxy for the amount of commonality between firm *i* and its five closest peers in the same auditor-industry-year. Higher similarity scores correspond to greater commonality.

5. Analysis of Similarity

5.1 Similarity Patterns

Panel A of Table 1 displays descriptive statistics for the financial statement and narrative disclosure similarity measures. SIM_{FS} is slightly skewed with a mean of -4.32 (median = -3.43). The sample contains the highest number of observations for this similarity measure (n = 32,412), because the financial statement components used to construct SIM_{FS} are available for nearly all company-years in Compustat.

For the narrative disclosures, the raw similarity scores are 0.19 for the business description, 0.19 for the MD&A, and 0.08 for the footnotes.[14] The footnotes tend to be the longest disclosures with 9,559 words, followed by 8,103 words in the MD&A, and 6,561 words in the business description. The average auditor-industry-year reference group size is 38 clients for financial statements, 22 for the business description and MD&A, and 12 for the footnotes (untabulated). After normalizing the raw scores using the observation document length, the average reference group document length, and the reference group size, the narrative similarity

[14] The scores are not directly comparable across disclosure types since they have different document lengths and are weighted using different sets of word frequencies (using the TF-IDF algorithm described in Appendix B).

scores (SIM_{BUS}, $SIM_{MD\&A}$, and SIM_{NOTES}) are centered around zero. Deviance from the mean of zero corresponds to the unexpected similarity conditional on document length and reference group characteristics.

Table 1, Panel B shows the Pearson pairwise correlations among the similarity measures. I limit the correlations to those observations with scores available for all four disclosures (narrative and financial), although the unrestricted correlations are similar. The four similarity scores are all positively correlated with one another, indicating they measure related constructs. The correlations are higher among the three narrative disclosures (0.59 to 0.73) than they are with the financial statements (0.08 to 0.16), possibly due to disparate approaches in measuring the two types of disclosure.

As a means of validation, Table 1, Panel B also shows the Pearson pairwise correlations between the similarity measures and various proxies for client differences, which I expect to be negatively correlated. For client *SIZE* and two of the components used to produce SIM_{FS} (*IRISK* and *LEV*), I calculate the client's absolute difference from the mean for the auditor-client-year, calling them |*SIZEDIFF*|, |*IRISKDIFF*|, and |*LEVDIFF*|.[15] The significant correlations with SIM_{FS} are all negative, even though they are not particularly large (the highest is -0.08). Negative correlations are also present for the narrative disclosure scores, which did not explicitly include any financial statement variables. Of all the difference variables, |*SIZEDIFF*| has the highest negative correlation with SIM_{FS}, so client size appears to be an important determinant of financial statement similarity, even though it was not directly included in the

[15] These difference variables should be negatively correlated with SIM_{FS} by design, but these correlations document the SIM_{FS} measure is working as expected. Performing a series of univariate correlation tests is also substantially different from the joint difference measure produced using the D^2 technique.

measure calculation. On the other hand, $|IRISKDIFF|$ has the most negative correlation with the narrative similarities, which is consistent with the importance of the numerous risk-related disclosures in annual report items accompanying the financial statements (e.g., Kravet and Muslu 2010; Campbell et al. 2010).

For an alternative measure of client difference, I use the absolute value of unexpected discretionary accruals. Following DeFond and Jiambalvo (1994), I estimate accruals with a cross-sectional modified Jones model run within SIC 2-digit industries. The residual from this model is $DACC$, the unexpected discretionary accruals. I do not interpret this residual other than as a difference from the modeled relationship. As expected, all the similarity scores are significantly negatively correlated with $|DACC|$.

As a final validation, I expect that client overlap will not change dramatically over short time periods because of the general stability in financial statements, related disclosures, and client portfolios. Large changes in the similarity measure from year to year are unlikely if the similarity measure is capturing the desired construct. In untabulated analysis, I calculate the autocorrelation coefficient for SIM_{FS} (0.66), SIM_{BUS} (0.83), $SIM_{MD\&A}$ (0.83), and SIM_{NOTES} (0.86), indicating a high level of time series stability in all four measures.

5.2 Audit Fee Model

The tests rely on the following base audit fee model developed from the audit fee meta-analysis in Hay et al (2006):

$$\begin{aligned} LNFEES = a_0 &+ a_1 SIZE + a_2 CSITEMS + a_3 IRISK + a_4 LOSS + a_5 LEV \\ &+ a_6 DELAY + a_7 NAS + a_8 BUSY + a_9 OPIN + a_{10} ICMW \\ &+ a_{11} TENURE + a_{12} INDSPEC + industry + year + \varepsilon \end{aligned} \quad (1)$$

I include controls for various client attributes, some of which are also used to calculate client similarity. The natural log of client assets is *SIZE*. I proxy for client complexity by counting the number of non-zero/non-missing items for that client-year in Compustat (*CSITEMS*). Inherent risk (*IRISK*) is receivables plus inventory, scaled by total assets. *LOSS*, a dummy set to one for negative net income, proxies for financial weakness. Finally, leverage (*LEV*) is long-term debt scaled by total assets.

I also control for auditor and engagement attributes. If the number of days between fiscal year end and the issuance of the 10-K is more than 90 days, then *DELAY* is set to one as a proxy for audit complexity. The log of the dollar amount of non-audit services is *NAS*.[16] Clients with a December 31 fiscal year end date could lead to increased resource constraints, so *BUSY* is a dummy set to one for these companies. Audits leading to anything other than a standard opinion might be associated with additional audit effort or risk. Therefore, *OPIN* is a dummy set to one for non-standard opinions, almost always a clean opinion with modified language. I construct a similar measure for internal control, setting *ICMW* to one if the auditor has noted a material weakness in internal control. *TENURE* is the number of years the client has been with the current auditor, according to Compustat. To ensure my measures capture a construct distinct from traditional proxies for industry specialization, I include *INDSPEC*, a dummy set to one when the current auditor receives at least 32.5% of the total fees available within the client's GICS industry and year.[17] Finally, I control for industry and year fixed effects. All controls are

[16] I first add 1 to the non-audit fees to avoid taking the log of 0.
[17] Since the literature has not extensively explored GICS industries in a specialization context, I use the sample's 75th percentile as a cutoff. I prefer GICS to SIC as the similarity scores are calculated using this categorization. Results are qualitatively unchanged when using a more typical 30% cutoff based on SIC 2-digit codes.

expected to be positive, except *INDSPEC* which is unpredicted. For these variables, Table 2 contains descriptive statistics in Panel A and correlations in Panel B. The patterns are consistent with prior audit fee literature.

5.3 Hypothesis 1

To test the first hypothesis regarding client similarity and audit fees, I augment the base model (1) with one or more of the similarity variables. H1 predicts the coefficients on these similarity variables will be negative. I begin by testing the model with SIM_{FS} to assess the relationship between fees and financial statement similarity, with the results in Table 3.[18] The coefficient on SIM_{FS} is significantly negative (t = -9.41), as predicted, so fees are lower as the financial statements of a client are more similar to other clients of its auditor. All control variables are significant and in the expected direction except for *LEV*, which is insignificant. While size, complexity, and risk are still important determinants of audit fees, it appears that financial statement overlap with other clients is also relevant.

I now turn to the three narrative disclosure similarity measures. For this test, I leave SIM_{FS} in the model as a control for underlying economic similarity and then alternately test the coefficients on SIM_{BUS}, $SIM_{MD\&A}$, and SIM_{NOTES}. Each of the narrative coefficients is significantly negative (t = -9.98, t = -4.77, and t = -2.86, respectively). The results are qualitatively unchanged if SIM_{FS} is excluded from these models. Once again, the control variables are as expected, with the exception of *LEV*. Overall, there is strong evidence that the similarity of client narrative disclosures is negatively related to audit fees, even after controlling for the similarity of the underlying financial statements.

[18] All standard errors are heteroscedasticity-consistent using a Huber-White adjustment.

As an evaluation of the economic significance of the of the effect, moving from the 25th to the 75th percentile of SIM_{FS} decreases audit fees by a range of 2.8% in the footnote model to 3.8% in the business description model. Corresponding changes in SIM_{BUS}, $SIM_{MD\&A}$, and SIM_{NOTES} are associated with additional declines in audit fees of 4.5%, 2.1%, and 1.5%, respectively. The largest combined effect is in the business description model where combined interquartile changes in both the financial statements and business description are associated with an 8.3% decrease in fees. Even the smallest economic effect—the footnotes—is a combined 4.3%. By comparison, the economic effect of being an industry specialist based on market share (*INDSPEC*) increases audit fees by a range of 4.7% to 8.8%, depending on which of the four models in Table 3 is considered. Therefore, the relation between client commonality and audit fees is both statistically and economically significant.

5.4 Hypothesis 2

The second hypothesis predicts the negative relation in H1 is magnified in industries that are more economically important to the auditor. Chung and Kallapur (2003) measure individual client importance by calculating the client's audit fees scaled by total fees received by the auditor in that year. My tests require a measure of the importance of a group of clients, rather than one specific client.[19] Therefore, as a proxy for economic importance I use portfolio share (*PORTSHR*), the audit fees received from a particular industry-year divided by the auditor's total fees from all industries in the same year. In keeping with the industry specialization literature, I include in the model the importance of the industry (*IMPIND*), which

[19] Using their measure directly would bias in favor of a result in my setting since fees would effectively be both a dependent and independent variable.

is a dummy set to one if the portfolio share exceeds 2.8% (the upper quartile of *PORTSHR*). The industry specialization literature has previously used this measure as a proxy for an industry's economic importance to the auditor (Neal and Riley Jr. 2004).

To test H2, I expand the model for the first hypothesis by adding a main term for *IMPIND* and its interaction with each similarity score. The interaction coefficients will be negative under H2's prediction that greater economic incentives accentuate the negative relation between similarity and fees. As shown in Table 4, the significant *IMPIND* main effects are positive, consistent with other studies, indicating that higher portfolio share is associated with higher fees.

Focusing first on financial statement similarity, the coefficient on SIM_{FS} remains negative (t = -7.68) as found in the test of H1. The interaction of *IMPIND* and SIM_{FS} is also negative (t = -3.54), supporting the economic incentive hypothesis. Moving on to the narrative disclosure similarities, I retain SIM_{FS} in the model to control for underlying financial statement similarity. I then add each narrative disclosure score and its interaction with *IMPIND* to the model. The main effects remain significant, as previously found, but none of the interactions are significant.

While I can easily reject the H2 null for financial statements, the results for the narrative disclosures are not significant. One ex-post interpretation for this outcome is that financial statements are quantitative, which should allow for specific technological improvements that would be more difficult to implement for soft, qualitative disclosures. In other words, while narratives *proxy* for underlying client similarity, it may be difficult to implement audit technology that specifically leverages this type of overlap. Financial statements are also likely to be more stable than text, which should more easily allow for technology improvements. In untabulated analysis, I use *PORTSHR* in place of *IMPIND* as an alternative measure of industry importance. The conclusions do not change for the financial statements, MD&A, or footnotes, but the business description has significantly negative coefficients on both the main and interaction terms. Supporting the importance of stability within the narrative disclosures, the business description is the most stable of the three textual items and it also has the strongest support for H2 in this alternative analysis.[20]

Overall, I find some empirical evidence that the negative similarity-fees relationship is incrementally negative as an industry becomes more important to the auditor's revenue stream. As portfolio share increases, the auditor may gain more knowledge and streamline its process for these clients. Alternatively, the auditor may not have a different cost structure due to technological investments, but is just more willing or able to charge lower fees to retain these economically important clients.

5.5 Hypothesis 3

The final hypotheses examine the consistency between financial statements and narrative disclosures. To test these hypotheses, I split the financial statement similarity and each of the narrative disclosure similarities into terciles. I am particularly interested in misalignment between the lowest and highest terciles of the financials and text, so I create dummies indicating when such misalignments occur. For each narrative disclosure type, I set the corresponding *POOLTEXT* variable to one when financial statement similarity is low (SIM_{FS} is in the bottom tercile) and narrative similarity is high (SIM_{BUS}, $SIM_{MD\&A}$, or SIM_{NOTES} is in the top tercile). These

[20] Using the raw year-over-year disclosure modification score from Brown and Tucker (2011), the business description has an average modification score of only 0.09, as compared to much larger annual modification scores of 0.16 for MD&A and 0.14 for the footnotes.

indicators correspond to the riskiest type of disclosure inconsistency—*pooled text*—since the financial statements portray a very atypical company for the industry, while its narrative disclosures are very similar to its peers.

I then create *DIFFTEXT* dummies set to one when financial statement similarity is high (SIM_{FS} is in the top tercile) and narrative similarity is low (SIM_{BUS}, $SIM_{MD\&A}$, or SIM_{NOTES} is in the bottom tercile). While still inconsistent, these *differentiated text* misalignments have potentially benign—and potentially beneficial—explanations. Examining each of the narrative disclosures in separate models, I expand the base audit fee model to include the respective *POOLTEXT* and *DIFFTEXT* dummy for that disclosure type. In each model, I also control for SIM_{FS} and the similarity of the textual disclosure being examined. H3a predicts the coefficient on *POOLTEXT* is positive and H3b predicts the coefficient on *DIFFTEXT* is nonzero (although a null result would not be unexpected).

Table 5 presents the results of the test. Consistent with the earlier tests of H1, all the financial statement and narrative disclosure similarity scores are significantly negative. As predicted by H3a, the coefficients on *POOLTEXT* for the business description, MD&A, and footnotes are all significantly positive (t = 5.05, t = 4.43, and t = 2.40, respectively). These results support the prediction that inconsistency in the form of pooled text (dissimilar financials, similar text) is associated with higher audit fees. Relative to other clients, the findings are consistent with pooled text clients either (1) representing higher idiosyncratic risk to the auditor or (2) leading to a lower willingness to implement technological improvements to take advantage of client commonality.

Turning to the test of differentiated narrative disclosures, *DIFFTEXT*$_{BUS}$ is significantly negative (t = -2.50, p-value = 0.012), as is *DIFFTEXT*$_{MD\&A}$ (t = -2.73). *DIFFTEXT*$_{NOTES}$ is negative, but insignificant (t = -1.07), potentially due to the much smaller sample size for the footnotes. The hypothesis makes only weak predictions about these coefficients because it is unclear whether differentiated text increases, decreases, or does not affect the risk and efficiency associated with auditing these inconsistent clients. However, the results support the idea that differentiated disclosures reduce risk or increase audit efficiency, even when they are inconsistent with the financial statements.

6. Alternative Measures and Sensitivity Analyses

6.1 Larger Reference Groups

The primary narrative disclosure measures are calculated based on the similarity to the five clients that are most similar to the observation. To test the sensitivity of the results to this choice, I construct textual similarities using *all* clients in the auditor-industry-year. These three alternative measures have correlations with the original measures that range from 0.91 to 0.94.

Compared with the original test of H1, the negative relation between client similarities and audit fees is qualitatively similar for the business description (t = -6.24), but somewhat weaker for the MD&A (t = -1.97; p-value = 0.048) and footnotes (t = -2.27; p-value = 0.023). The patterns for the second hypothesis are unchanged. For the final hypotheses regarding disclosure consistency, the results are qualitatively unchanged except that *DIFFTEXT*$_{MD\&A}$ is now slightly less significant (t = -2.02; p-value = 0.044) and *DIFFTEXT*$_{BUS}$ is no longer significant.

Overall, the results are slightly weaker in a few cases as the reference group is expanded to include more dissimilar clients. The changes in significance could be due to additional measurement error in the proxies as less relevant peers affect the calculations. This pattern is also consistent with auditors either explicitly or implicitly taking into account the similarity of more narrowly constructed client subgroups than the GICS industry as defined by Standard & Poor's.

6.2 Minimum Reference Group Size

Rather than requiring the auditor-industry-year to have at least five clients, I alternatively require at least ten clients from which to choose the five most similar. The results are qualitatively unchanged for the business description and MD&A samples, but weaker for the footnote sample. These changes in significance seem to be attributable to a reduction in sample size from 9,806 observations to only 6,317.

6.3 Accounting System Comparability

While I consider a broad notion of client overlap using multiple financial statement variables and narrative disclosure language, De Franco et al (2011) specifically examine the comparability of accounting systems between companies. For each company, they regress 16 quarters of earnings (an accounting system output) on returns (the net 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*). As a sensitivity test to my primary D^2 metric, I calculate these two measures as described in more detail in De Franco et al (2011) as an alternative to SIM_{FS}.

The *COMPACCT-IND* variable is uncorrelated with the four primary similarity scores I use in the current study. Nor is it correlated with any of the proxies for client differences (|*SIZEDIFF*|, |*IRISKDIFF*|, |*LEVDIFF*|, and *DACC*). In contrast, *COMPACCT-R2* has a correlation of -0.10 with SIM_{FS} and correlations of 0.06, 0.02, and 0.04 with SIM_{BUS}, $SIM_{MD\&A}$ and SIM_{NOTES}, respectively. As an alternative test of H1, I separately include the two accounting comparability measures in the base audit fee model. They are both significantly negative (t = -5.01 for *COMPACCT-IND* and t = -2.14 for *COMPACCT-R2*). These results hold whether or not I include SIM_{FS} in the model, although SIM_{FS} has a much higher economic magnitude and a more negative t-statistic in both cases. I find no support for the second hypothesis when using these alternative measures. However, they strongly support H3a and H3b regarding disclosure consistency. Using *COMPACCT-IND*, all of the *POOLTEXT* and *DIFFTEXT* coefficients are qualitatively similar to the original tests except that $DIFFTEXT_{NOTES}$ also becomes significantly negative, making it consistent with the business description and MD&A results.

Using the relation between earnings and returns, De Franco et al (2011) develop an empirical proxy that is directly related to their theoretical construct. However, even though the statistical significance of their measures are similar to mine, SIM_{FS} is much more strongly related to audit pricing in terms of economic magnitude than their measures of accounting comparability. Therefore, depending on the context, each approach could provide unique insights as proxies for company similarity. The advantages of my approach are that it requires no knowledge about the functional form of the relationship, requires less time series data, and can include an arbitrary number of economic dimensions in the similarity score.[21]

7. Conclusion

I introduce measures of financial statement and narrative disclosure similarity as proxies for audit client overlap. As predicted, I find higher commonality among clients is associated with lower audit fees, which I interpret as reduced production costs arising from increased audit efficiency and reduced risk due to greater potential for improved audit technology and shared knowledge. These patterns are stronger when the auditor has higher financial incentives to profit from the non-idiosyncratic elements of the audit. I also find that inconsistencies between financial statements and narrative disclosures are associated with higher fees when these differences are consistent with the client attempting to reduce its apparent financial differences with peer companies. In contrast, I find lower audit fees when the narrative disclosures differ from financial statements in a manner consistent with the client revealing differentiating firm-specific information.

The measures I develop in this paper have additional potential applications in audit research. For example, client commonality could be relevant to a company that is choosing whether to keep their incumbent auditor or switch to a new one. A company might look for an auditor that already audits similar firms (or dissimilar firms if knowledge spillover is a

---

[21] The DeFranco et al (2011) approach can only be used as an alternative to SIM_{FS}, the financial statement similarity, and not as a proxy for narrative disclosure similarity.

competitive concern). Outside of audit research, the financial statement and narrative disclosure inconsistency result could have implications for a company's information environment. There could also be econometric applications when the research design calls for a control company that is very similar to the original observation. While somewhat related to other measures of company similarity, such as the accounting comparability measure in De Franco et al (2011), I provide a broader alternative that could be preferable in certain research contexts.

In addition to the contribution provided by the measures themselves, the proxies allow me to explore topics that were previously difficult to examine empirically. I provide a more direct proxy for the potential of specialization than merely using the prevalent industry market share measures, which can be difficult to interpret. This paper is also one of the few to integrate narrative disclosures as an empirical proxy for elements of the audit.

References

Amir, E. 1993. The Market Valuation of Accounting Information: The Case of Postretirement Benefits Other Than Pensions. *The Accounting Review* 68 (4): 703-724.

Bamber, L. S., and Y. S. Cheon. 1998. Discretionary Management Earnings Forecast Disclosures: Antecedents and Outcomes Associated with Forecast Venue and Forecast Specificity Choices. *Journal of Accounting Research* 36 (2): 167-190.

Beck, P. J., and M. G. H. Wu. 2006. Learning by Doing and Audit Quality. *Contemporary Accounting Research* 23 (1): 1-30.

Berry, H., M. F. Guillén, and N. Zhou. 2010. An Institutional Approach to Cross-National Distance. *Journal of International Business Studies* 41 (9): 1460-1480.

Blokdijk, H., F. Drieenhuizen, D. A. Simunic, and M. T. Stein. 2006. An Analysis of Cross-Sectional Differences in Big and Non-Big Public Accounting Firms' Audit Programs. *Auditing* 25 (1): 27-48.

Brown, S. V., and J. W. Tucker. 2011. Large-Sample Evidence on Firms' Year-over-Year MD&A Modifications. *Journal of Accounting Research* 49 (2): 309-346.

Cahan, S., J. Godfrey, J. Hamilton, and D. Jeter. 2008. Auditor Specialization, Auditor Dominance, and Audit Fees: The Role of Investment Opportunities. *The Accounting Review* 83 (6): 1393.

Campbell, J. L., H. Chen, D. S. Dhaliwal, H.-M. Lu, and L. Steele. 2010. The Information Content of Mandatory Risk Factor Disclosures in Corporate Filings. *SSRN eLibrary*.

Causholli, M., M. De Martinis, D. Hay, and W. R. Knechel. 2010. Audit Markets, Fees and Production: Towards an Integrated View of Empirical Audit Research. *Journal of Accounting Literature* 29: 167-215.

Chung, H., and S. Kallapur. 2003. Client Importance, Nonaudit Services, and Abnormal Accruals. *The Accounting Review* 78 (4): 931-955.

Craswell, A. T., J. R. Francis, and S. L. Taylor. 1995. Auditor brand name reputations and industry specializations. *Journal of Accounting and Economics* 20 (3): 297-322.

Darr, E. D., L. Argote, and D. Epple. 1995. The Acquisition, Transfer, and Depreciation of Knowledge in Service Organizations: Productivity in Franchises. *Management Science* 41 (11): 1750-1762.

Davis, L. R., D. N. Ricchiute, and G. Trompeter. 1993. Audit Effort, Audit Fees, and the Provision of Nonaudit Services to Audit Clients. *The Accounting Review* 68 (1): 135-150.

DeFond, M. L., and J. Jiambalvo. 1994. Debt Covenant Violation and Manipulation of Accruals. *Journal of Accounting and Economics* 17 (1-2): 145-176.

Dowling, C. 2009. Appropriate Audit Support System Use: The Influence of Auditor, Audit Team, and Firm Factors. *The Accounting Review* 84 (3): 771.

Dunn, K. A., and B. W. Mayhew. 2004. Audit Firm Industry Specialization and Client Disclosure Quality. *Review of Accounting Studies* 9 (1): 35-58.

Feldman, R., S. Govindaraj, J. Livnat, and B. Segal. 2010. Management's Tone Change, Post Earnings Announcement Drift and Accruals. *Review of Accounting Studies* 15 (4): 915-953.

De Franco, G., S. P. Kothari, and R. S. Verdi. 2011. The Benefits of Financial Statement Comparability. *Journal of Accounting Research* 49 (4): 895-931.

Gramling, A. A., and D. N. Stone. 2001. Audit firm industry expertise: A review and synthesis of the archival literature. *Journal of Accounting Literature* 20: 1.

Guilding, C., and L. McManus. 2002. The Incidence, Perceived Merit and Antecedents of Customer Accounting: An Exploratory Note. *Accounting, Organizations and Society* 27 (1-2): 45-59.

Hackenbrack, K., and W. R. Knechel. 1997. Resource allocation decisions in audit engagements. *Contemporary Accounting Research* 14 (3): 481.

Hair, J. F., W. C. Black, B. J. Babin, R. E. Anderson, and R. L. Tatham. 2006. *Multivariate Data Analysis*. 6th ed. Upper Saddle River, NJ: Pearson Prentice Hall.

Hanley, K. W., and G. Hoberg. 2012. Litigation risk, strategic disclosure and the underpricing of initial public offerings. *Journal of Financial Economics* 103 (2): 235-254.

Hay, D. C., W. R. Knechel, and N. Wong. 2006. Audit Fees: A Meta-analysis of the Effect of Supply and Demand Attributes. *Contemporary Accounting Research* 23 (1): 141-191.

Iyer, V. M. 1998. Characteristics of Accounting Firm Alumni Who Benefit Their Former Firm. *Accounting Horizons* 12 (1): 18-30.

Jaffe, A. B. 1986. Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits, and Market Value. *The American Economic Review* 76 (5): 984-1001.

Knechel, W. R., L. Niemi, and M. Zerni. 2011. Empirical Evidence on the Implicit Determinants of Compensation in Big four Audit Partnerships. *Working Paper*.

Kravet, T. D., and V. Muslu. 2010. Informativeness of Risk Disclosures in Corporate Annual Reports. *SSRN eLibrary*.

Li, F. 2008. Annual Report Readability, Current Earnings, and Earnings Persistence. *Journal of Accounting and Economics* 45 (2-3): 221-247.

———. 2010. The Information Content of Forward-Looking Statements in Corporate Filings-A Naïve Bayesian Machine Learning Approach. *Journal of Accounting Research* 48 (5): 1049-1102.

Low, K.-Y. 2004. The Effects of Industry Specialization on Audit Risk Assessments and Audit-Planning Decisions. *The Accounting Review* 79 (1): 201.

De Maesschalck, R., D. Jouan-Rimbaud, and D. L. Massart. 2000. The Mahalanobis Distance. *Chemometrics and Intelligent Laboratory Systems* 50 (1): 1-18.

Mahalanobis, P. C. 1936. On the Generalized Distance in Statistics. *Proceedings of the National Institute of Sciences of India* 2 (1): 49–55.

Manning, C. D., and H. Schütze. 1999. *Foundations of Statistical Natural Language Processing*. Cambridge, Mass: MIT Press.

Mayhew, B. W., and M. S. Wilkins. 2003. Audit Firm Industry Specialization as a Differentiation Strategy: Evidence from Fees Charged to Firms Going Public. *Auditing* 22 (2): 33-52.

Mimmack, G. M., S. J. Mason, and J. S. Galpin. 2001. Choice of Distance Matrices in Cluster Analysis: Defining Regions. *Journal of Climate* 14 (12): 2790-2797.

Minutti-Meza, M. 2010. Does Auditor Industry Specialization Improve Audit Quality? Evidence from Comparable Clients. *Working Paper*.

———. 2012. Does Using Information From Peer Firms Improve Account-Level Expectation Models? *Working Paper*.

Neal, T. L., and R. R. Riley Jr. 2004. Auditor Industry Specialist Research Design. *Auditing* 23 (2): 169-177.

Numan, W., and M. Willekens. 2011. An empirical test of spatial competition in the audit market. *Journal of Accounting and Economics* Forthcoming.

O'Keefe, T. B., R. D. King, and K. M. Gaver. 1994. Audit Fees, Industry Specialization, and Compliance with GAAS Reporting Standards. *Auditing* 13 (2): 41-55.

O'Keefe, T. B., D. A. Simunic, and M. T. Stein. 1994. The Production of Audit Services: Evidence from a Major Public Accounting Firm. *Journal of Accounting Research* 32 (2): 241-261.

Owhoso, V. E., W. F. Messier Jr., and J. G. Lynch Jr. 2002. Error Detection by Industry-Specialized Teams during Sequential Audit Review. *Journal of Accounting Research* 40 (3): 883-900.

Palmrose, Z.-V. 1986. Audit Fees and Auditor Size: Further Evidence. *Journal of Accounting Research* 24 (1): 97-110.

Public Company Accounting Oversight Board. 2011. *Concept Release on Possible Revisions to Pcaob Standards Related to Reports on Audited Financial Statements*. *PCAOB Release No. 2011-003*.

Rege, U. P. 1984. Accounting Ratios to Locate Take-Over Targets. *Journal of Business Finance & Accounting* 11 (3): 301-311.

Riedl, E., and S. Srinivasan. 2010. Signaling Firm Performance Through Financial Statement Presentation:: An Analysis Using Special Items. *Contemporary Accounting Research* 27 (1): 289.

Salton, G., A. Wong, and C. S. Yang. 1975. A Vector Space Model for Automatic Indexing. *Communications of the ACM* 18 (11): 613-620.

Securities and Exchange Commission. 1987. *Concept Release on Management's Discussion and Analysis of Financial Condition and Operations*. *Release No. 33-6711*.

———. 1989. *Management's Discussion and Analysis of Financial Condition and Results of Operations; Certain Investment Company Disclosures*. *Release No. 33-6835*.

———. 2003. *Commission Guidance Regarding Management's Discussion and Analysis of Financial Condition and Results of Operations*. *Release No. 33-8350*.

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

Singhal, A. 2001. Modern Information Retrieval: A Brief Overview. *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering* 24 (4): 35-43.

Sirois, L.-P., and D. A. Simunic. 2010. Auditor Size and Audit Quality Revisited: The Importance of Audit Technology. *SSRN eLibrary*.

Sohn, M.-W. 2001. Distance and Cosine Measures of Niche Overlap. *Social Networks* 23 (2): 141-165.

Sun, Y. 2010. Do MD&A Disclosures Help Users Interpret Disproportionate Inventory Increases? *The Accounting Review* 85 (4): 1411.

Tasker, S. C. 1998. Bridging the Information Gap: Quarterly Conference Calls as a Medium for Voluntary Disclosure. *Review of Accounting Studies* 3 (1): 137-167.

Wahlen, J. M. 1994. The Nature of Information in Commercial Bank Loan Loss Disclosures. *The Accounting Review* 69 (3): 455-478.

Ward, D. D., R. J. Elder, and S. C. Kattelus. 1994. Further Evidence on the Determinants of Municipal Audit Fees. *The Accounting Review* 69 (2): 399-411.

Willenborg, M. 2002. Discussion of "Brand Name Audit Pricing, Industry Specialization, and Leadership Premiums post-Big 8 and Big 6 Mergers." *Contemporary Accounting Research* 19 (1): 111-116.

Appendix A
Extraction of Annual Report Items

To gather the business description, MD&A, and footnotes sample, I begin by downloading all 10-K's and 10-K405's available on the SEC's EDGAR system that meet the following requirements: (1) fiscal years between 2000 and 2009, (2) assets greater than $1 million, (3) no change in fiscal year-end, (4) not in the financial and utility industries, and (5) engaging a Big4 auditor. As described in Table A1, this initial screen leaves 29,874 annual reports (45,304 Compustat observations minus 15,430 annual reports missing on EDGAR).

I next screen out any unusually short annual reports since these typically belong to holding companies, firms that are winding down, and other atypical observations. I use a cutoff of 50,000 characters for this purpose (approximately the 4^{th} percentile of 10-K length). This value seems to filter out most of the unwanted observations without losing a substantial number of desired reports. I 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. These filters leave 29,205 annual reports.

I begin the item extraction process by stripping all HTML formatting and data tables as in Li (2008; 2010). I then split each annual report into its component items, keeping only the business description, MD&A, and footnotes (the financial statements were removed when data tables were discarded).

I remove any narrative disclosures that contain language indicating the relevant section has been omitted as permitted by regulation. I 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. I 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 my initial string search did not recognize.

I now 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, I 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.

Appendix B
Calculation of Narrative Disclosure Similarity Score

As described in Brown and Tucker (2011), the Vector Space Model (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:

$$v = (w_1, w_2, \dots, w_n)$$

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:

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

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

The vectors allow for various 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 zero (completely dissimilar documents) to one (identical documents). I stem all words using the Porter stemming algorithm to reduce the dimensionality of the data, which in turn limits the computing time and resources required (e.g., "earnings," "earned," and "earn" are all converted to "earn").[22] Consistent with Brown and Tucker (2011), I use the term frequency-inverse document frequency (TF-IDF) algorithm to decrease the weight on frequently used words and increase the weight on uncommon words.[23] 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) are interested in the differences between just two documents at a time, they only calculate pairwise similarity scores. In contrast, I aggregate these pairwise scores 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. Aggregating the pairwise similarities yields a summary measure that is compatible with SIM_{FS}. To combine the pairwise $SIM_{DOC,i,j}$ scores between client i and all other clients j in the same auditor, industry, and year, I average the five highest pairwise similarities to get $SIM_{DOC,i}$ for each observation in my sample. I use only the most similar other clients because the efficiency gains and knowledge spillovers are likely to increase at a declining rate, implying a potential threshold effect.[24]

I calculate the $SIM_{DOC,i}$ similarity measure for each observation in the business description ($RAWSIM_{BUS}$), MD&A ($RAWSIM_{MD\&A}$) and footnote ($RAWSIM_{NOTES}$) samples.

---

[22] 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 6 gigabytes of working memory.

[23] I 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. I generate the TF-IDF weights independently for each type of narrative disclosure.

[24] I alternatively use the average of all other clients in a sensitivity analysis.

However, Brown and Tucker (2011) show that these raw scores are positively related to document length because of the mechanics of the calculation, rather than due to any meaningful underlying relation. 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 the current study). Since I do not use consecutive-year documents, which have a relatively constant length, I also include the first five powers of the average length of the reference group ($GRPLEN_{BUS}$, $GRPLEN_{MD\&A}$, and $GRPLEN_{NOTES}$).[25] Finally, I include the first five powers of the number of clients in the auditor-industry-year ($GRPSIZE_{BUS}$, $GRPSIZE_{MD\&A}$, and $GRPSIZE_{NOTES}$). This correction is necessary because the raw score is based on the average similarity of the five most similar clients, which will mechanically increase with the size of the auditor-industry-year from which the five are chosen.[26] Regressing the raw similarity scores on the first five powers of the observation document length and the average reference group length yields a residual that represents the variation in the raw similarity scores that cannot be explained by these factors. I label these residuals SIM_{BUS}, $SIM_{MD\&A}$, and SIM_{NOTES}, producing the similarity scores I use in my analysis.

[25] Hanley and Hoberg (2012) use the VSM to measure the similarity of an IPO prospectus to all the recent IPO's experiencing litigation problems. However, they do not control for document length, making it difficult to ascertain the validity of their measure.
[26] For example, if the auditor-industry-year only has five other clients, some of those clients might be quite dissimilar to the observation of interest, leading to a lower raw similarity measure. On the other hand, if the auditor-industry-year has 30 other clients, the probability is high that the five most similar clients are fairly close to the observation of interest, which leads to a higher raw similarity score.

Table A1
Narrative Disclosure Sample Selection Process

	Reports		
10-K available on EDGAR; fiscal years 2000-2009; Compustat assets > $1M; no FYE change; excl. financials and utilities; Big4 auditor	30,906		
Less: Short reports (<50,000 characters)	(693)		
Total annual reports available	30,213		
	Bus Desc	**MD&A**	**Footnotes**
Less: Item not succesfully extracted	(1,233)	(511)	(659)
Less: Item specifically omitted	(13)	(10)	(67)
Less: Item included by reference	(16)	(1,664)	(7,454)
Less: Short items (<150 characters)	(1,143)	(1,032)	(2,519)
Less: < 500 or > 20,000 words	(2,720)	(2,946)	(6,449)
Less: < 5 other clients in auditor-industry-year	(1,942)	(1,904)	(2,399)
Total items available	23,146	22,146	10,666

Table 1
Similarity Measures

Panel A: Descriptive Statistics

Variable	Mean	Std Dev	25%	Median	75%	Obs
***SIM**$_{FS}$*	(4.316)	3.250	(4.963)	(3.425)	(2.436)	32,412
***SIM**$_{BUS}$*	(0.000)	0.133	(0.093)	(0.031)	0.062	23,146
***SIM**$_{MD\&A}$*	0.000	0.145	(0.093)	(0.038)	0.051	22,146
***SIM**$_{NOTES}$*	0.000	0.080	(0.048)	(0.022)	0.019	10,666
***RAWSIM**$_{BUS}$*	0.190	0.142	0.083	0.153	0.261	25,088
***RAWSIM**$_{MD\&A}$*	0.185	0.154	0.074	0.136	0.246	24,050
***RAWSIM**$_{NOTES}$*	0.083	0.085	0.030	0.054	0.102	13,065
***LEN**$_{BUS}$*	6,561	3,792	3,787	5,671	8,564	25,088
***LEN**$_{MD\&A}$*	8,103	3,771	5,289	7,619	10,393	24,050
***LEN**$_{NOTES}$*	9,559	3,930	6,480	9,123	12,179	13,065
***GRPLEN**$_{BUS}$*	6,547	2,119	4,945	6,257	7,844	25,088
***GRPLEN**$_{MD\&A}$*	8,096	2,031	6,899	8,291	9,482	24,050
***GRPLEN**$_{NOTES}$*	9,543	2,418	7,820	9,547	11,255	13,065
***GRPSIZE**$_{BUS}$*	21.915	15.533	10.000	19.000	29.000	25,088
***GRPSIZE**$_{MD\&A}$*	22.172	16.572	10.000	19.000	29.000	24,050
***GRPSIZE**$_{NOTES}$*	11.974	8.780	6.000	10.000	16.000	13,065

Subscripts: *FS* = Financial Statements, *BUS* = Business Description, *MD&A* = Management's Discussion & Analysis, *NOTES* = Footnotes to Financial Statements. *SIM* = similarity of observation to other clients in the reference group, standardized for *LEN*, *GRPLEN*, and *GRPSIZE*; higher value indicates more similarity. *RAWSIM* = *SIM* before standardization. *LEN* = # of words in the observation's text. *GRPLEN* = mean # of words in reference group text. *GRPSIZE* = # of observations in the reference group.

Table 1 (continued)

Panel B: Pairwise Pearson Correlations of Similarity Measures and Proxies for Differences

	SIM_FS	*SIM_BUS*	*SIM_MD&A*	*SIM_NOTES*	|*SIZEDIFF*|	|*IRISKDIFF*|	|*LEVDIFF*|
SIM_BUS	**0.08**						
SIM_MD&A	**0.08**	**0.73**					
SIM_NOTES	**0.16**	**0.59**	**0.70**				
|*SIZEDIFF*|	**-0.08**	**-0.05**	**-0.07**	**-0.03**			
|*IRISKDIFF*|	0.01	**-0.10**	**-0.11**	**-0.11**	**0.10**		
|*LEVDIFF*|	**-0.03**	0.00	**-0.03**	-0.01	**0.01**	**0.00**	
|*DACC*|	**-0.05**	**-0.07**	**-0.06**	**-0.05**	**0.05**	**0.08**	**0.01**

Correlations in bold are significant at the 5% level. Those within dashed box are generally expected to be negative. Based only on observations with at least five observations in all four reference groups. |*SIZEDIFF*| = absolute difference from mean of reference group *SIZE* (log of total assets). |*IRISKDIFF*| = absolute difference from mean of reference group *IRISK* ([receivables + inventory] / assets). |*LEVDIFF*| = absolute difference from mean of reference group *LEV* (long-term debt / assets). |*DACC*| = absolute unexpected discretionary accruals from cross-sectional modified Jones model.

Table 2
Audit Fee Model Variables

Panel A: Descriptive Statistics

Variable	Mean	Std Dev	25%	Median	75%
LNFEES	13.384	1.332	12.403	13.367	14.269
AT	3,590.358	13,797.860	110.045	439.236	1,771.330
SIZE	6.119	2.050	4.701	6.085	7.479
CSITEMS	165.494	33.216	142.000	164.000	188.000
IRISK	0.235	0.185	0.085	0.199	0.341
LOSS	0.383				
LEV	0.198	0.302	0.001	0.119	0.300
DELAY	0.253				
NAS	11.260	3.575	10.692	12.024	13.190
BUSY	0.710				
OPIN	0.474				
ICMW	0.030				
TENURE	9.712	8.567	3.000	7.000	13.000
INDSHR	0.254	0.111	0.171	0.244	0.325
PORTSHR	0.022	0.015	0.011	0.019	0.028
INDSPEC	0.249				
IMPIND	0.282				
Observations	**33,006**				

For each client: *LNFEES* = log of audit fees. *AT* = total assets. *SIZE* = log of *AT*. *CSITEMS* = # of non-missing/non-zero variables in Compustat. *IRISK* = receivables plus inventory, scaled by *AT*. *LOSS* = 1 if net income < 0. *LEV* = long-term debt, scaled by *AT*. *DELAY* = 1 if 10-K filed > 90 days after fiscal year end. *NAS* = log of non-audit fees. *BUSY* = 1 if 12/31 fiscal year end. *OPIN* = 1 for non-standard opinion. *ICMW* = 1 if material weakness in internal control. *TENURE* = number of years with the current auditor. *INDSHR* = % of industry's fees provided to the current auditor. *PORTSHR* = % of current auditor's fees provided by the client's industry-year. *INDSPEC* = 1 if *INDSHR* >= 32.5%. *IMPIND* = 1 if *PORTSHR* >= 2.8% (economically important industry).

Table 2 (continued)

Panel B: Pairwise Pearson Correlations of Continuous Variables in Audit Fee Model

	LNFEES	*SIZE*	*CSITEMS*	*IRISK*	*LEV*	*NAS*	*TENURE*	*INDSHR*	*PORTSHR*
SIM_{FS}	0.02	**0.13**	**0.11**	**0.10**	**0.07**	**0.02**	**0.05**	0.01	**-0.10**
SIM_{BUS}	**0.11**	**0.27**	**0.06**	**-0.17**	**0.10**	**-0.05**	**0.03**	**0.09**	**0.09**
$SIM_{MD\&A}$	0.02	**0.18**	**-0.02**	**-0.16**	**0.03**	**-0.05**	**0.03**	0.01	**-0.02**
SIM_{NOTES}	0.01	**0.20**	**-0.03**	**-0.16**	**0.08**	**-0.04**	**0.06**	**0.04**	**0.13**
LNFEES		**0.75**	**0.77**	**0.03**	**0.12**	**0.28**	**0.31**	**0.09**	**0.03**
SIZE			**0.65**	**-0.07**	**0.19**	**0.38**	**0.33**	**0.06**	**0.10**
CSITEMS				**0.12**	**0.16**	**0.28**	**0.36**	**0.04**	**-0.03**
IRISK					**-0.09**	**0.06**	**0.10**	**-0.04**	**-0.09**
LEV						**0.06**	-0.01	**-0.03**	**0.07**
NAS							**0.18**	**0.02**	-0.02
TENURE								**0.08**	0.01
INDSHR									**0.32**

Correlations in bold are significant at the 5% level. *SIM* = similarity of observation to other clients in the financial statement (*FS*), business description (*BUS*), MD&A (*MD&A*), and footnote (*NOTES*) reference groups. See Table 1, Panel B for correlations of *SIM* measures.

Table 3
OLS Regression of Audit Fees on Client Similarity

		LNFEES		*LNFEES*		*LNFEES*		*LNFEES*	
	Exp	**Coef**	**t-stat**	**Coef**	**t-stat**	**Coef**	**t-stat**	**Coef**	**t-stat**
(Intercept)		8.169	296.35***	8.275	268.65***	8.392	268.73***	8.426	182.00***
SIZE	+	0.414	135.46***	0.391	109.64***	0.371	100.40***	0.383	73.19***
CSITEMS	+	0.009	44.70***	0.009	42.46***	0.009	42.87***	0.009	28.30***
IRISK	+	0.534	21.25***	0.463	17.14***	0.440	16.09***	0.446	11.47***
LOSS	+	0.173	21.85***	0.170	19.36***	0.157	17.68***	0.160	12.61***
LEV	+	-0.016	-0.84	-0.034	-2.12**	-0.026	-1.69*	0.005	0.23
DELAY	+	0.022	2.40**	0.112	7.75***	0.117	7.86***	0.064	2.89***
NAS	+	0.021	16.14***	0.018	12.28***	0.018	12.01***	0.013	6.50***
BUSY	+	0.071	9.21***	0.093	10.98***	0.088	10.04***	0.100	8.12***
OPIN	+	0.105	13.97***	0.100	11.74***	0.095	10.95***	0.079	6.49***
ICMW	+	0.512	24.07***	0.445	18.45***	0.446	18.15***	0.448	13.19***
TENURE	+	0.003	6.50***	0.002	5.50***	0.002	3.75***	0.002	2.39**
INDSPEC	?	0.084	10.16***	0.073	7.79***	0.067	7.03***	0.046	3.36***
SIM_{FS}	-	-0.016	-9.41***	-0.015	-8.52***	-0.013	-7.41***	-0.011	-4.35***
SIM_{BUS}	-			-0.339	-9.98***				
$SIM_{MD\&A}$	-					-0.158	-4.77***		
SIM_{NOTES}	-							-0.245	-2.86***
Years		Yes		Yes		Yes		Yes	
Industries		Yes		Yes		Yes		Yes	
Adj R^2		0.808		0.822		0.810		0.803	
Model F		5,729		4,111		3,666		1,632	
Obs		32,412		21,450		20,530		9,806	

H1 predicts negative coefficients on the *SIM* measures. *LNFEES* = log of audit fees. *SIZE* = log of total assets. *CSITEMS* = # of non-missing/non-zero variables in Compustat. *IRISK* = receivables plus inventory, scaled by total assets. *LOSS* = 1 if net income < 0. *LEV* = long-term debt, scaled by total assets. *DELAY* = 1 if 10-K filed > 90 days after fiscal year end. *NAS* = log of non-audit fees. *BUSY* = 1 if 12/31 fiscal year end. *OPIN* = 1 for non-standard opinion. *ICMW* = 1 if material weakness in internal control. *TENURE* = number of years with the current auditor. *INDSPEC* = 1 if the auditor is a market share specialist. *SIM* = similarity of observation to other clients in the financial statement (*FS*), business description (*BUS*), MD&A (*MD&A*), and footnote (*NOTES*) reference groups. Standard errors are Huber-White-adjusted. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively, in a two-tailed test.