AUDITING: A JOURNAL OF PRACTICE & THEORY
Vol. 25, No. 2
November 2006
pp. 53–84

Peer-Based Approach for Analytical Procedures

Rani Hoitash, Alexander Kogan, and Miklos A. Vasarhelyi

SUMMARY: This study extends the existing research in analytical procedures by allowing for learning from contemporaneous information transfers among peer companies. We introduce an approach for selecting peers for each client and perform tests to examine the contribution of peers' information to the performance of analytical procedures. We find that peer data are imperfect substitutions for contemporaneous firm-specific variables when such variables are not in error. However, we observe that contemporaneous peer specific data are especially beneficial when coordinated errors exist in multiple accounts. We demonstrate that when errors are seeded into two contemporaneous accounts, peer models perform better at detecting errors. We also find that fast-changing companies experience inferior prediction and error detection accuracy, and that larger companies experience more accurate prediction, lower Type II errors, and higher Type I errors. Additionally, we observe that significant improvements in the performance of analytical procedures are associated with larger clients indicating that auditors of larger companies can potentially benefit more from the use of peer data.

Keywords: analytical procedures; data management; information sharing; peers.

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

INTRODUCTION

The Statement of Auditing Standard (SAS) No. 56 states that analytical procedures are required in the planning and overall review stages of the audit, and are recommended during substantive testing (American Institute of Certified Public Accountants [AICPA] 1988). Analytical procedures are defined as the diagnostic process of identifying and determining the cause of unexpected fluctuations in account balances and financial ratios. The benefits associated with analytical procedures are considered substantial if they are proven to reduce the most expensive audit task (namely, the test of details), decrease the risk that a material error will go undetected, and if they are constructed to be stable across companies and time horizons.

Rani Hoitash is an Assistant Professor at Suffolk University, and Alexander Kogan and Miklos A. Vasarhelyi are Professors at Rutgers University.

We appreciate helpful comments and suggestions from Robert Knechel (associate editor), and two anonymous referees. We also appreciate helpful comments from Michael Alles, Lynford Graham, Udi Hoitash, and Adrian Kester.

Submitted: June 2004
Accepted: June 2006

The demise of Arthur Andersen and recent business scandals within Enron and WorldCom have also prompted the profession to address the issue of audit effectiveness. Though it was always in the interest of policy makers to improve the quality of audits, the recent scandals provide outside pressure to develop and refine innovative audit techniques including continuous auditing and information sharing in the auditing process. The current paper examines the potential use of a data management system that allows for learning from contemporaneous information transfers among peer companies to improve the performance of analytical procedures. Specifically, we examine whether the inclusion of contemporaneous account balances from peer companies as independent variables in the expectation model contributes to the improvement of prediction accuracy and of the performance of error detection.

Previous studies examined various ways to improve the performance of analytical procedures. A number of these studies document that statistical techniques such as X-11 (Dugan et al. 1985), Vector Autoregression (Dzeng 1994), and structural models (Chen and Leitch 1998; Wild 1987) can be used to improve the prediction performance of account balances and ratios. Additional studies concentrate on other factors that result in prediction improvements such as looking at single industry companies (Wheeler and Pany 1990), and using disaggregated multilocation data (Allen et al. 1999). Lev (1980), Loebbecke and Steinbart (1987), Wild (1987), and Allen (1992) offer evidence to support the fact that exogenous industrial and economy wide data improve the predictive ability of regression based analytical procedures. Further papers investigate the increased effectiveness of using higher frequency data. Cogger (1981), Knechel (1988), and Dzeng (1994) find that the use of monthly data greatly enhances the effectiveness of these models.

The objective of this study is to use an empirical approach to examine whether information transfer of contemporaneous variables among peer companies results in improved prediction and error detection. The study investigates whether account balance expectations that are generated using contemporaneous peer data result in superior prediction and error detection performance in comparison to those that are derived using only company specific data. This study examines three research questions. We first examine whether models that include peer data lead to improved prediction accuracy. We observe that peer data is extremely useful when no other contemporaneous variables are included. However, the inclusion of contemporaneous data from peer companies does not always contribute to the prediction performance when other contemporaneous variables are included. Our second research question looks at the contribution of peer data to the prediction performance when coordinated errors exist.[1] We find that when coordinated errors are present, models that incorporate peer data are better able to moderate the impact of these errors yielding marginally better prediction results. Our third research question examines the error detection performance when coordinated errors exist. Our results indicate that the inclusion of peer data helps considerably in detecting errors. Thus, while peer data does not always improve the prediction accuracy, its contribution to the error detection performance is substantial.

The major contributions of this study include, first, an investigation of the performance of industry-wide analytics, including analytics for rapidly changing industries. Previous studies have primarily documented the performance of analytical procedures within a narrowly defined scope, concentrating on single companies within stable industries. Secondly, this paper presents a new approach for peer-based information transfer within the context

[1] We define coordinated errors as errors/irregularities that exist in multiple related accounts.

of analytical procedures. Finally, this paper introduces a dynamic peer selection approach for the purpose of performing analytical procedures.

BACKGROUND, MOTIVATION, AND RESEARCH QUESTIONS

Many previous studies concentrate on various ways to improve the performance of analytical procedures. Several studies look at the ideal level of temporal aggregation for analytical procedures by performing analytics on annual, quarterly, and monthly observations. In general, prior research finds that the higher the frequency of the data used, the better the prediction and error detection results. Wild (1987) and Dzeng (1994) show that using monthly observations, rather than quarterly and annual data, results in improved account balance predictions. However, Wheeler and Pany (1990) suggest that using monthly data might result in inferior predictions due to the reduction in data reliability. They state that although quarterly data are not fully audited, the auditor does review them, whereas monthly data are not reviewed at all. In contrast, Chen and Leitch (1998) assert that the use of monthly data is more in line with the business cycle of many companies, and therefore month-to-month relations should yield better predictions.

Additional studies concentrate on using more sophisticated statistical techniques to generate more accurate and precise predictions. Dugan et al. (1985) propose the X-11 Model as an analytical procedure technique for auditors. This model was later used by Wheeler and Pany (1990) and by Chen and Leitch (1998), and was found not to be superior to multivariate regression models. Another study (Dzeng 1994) introduces Vector Autoregression (VAR) as a possible tool for performing analytical procedures and finds that VAR performed slightly better than a multivariate regression model. Structural equations were used as an analytical procedure tool by Wild (1987) and by Chen and Leitch (1998). Both studies conclude that the prediction performance of the structural model is not significantly better than that of multivariate stepwise regression models. More recently, Leitch and Chen (2003) examine the error detection performance of analytical procedures by looking at coordinated errors. Specifically, they study errors jointly by examining error patterns that are seeded into more than one account. In the current paper, we adopt a similar approach and examine the marginal contribution of peer data in improving the error detection capability for accounts with coordinated errors.

Data availability to auditors has historically been a hurdle for including more contemporaneous variables in expectation models. Therefore, auditors and researchers have traditionally used company specific contemporaneous data in addition to publicly available data but did not use contemporaneous data from peer companies. Using data from peer companies in similar industries requires auditors to have these data readily available prior to the audit. However, at the time of the audit, contemporaneous data points from peer companies are not publicly available. In this study, a simulated database repository[2] for industry auditing is constructed using industry-wide data. The data management approach is motivated by the fact that many Big 4 accounting firms audit multiple companies in each industry, and consequently have access to their proprietary information. These auditors could theoretically use data from one or more peer companies to improve predictions for another peer company. Thus, once the constraint of data availability is relaxed, and the use of information systems to facilitate data exchange is introduced, it becomes valuable to assess the potential contribution of contemporaneous peer information to the performance

[2] Database repository in this study refers to a central database system managed by a trusted third party. Account balance information from multiple companies that are audited by multiple CPA firms is transmitted and stored in that database.

of analytical procedures. An additional desirable characteristic of this approach is that the risk associated with the use of un-audited information (Wild and Biggs 1990) is somewhat mitigated by the fact that financial information is drawn from a different distribution and, therefore, is independent.[3] Specifically, developing expectations using firm-specific data as well as data from peer companies could help auditors in detecting errors when coordinated errors exist.

We examine three research questions concerning the relative benefits of using peer models in analytical procedures with respect to predictive performance and error detection. We also evaluate the impact of company specific characteristics and structural changes on the performance of analytical procedures. This is done given the hypothetical scenario that sharing information across audit firms is legal, or alternatively, that auditor-industry groups are highly concentrated. We discuss the legal constraints as well as the technology requirements in the concluding section of the paper.

A number of previous studies that concentrate on analytical procedures examine the use of multiple companies in similar industries (Allen 1992; AICPA 1988; Lev 1980; Wheeler and Pany 1990) as well as the use of multi-location data (Allen et al. 1999) to generate expectations. Other studies use lagged industry variables (Wheeler and Pany 1990) and contemporaneous industry indexes (Chen and Leitch 1998). Allen (1992) employs a joint expectation model with one set of coefficients for five electric utility companies and observes mixed results. He finds that pooled industry models are better at predicting expense accounts but inferior at predicting revenue accounts.

Additional evidence relating to the usefulness of industry data exists in the literature of industrial organization economics. A major argument in that literature is that the structural characteristics of industries are the primary determinants of firms' performance (Hawawini et al. 2003). Other studies within the accounting literature recognize that individual companies experience structural similarities. These studies examine the issue of accruals and abnormal accruals measurement under the assumption that companies should be matched based on industry characteristics (DeFond and Jiambalvo 1994), size within industries (Perry and Williams 1994), or industry and performance (Kothari et al. 2005). Thus, the recognition that both economic and industry factors impact companies is widely accepted but seldom used within auditing research.

Both financial accounting and auditing research have studied information transfer and auditor specialization across companies in similar industries. The financial accounting literature primarily looks at earnings announcements' effect on the stock prices of other firms in the same industry. Foster (1981) looks at intra-industry share price effects and earnings' estimates revisions on non-announcing firms. In the auditing literature, many studies (Taylor 2000; Kwon 1996; Wright and Wright 1997; Hogan and Jeter 1999) look at industry specialization and the effects of such specialization on audit performance, audit fees, and economies of scale among other factors. Thus, while a number of auditing papers examine the notion of knowledge transfer, our study concentrates on information transfer that is achieved through the transfer of raw data. Many of the Big 4 CPA firms have adopted an industry-auditing approach.[4] Using this approach, auditors share knowledge that they collect

[3] Earlier work by Ijiri and Leitch (1980) examined Stein's paradox (see also Efron and Morris 1973, 1977) in the context of audit sampling. This research suggests that auditors could potentially reduce their composite risk by using information and experience from multiple clients.

[4] Industry-auditing approach refers to the notion of sharing data, analysis, and experience across teams that work in similar industries.

in various audits in similar and interrelated industries but do not use proprietary data from peer companies in the context of analytical procedures.

As Lev (1980) states, "firms do not operate in a vacuum." All firms are affected by the same economy-wide factors such as inflation rates and changes in the fiscal policy. Additionally, firms that operate in similar industries are affected by similar industry-wide effects such as raw materials, wages, and energy costs. Therefore, when constructing a prediction model, it is important to consider economy and industry-wide effects as well as firm-specific effects. Our first research question examines whether models with peer data lead to different mean absolute percentage error (MAPE) in comparison to models that do not incorporate peer data. The inclusion of contemporaneous industry variables is one way to capture the well documented economic and industry-wide factors affecting companies' performance. Though the use of contemporaneous industry variables has only been utilized to a limited extent, past research suggests that the inclusion of such variables would contribute to the performance of analytical procedures. The use of peer models is expected to exploit cross sectional commonalities and better preserve the structural relationships that are generated in the estimation period through the holdout period.

> **Research Question 1:** Do models that incorporate peer data generate smaller mean absolute percentage errors?

It is often the case that contemporaneous account information is the best predictor of another account within the same company. Specifically, the correlation between accounts such as AR and Sales or COGS and AP is fairly high.[5] Therefore, it is possible that additional independent variables beyond the inclusion of contemporaneous accounts could add little to the prediction accuracy. Yet, the use of unaudited information from other accounts could potentially introduce an additional risk (Wild and Biggs 1990) that could be mitigated if additional financial information is drawn from a different distribution and, therefore, is independent.[6] This risk could arise when an account that is in error is used to predict another account. Therefore, in the second research question we examine the effect of using unaudited information from one account that is in error to predict another account.

> **Research Question 2:** Can peer data moderate the impact of materially misstated account balances on the prediction accuracy of related accounts?

The purpose of the third research question is to test the error detection performance when coordinated errors are present. If two accounts are in error (i.e., both sales and accounts receivable are overstated), then utilizing one account to predict the value of another account could potentially result in inferior error detection performance. The inclusion of contemporaneous peer data has the potential of moderating this effect and helping auditors identify errors more efficiently and effectively.

> **Research Question 3:** Are models that incorporate peer data better able to detect errors when coordinated errors exist?

---

[5] In our sample, the average Pearson correlation between *AR* and *SALES* and between *COGS* and *AP* across our industries is 0.939 and 0.844, respectively.

[6] We should note that using un-audited data from other companies introduces a different kind of risk.

METHODOLOGY

Data

For the purposes of this study, 18 industries that experienced various sales growth rates from 1991–2002 were selected. These industries were chosen due to their diverse growth characteristics and their membership in various economic sectors. For example, the selected industries include the Steel Works & Blast Furnaces industry (Standard Industrial Classification [SIC] 3312) that experienced on average 3.4 percent annual growth during the sample period, and the Pharmaceutical Preparations industry (SIC 2834) that experienced an average annual growth of 23.4 percent. Quarterly information for the total revenues, cost of revenues, accounts receivable, and accounts payable was extracted from the Compustat quarterly files for the period of 1991–2002. These accounts were chosen because of the common treatment that they often receive across companies and their participation in the two major business processes, the revenue, and the purchasing processes. Accounts such as inventory[7] would have presented additional data constraints and would have further limited our pool of peers for each audit client. To remain in the sample, firms had to have uninterrupted quarterly data for four years for each estimation as well as have year-to-year sales growth of no more than 500 percent. Our final sample includes 5,747 quarterly observations. The selected industries are presented together with their average sales growth in Table 1.

In order to simulate the data management system, we chose to group companies into dynamic peer clusters. Companies in the same industry group are likely to have many common characteristics. Therefore, knowledge collected from client X that is audited by firm i can potentially be used to perform analytical procedures for client Y that is audited by any auditor that participates in the knowledge sharing system. At this stage, our implementation presumes that all auditors share information regarding their clients, and therefore, the entire industry, as defined by the four-digit SIC code is examined.

Since even the most specific industry classification coding system is too general, the current four-digit SIC coding system is further partitioned and peer companies are dynamically chosen for each audit client. The process of identifying peer companies is done as follows: Within each four-digit SIC code, firms are ranked based on their sales (size proxy) and their sales growth (change proxy). This ranking is done using the last audited period of the estimation sample. Peers are selected for each company based on their size and growth proximity to that company at a given time. The iterative process of assigning peer companies for each audit client within an industry persists until peers are selected or the determination is made that there are no appropriate peers for the audit client. Using this approach, the peer selection process results in a relatively homogenous group of peers for each audit client. The current peer selection criteria are different from the traditional clustering approach in that while company A and C may be chosen as peers for company B, company C may not be the best peer for company A. An illustration of the peer selection process is described in Table 2 in which we demonstrate the process of assigning peers for each company in an industry of ten companies for a specific audit year. A company is assigned specific peers only if their size and growth rankings are both comparable to the rankings of that company. This process may result in companies with no peers, and those companies are subsequently dropped from the sample. While the companies examined may differ along many dimensions, including their products, geographic locations, and other economic factors, they are expected to share the same industry and economy-wide effects.

---

[7] Inventory and fixed assets accounts are often handled differently across companies.

TABLE 1
Descriptive Statistics Sample Companies from 1991–2002

Sic Code	Industry Name	Number of Firm-Years	Revenues	Revenue Growth	Cost of Sales	Accounts Receivable	Accounts Payable
1381	Drilling Oil And Gas Wells	132	100.79	0.27	65.54	75.41	30.12
2821	Plastics, Resins, Elastomers	98	774.35	0.09	520.94	637.74	296.02
2834	Pharmaceutical Preparations	391	774.06	0.23	239.69	461.39	224.32
2911	Petroleum Refining	250	4,236.01	0.11	3,373.71	1,718.86	1,719.36
3089	Plastics Products, Nec	155	124.91	0.12	84.12	78.67	38.11
3312	Steel Works & Blast Furnaces	192	405.61	0.03	354.58	177.54	144.92
3576	Computer Communication Equip	159	22.92	0.13	11.57	16.95	7.67
3661	Tele & Telegraph Apparatus	214	166.62	0.19	93.93	184.11	84.76
3663	Radio, TV Broadcast, Comm Eq	272	429.19	0.12	263.49	357.55	115.39
3674	Semiconductor, Related Device	327	173.15	0.20	90.29	106.75	60.19
3714	Motor Vehicle Part, Accessory	213	511.31	0.12	402.36	362.97	218.9
4911	Electric Services	562	566.7	0.10	393.86	275.69	192.15
5812	Eating Places	433	157.31	0.09	113.99	22.86	27.89
6021	National Commercial Banks	567	688.41	0.12	300.87	16,305.99	18,253.88
6022	State Commercial Banks	563	181.86	0.12	77.96	4,343.87	5,378.52
6331	Fire, Marine, Casualty Ins	403	641.44	0.13	574.93	1,771.26	694.61
7370	Cmp Programming, Data Process	262	973.22	0.16	549.95	1,105.7	633.86
7372	Prepackaged Software	554	94.67	0.17	24.3	65.13	21.48

This table presents descriptive statistics for 18 industries between the years 1991–2002. The mean account value for the total revenues, cost of sales, total assets, accounts receivable and accounts payable are broken by the four-digit SIC code. The mean sales growth for each four-digit SIC code is presented together with the number of firm quarter observations that met the data availability criteria.

TABLE 2
An Illustration of the Peer Selection Criteria

Company	Sales Rank	Sales Growth Rank	Selected Peers
A	1	8	C
B	2	3	C,D
C	3	5	B,D,E
D	4	2	B,C,E
E	5	4	C,D,G
F	6	9	G,H
G	7	7	E,F,H
H	8	6	G,F
I	9	2	J
J	10	1	I

Within each four-digit SIC code, companies are ranked by their total revenues and revenue growth accounts. The total size of the SIC group n represents the number of companies within each SIC code for a particular audit year. In the above table n = 10 for SIC $$$$ and year YYYY. The allowable proximity for each year is determined as follows: sales have to be within Integer (n/5) of each peer. In the above example, looking for peers for client E (sales rank = 5), potential peers will have a rank value between 3 and 7. Additionally, the sales growth rank was set to be Integer (n/4) allowing for more variation in the growth in comparison to the size proxy. Therefore, potential peers for client E will have sales growth rank between 1 and 7. Both criteria—size and growth have to be met for each audit client in order to be considered as peers. For example, in the above table peers for client E will be companies C, D, and G, whereas peers for client G will be E, F, and H. As illustrated, each client is assigned peer companies in such a way that the peer relation is not symmetric, and peer groups are therefore different from clusters. This process is repeated for each client in each industry group, and is performed iteratively for each audit year.

The peer selection method that we propose is relatively crude in comparison to the analysis that can be performed by practicing auditors.

Generating Monthly Observations

In previous studies, disaggregate monthly data performed better in analytical procedures than did quarterly data points (Wild 1987; Chen and Leitch 1998; Cogger 1981; Knechel 1988; Dzeng 1994). A common difficulty with time series analysis is the tradeoff between the need for sample size and the accuracy of the estimation models given the model's stability over long periods of time. Through using higher frequency data in the form of monthly observations, it is expected that the prediction accuracy will be better. Using quarterly data requires the use of data that represents information that spans over eight years. However, companies are operating in a dynamic environment, in which changes in technology, productivity, and labor cost constantly occur. The use of quarterly data may result in a well-specified model during the estimation period, but most likely, will result in inferior out-of-sample predictions. Therefore, using monthly observations seems to be an appropriate choice. Given that monthly data is not readily available for a large number of companies, a data interpolation technique was used in this study.

A cubic splines interpolation was introduced within the auditing literature by Chen and Leitch (1998) and Leitch and Chen (1999). In the current study, cubic splines are used to interpolate monthly observations through the use of publicly available quarterly observations. From each of the four quarterly observations, 12 monthly points are generated and later used as monthly data points. This process is performed for the 48 quarters from 1991–2002.[8]

Within the process of interpolating accounting data, it is important to distinguish between variables that are measured at points in time and variables that represent totals or averages over an interval. The algorithm used for interpolating income statement accounts must guarantee that the interpolated values sum up to the original value. In other words, the totals of the three months in each quarter are summed up to be equal to the quarterly value. In the mathematical equation, the weights are the coefficients of the cubic polynomials used to interpolate the data. The coefficients define the line so that it passes through each of the data points in a smooth way. The basic third degree polynomial is defined in Equation (1) as follows:

$$S_i(X) = a_i(X - X_i)^3 + b_i(X - X_i)^2 + c_i(X - X_i) + d_i. \qquad (1)$$

Equation (1) can be used in its basic form to interpolate accounts that are stocks (balance sheet accounts); however, it needs to be slightly modified for the purpose of interpolating flows (income statement accounts). This is done by constraining the three monthly observations in the income statement accounts to sum up to the quarterly value.

Model Specification

As in most studies, we first use our models to predict the monthly account balances and subsequently test the ability of these models to detect errors. The research questions are tested by comparing the performance of peer-based models with a benchmark model. We compare the benchmark models in which expectations are derived using archival data and firm-specific contemporaneous data with prediction models that incorporate contemporaneous peer data.

The peer model specification examines the commonalities between a given company and its group of peer companies by including a standardized industry/peer average as an independent variable. To avoid the impact of company size on the peer average, a standard score is calculated for each company as follows $Z = \frac{Y - \mu y}{\sigma y}$, where Y represents a monthly account balance, and the mean and the standard deviation of Y are calculated during the estimation period and are used to calculate the standard score for each data point during the estimation and holdout periods. The advantage of this approach stems from the fact that on average companies behave similarly to their peers. In other words, companies' accounts will, on average, experience similar changes and will therefore contribute to the prediction performance for that company. Conversely, the assumption that the audit client will typically experience similar changes to all of its peers is somewhat constraining because this approach results in the loss of information due to the aggregation of peers' data.[9]

We present our models in Table 3 in which peer models are estimated as depicted in Models 2, 3, 4, and 5 and are compared to the benchmark Models 6, 7, 8, and 9 respectively. *SALES, COGS, AR, and AP* represent total revenue, cost of goods sold, accounts receivable and accounts payable balances for month t. The *IND* term in the peer models represents

[8] To avoid issues with boundary conditions we actually use data from 1990–2003 for the interpolation process and drop the first and the last years of data.

[9] We also used an alternative peer model specification per Allen et al. (1999) in which the raw contemporaneous account balances from peer companies are included as independent variables. This approach allowed for different relationships between audit client and each peer but often led to inferior predictions and error detection performance.

TABLE 3
Specification of Models

$$IND_t = \frac{\sum_{1}^{i} Z_t}{i}$$

$$SALES_t = \alpha + \beta_1 SALES_{t-12} + \beta_2 IND_t + \varepsilon_t \tag{2}$$
$$SALES_t = \alpha + \beta_1 SALES_{t-12} + \beta_2 IND_t + \beta_3 AR_t + \varepsilon_t \tag{3}$$
$$COGS_t = \alpha + \beta_1 COGS_{t-12} + \beta_2 IND_t + \varepsilon_t \tag{4}$$
$$COGS_t = \alpha + \beta_1 COGS_{t-12} + \beta_2 IND_t + \beta_3 AP_t + \varepsilon_t \tag{5}$$
$$SALES_t = \alpha + \beta_1 SALES_{t-12} + \varepsilon_t \tag{6}$$
$$SALES_t = \alpha + \beta_1 SALES_{t-12} + \beta_2 AR_t + \varepsilon_t \tag{7}$$
$$COGS_t = \alpha + \beta_1 COGS_{t-12} + \varepsilon_t \tag{8}$$
$$COGS_t = \alpha + \beta_1 COGS_{t-12} + \beta_2 AP_t + \varepsilon_t \tag{9}$$
$$MAPE_t = \alpha + \beta_1 CHANGE_t + \beta_2 SIZE_t + \varepsilon_t \tag{10}$$
$$ERROR_t = \alpha + \beta_1 CHANGE_t + \beta_2 SIZE_t + \varepsilon_t \tag{11}$$
$$MAPE_t - DIFF_t = \alpha + \beta_1 CHANGE_t + \beta_2 SIZE_t + \varepsilon_t \tag{12}$$
$$ERROR_t - DIFF_t = \alpha + \beta_1 CHANGE_t + \beta_2 SIZE_t + \varepsilon_t \tag{13}$$

SALES, COGS, AR, and AP represent total revenue, cost of goods sold, accounts receivable and accounts payable balances for month *t*. The *IND* term in the peer models represents the average standard score for a group of peers and is calculated as depicted above. *MAPE* is the Mean Absolute Percentage Error for each company year and *ERROR* is the sum of the monthly Type I or Type II errors for each company year. We define *CHANGE* by using either the absolute average change in sales (Compustat quarterly data item 2) or *EPS* (Compustat quarterly data item 19) of a company during each estimation period. *SIZE* is measured as the log of the quarterly total assets (Compustat quarterly data item 44). The dependent variable in Model 12, *MAPE-DIFF*, is the difference in the *MAPE* between the benchmark model and the peer model. Similarly, the dependent variable in Model 13, *ERROR-DIFF*, is the difference in the number of Type I or Type II errors between the benchmark model and the peer model.

the average standard score for a group of peers and is calculated as described above. In all the models, we use a 12-month lag term as an independent variable.

Test of Research Questions

Each regression model is estimated over 36 months or three years and is tested over the subsequent 12 months. Every model is estimated separately for each company based on its unique set of peer companies. The selection of 36 months as the training period and 12 months as the holdout period is similar to the research design employed by many previous studies.

The peer regression models are estimated separately for each company based on its unique set of peers. The peer selection algorithm matches peers for each company in each year throughout our sample period. For example, for the purpose of predicting account balances for 1994, peers are identified based on data from the last quarter of 1993, whereas data from 1991–1993 is used to generate predictions. In that manner, the process of selecting peers and estimating the models is done separately nine times for each company for the years 1994–2002. Subsequently, we generate 12 monthly predictions for each company in our sample for each year-account from 1994–2002. A total of 12 monthly predictions are created for each estimation year for each account totaling 216 predicted observations for each company-model.[10] Figure 1 illustrates the steps that we follow in generating our test data.

Prediction performance is evaluated based on examining the mean absolute percentage error[11] (MAPE) for each account-model. The MAPE is calculated for the out-of-sample prediction for each account-company-month. The MAPEs for the 12-month period are aggregated over company-year resulting in an aggregated measure of MAPE for each company-account-model. To evaluate the prediction performance of each model, results are aggregated over each account-industry, resulting in one MAPE for every account-model-industry. For example, each model is estimated separately for every company with its own unique set of peers. For each company, a forecast for each account balance (such as revenues) is generated for every month. An average of the MAPE is calculated over the entire industry and is tested separately. To test whether the results generated by the peer model are superior to the benchmark model, a Wilcoxon Rank-Sum test is performed separately for each industry over the prediction period of nine years. To assess the goodness of fit of the models, the adjusted R^2 statistics for each company model is calculated and aggregated to an industry average.

The prediction performance is evaluated for Models 2 through 9. In each case, we estimate these models using independent and dependent variables that are not in error. However, for a subset of our models, we do test the impact of the existence of errors in the independent variables on the prediction accuracy. We therefore seed material errors into the AR and AP accounts in Models 3, 5, 7, and 9 for the prediction year and use the estimated coefficients[12] for calculating our expectations. The materiality definition that we use is similar to the one used by Knechel (1988) and is set to be equal to 2 percent of the account value. We subsequently use MAPE to examine the impact of using an account balance with material error to predict another account balance.

The third research question is examined by seeding errors into account balances and evaluating the error detection performance of Models 3 and 5 in comparison to Models 7 and 9. Similarly to the approach adopted by Leitch and Chen (2003), we seed errors into more than one account and evaluate which model can better identify these errors. We concentrate on the following coordinated errors:

(1) Overstatement of AR and Total Revenues.
(2) Understatement of AP and Cost of Revenues.

Errors in these accounts are of particular interest because these accounts are often targeted for fraud and manipulation.[13] We first estimate Models 3, 5, 7, and 9 using independent and dependent variables with no errors. However, in order to test the third research question we seed errors into firm-specific independent variables (i.e., seed errors into AR in Model 3 during the prediction year) and use the estimated coefficients to calculate the

---

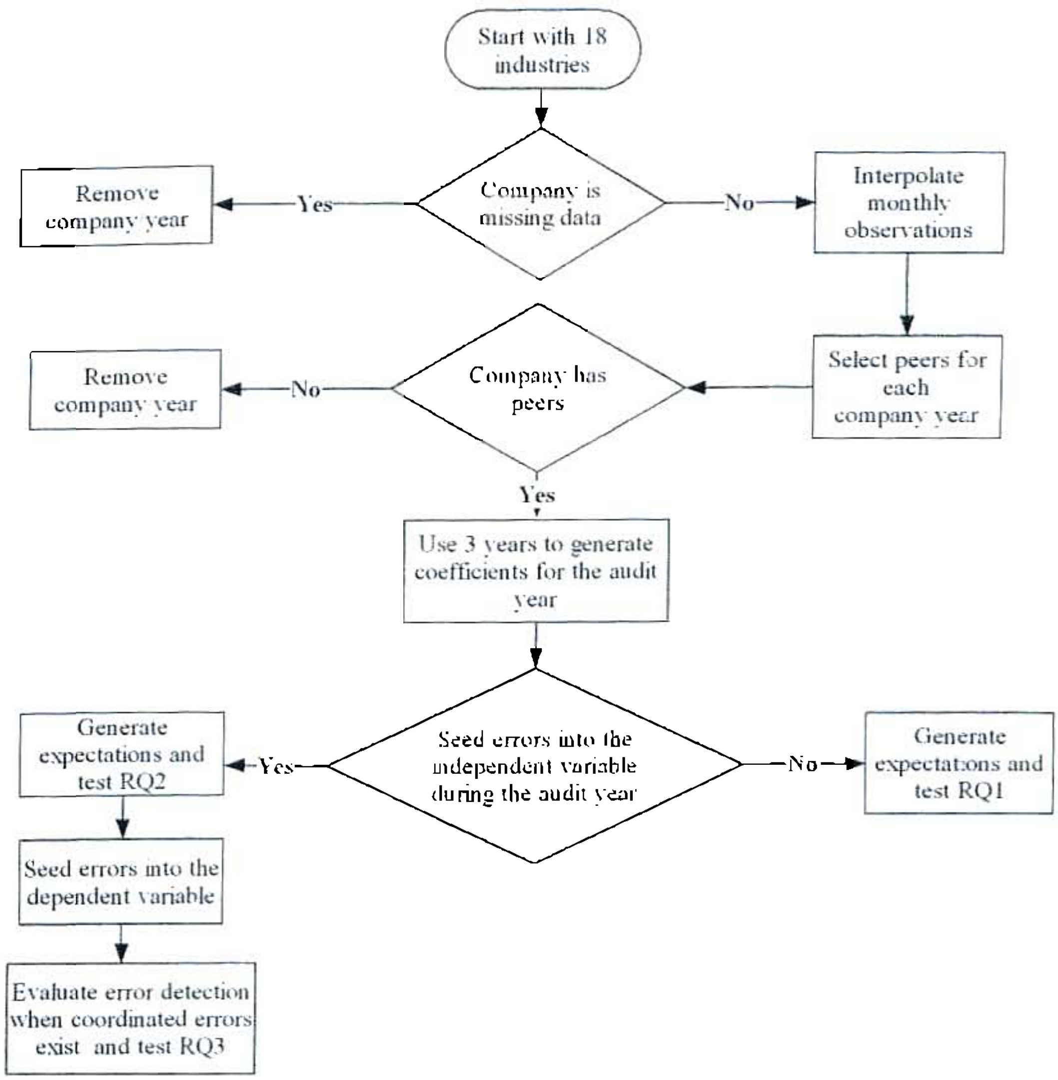
[10] The number of monthly predictions can vary based on data availability for each company.

[11] $MAPE = \frac{1}{i^{*}12}\sum_{i=1}^{i}\sum_{j=1}^{12}\left|\frac{P_{ij} - A_{ij}}{A_{ij}}\right|$, where P represents the predicted value, A represents the actual value, i is the number of companies in each industry and j is the number of months predicted.

[12] These coefficients are generated based on data points that are not in error.

[13] A study by the Government Accountability Office (GAO 2002) found that the reason for 37.9 percent and 15.7 percent of nontechnical restatements between 1997 and June 2002 involve the revenue and expense accounts, respectively.

FIGURE 1
Illustration of the Sample Selection and the Test Sample



forecasted account balance of the dependent variable. Following this approach, the predicted value is generated using data that is in error. Subsequently, we seed errors into the actual dependent variable and evaluate the performance of our models in detecting the existence/absence of errors in that account. To perform this test we first calculate the confidence interval around the estimated value and use a statistical investigation rule to evaluate the error detection performance of the peer model and the benchmark model.[14]

[14] Each confidence interval is calculated separately for each account. This is different from Leitch and Chen (2003) who generate joint bivariate confidence intervals for two accounts.

The statistical investigation rule signals when the standardized difference between the actual and the predicted account balance exceeds the critical Z-value that is based on the auditor's specified risk level α. In the current study $\alpha = 0.05$ and $\alpha = 0.33$ were used.[15] An investigation would take place if $(Y_t - \hat{Y}_t)/S_y > Z_{1-\alpha}$ where S_y is the base period standard deviation of the series, Y_t. If the statistical investigation rule signaled that an error was present when no error was seeded into the account balance, than a Type I error occurred. If the investigation rule failed to signal that a material error was present when an account balance was seeded with error, than a Type II error occurred.

RESULTS

In the first research question, we conjecture that account balance expectations that are generated using peer models should result in more accurate predictions than expectations derived from company specific models. For that purpose, we evaluate the predictions from Models 2, 4, 6, and 8 that estimate the balance of the revenues series and the cost of sales series. Prediction performance is evaluated by using a nonparametric test to evaluate the differences between the mean absolute percentage errors of the peer and the benchmark models. Panels A and B in Table 4 contain the results for the test of the first research question for the total revenue and the cost of revenue series. As displayed in Table 4, results suggest that peer data improves the prediction accuracy when such data is added to relatively simple models. Panel A presents the aggregate MAPE for each one of the 18 industries using predictions from Models 2 and 6. In Panel A of Table 4, we observe that, when comparing peer models to benchmark models, 14 of the 18 industries experience prediction improvements out of which eight differences are significant. Panel B of Table 4 displays the prediction results for the cost of sales series. Using Models 4 and 8, we observe that 11 industries experience more accurate predictions, five of which are statistically significant. However, the benchmark model is statistically superior to the peer model for four of the 18 industries.

Table 5 displays the results for the second research question based on predictions from Models 3, 5, 7, and 9. In the left side of Panel A we first display the prediction results for the revenue series using Models 3 and 7. Using these models, we observe that in comparison to the results of the revenue series in Panel A of Table 4 fewer industries experience statistically significant improvement in prediction accuracy. In fact, the benchmark model yields better results for a larger number of industries. This suggests that the contribution of the peer data is diluted when contemporaneous firm-specific data is included in the prediction model. However, because of the inherent risk associated with using independent variables from the same distribution, we examine the prediction performance of Models 3 and 7 when the contemporaneous independent variable is materially misstated during the holdout period. In the right side of Panel A we observe that the prediction performance of peer models relative to the benchmark model improves when AR has a material error.[16]

Panel B of Table 5 presents the prediction performance for the cost of sales series using Models 5 and 9 and depicts a surprisingly similar pattern to the results in Panel B of Table 4. The same five industries experience a statistically significant prediction improvement in comparison to the peer model. However, we observe that many industries experience superior prediction performance using the benchmark model rather than the peer model. In

[15] We also used $\alpha = 0.10$ and obtained similar results.

[16] We have assessed the potential presence of multicollinearity by examining the significance of our independent variables and calculating the variance inflation factors (VIF) for a subset of our regressions. None of the calculated VIF values exceeded 2.4 indicating that if multicollinearity is present, it should not impact our results.

TABLE 4

SIC-Aggregated Prediction Performance for the Total Revenue and Cost of Sales Series (Based on Models 2, 4, 6, and 8)

		Panel A: Prediction and Goodness of Fit Results of the *Total Revenue* Series Based on Models 2 and 6						Panel B: Prediction and Goodness of Fit Results of the *Cost of Sales* Series Based on Models 4 and 8					
SIC Code	Number of Firm-Years	Peer-Model *MAPE*	*MAPE*	Left-Sided Wilcoxon Rank-Sum Test	Right-Sided Wilcoxon Rank-Sum Test	Peer-Model ADJRSQ	ADJRSQ	Peer-Model *MAPE*	*MAPE*	Left-Sided Wilcoxon Rank-Sum Test	Right-Sided Wilcoxon Rank-Sum Test	Peer-Model ADJRSQ	ADJRSQ
1381	132	0.240	0.290	0.000	—	0.602	0.297	0.245	0.265	0.000	—	0.500	0.274
2821	98	0.159	0.162	—	0.333	0.459	0.237	0.182	0.166	—	0.058	0.402	0.194
2834	391	0.177	0.180	0.121	—	0.551	0.438	0.199	0.187	—	0.007	0.491	0.361
2911	250	0.159	0.212	0.000	—	0.660	0.174	0.178	0.231	0.000	—	0.630	0.151
3089	155	0.162	0.161	—	0.108	0.607	0.510	0.144	0.145	0.240	—	0.555	0.441
3312	192	0.143	0.150	0.013	—	0.431	0.217	0.132	0.130	0.301	—	0.418	0.222
3576	159	0.230	0.232	0.307	—	0.469	0.325	0.242	0.247	0.284	—	0.423	0.304
3661	214	0.208	0.201	—	0.158	0.562	0.430	0.231	0.223	—	0.294	0.543	0.399
3663	272	0.227	0.230	0.146	—	0.476	0.346	0.243	0.236	—	0.092	0.449	0.313
3674	327	0.219	0.234	0.000	—	0.584	0.365	0.230	0.236	0.227	—	0.524	0.345
3714	213	0.170	0.170	0.446	—	0.601	0.419	0.165	0.167	0.166	—	0.578	0.404
4911	562	0.121	0.130	0.000	—	0.624	0.500	0.142	0.144	0.001	—	0.437	0.280
5812	433	0.121	0.124	0.429	—	0.663	0.593	0.132	0.135	0.291	—	0.626	0.539
6021	567	0.127	0.137	0.000	—	0.686	0.472	0.187	0.241	0.000	—	0.657	0.328
6022	563	0.113	0.132	0.000	—	0.711	0.502	0.170	0.235	0.000	—	0.665	0.357
6331	403	0.129	0.133	0.027	—	0.522	0.357	0.172	0.164	—	0.044	0.440	0.292
7370	262	0.213	0.212	—	0.492	0.609	0.458	0.232	0.240	—	0.437	0.516	0.375
7372	554	0.211	0.213	0.137	—	0.588	0.494	0.260	0.255	—	0.105	0.483	0.368

This table displays the prediction results for the total revenue and cost of sales series. The table displays the number of firm years in each SIC group, the MAPEs from the peer models (that are based on Models 2 ($SALES_t = \alpha + \beta_1 SALES_{t-12} + \beta_1 IND_t + \varepsilon_t$) and 4 ($COGS_t = \alpha + \beta_1 COGS_{t-12} + \beta_2 IND_t + \varepsilon_t$)) and from the benchmark models (Models 6 ($SALES_t = \alpha + \beta_1 SALES_{t-12} + \varepsilon_t$) and 8 ($COGS_t = \alpha + \beta_1 COGS_{t-12} + \varepsilon_t$)), two one-sided Wilcoxon rank-sum tests probabilities for the differences in the monthly *MAPE* values and the average Adj-R^2 for each model. The left side probabilities are for testing the hypothesis that *MAPE*s from the peer models are lower than the *MAPE*s from the benchmark models while the right side probabilities are for testing the hypothesis that the benchmark models perform better.

TABLE 5
SIC-Aggregated Prediction Performance

Panel A: Total Revenue Series (Based on Models 3 and 7)[a]

		Prediction Results Based on Models 3 and 7				Prediction Results Based on Models 3 and 7 *with Material Error Seeded into AR*				Goodness of Fit Models 3 and 7	
SIC Code	Number of Firm-Years	Peer-Model *MAPE*	*MAPE*	Left-Sided Wilcoxon Rank-Sum Test	Right-Sided Wilcoxon Rank-Sum Test	Peer-Model *MAPE*	*MAPE*	Left-Sided Wilcoxon Rank-Sum Test	Right-Sided Wilcoxon Rank-Sum Test	Peer-Model ADJRSQ	ADJRSQ
1381	132	0.182	0.179	—	0.482	0.207	0.199	—	0.171	0.754	0.660
2821	98	0.137	0.137	0.231	—	0.154	0.156	0.299	—	0.616	0.517
2834	391	0.166	0.159	—	**0.035**	0.177	0.175	—	0.110	0.671	0.611
2911	250	0.143	0.155	**0.000**	—	0.170	0.199	**0.000**	—	0.744	0.516
3089	155	0.126	0.128	—	0.235	0.150	0.147	—	**0.057**	0.759	0.706
3312	192	0.110	0.112	0.324	—	0.141	0.144	0.366	—	0.655	0.539
3576	159	0.187	0.168	—	**0.000**	0.200	0.187	—	**0.019**	0.687	0.625
3661	214	0.184	0.170	—	**0.018**	0.194	0.187	—	0.205	0.719	0.653
3663	272	0.201	0.186	—	**0.001**	0.213	0.199	—	**0.005**	0.663	0.612
3674	327	0.182	0.178	—	0.187	0.198	0.196	—	0.175	0.750	0.661
3714	213	0.140	0.138	—	0.313	0.162	0.161	—	0.334	0.737	0.636
4911	562	0.117	0.122	**0.000**	—	0.134	0.143	**0.000**	—	0.718	0.636
5812	433	0.123	0.125	0.351	—	0.196	0.208	**0.026**	—	0.714	0.661
6021	567	0.117	0.118	0.137	—	0.117	0.118	0.116	—	0.770	0.673
6022	563	0.106	0.115	**0.000**	—	0.106	0.116	**0.000**	—	0.777	0.683
6331	403	0.132	0.134	0.201	—	0.135	0.140	**0.033**	—	0.614	0.516
7370	262	0.177	0.167	—	**0.011**	0.191	0.185	—	0.119	0.739	0.686
7372	554	0.173	0.167	—	**0.074**	0.181	0.180	0.463	—	0.733	0.692

(*continued on next page*)

TABLE 5 (continued)

Panel B: Cost of Sales Series (Based on Models 5 and 9)[b]

		Prediction Results Based on Models 5 and 9				Prediction Results Based on Models 5 and 9 *with Material Error Seeded into AP*				Goodness of Fit Models 5 and 9	
SIC Code	Number of Firm-Years	Peer-Model *MAPE*	*MAPE*	Left-Sided Wilcoxon Rank-Sum Test	Right-Sided Wilcoxon Rank-Sum Test	Peer-Model *MAPE*	*MAPE*	Left-Sided Wilcoxon Rank-Sum Test	Right-Sided Wilcoxon Rank-Sum Test	Peer-Model ADJRSQ	ADJRSQ
1381	132	0.215	0.225	**0.044**	—	0.232	0.239	0.224	—	0.500	0.274
2821	98	0.171	0.154	—	**0.012**	0.172	0.153	—	**0.005**	0.402	0.194
2834	391	0.191	0.177	—	**0.000**	0.198	0.185	—	**0.003**	0.491	0.361
2911	250	0.161	0.178	**0.000**	—	0.175	0.200	0.000	—	0.630	0.151
3089	155	0.127	0.122	—	0.466	0.140	0.137	—	0.420	0.555	0.441
3312	192	0.121	0.114	—	0.113	0.136	0.134	—	0.241	0.418	0.222
3576	159	0.230	0.236	0.491	—	0.236	0.245	0.285	—	0.423	0.304
3661	214	0.222	0.199	—	**0.000**	0.227	0.208	—	**0.013**	0.543	0.399
3663	272	0.235	0.216	—	**0.000**	0.245	0.231	—	**0.006**	0.449	0.313
3674	327	0.221	0.211	—	**0.015**	0.223	0.216	—	**0.038**	0.524	0.345
3714	213	0.151	0.139	—	**0.012**	0.170	0.161	—	0.099	0.578	0.404
4911	562	0.136	0.138	**0.003**	—	0.148	0.148	0.145	—	0.437	0.280
5812	433	0.131	0.130	—	0.220	0.152	0.157	0.091	—	0.626	0.539
6021	567	0.186	0.226	**0.000**	—	0.186	0.227	0.000	—	0.657	0.328
6022	563	0.161	0.219	**0.000**	—	0.161	0.220	0.000	—	0.665	0.357
6331	403	0.184	0.178	—	0.127	0.198	0.196	0.370	—	0.440	0.292
7370	262	0.217	0.221	—	0.485	0.227	0.240	0.094	—	0.516	0.375
7372	554	0.259	0.252	—	**0.008**	0.261	0.254	—	**0.010**	0.483	0.368

(*continued on next page*)

TABLE 5 (continued)

[a] Panel A displays the prediction results for the total revenue series. The table displays the number of firm years in each SIC group, the *MAPE* from the peer model (that is based on Models 3 ($SALES_t = \alpha + \beta_1 SALES_{t-12} + \beta_2 IND_t + \beta_3 AR_t + \varepsilon_t$)) and from the benchmark model (based on Model 7 ($SALES_t = \alpha + \beta_1 SALES_{t-12} + \beta_2 AR_t + \varepsilon_t$)). The table also displays the *MAPE*s that are based on Models 3 and 7 with material error seeded into the AR account. Additionally, the results for two one-sided Wilcoxon rank-sum tests probabilities for the differences in the monthly *MAPE* values (between Models 3 and 7) are presented and the average Adj-R^2 for each model is displayed in the last two columns. The left side probabilities are for testing the hypothesis that *MAPE*s from the peer models are lower than the *MAPE*s from the benchmark models, while the right side probabilities are for testing the hypothesis that the benchmark models perform better. Significant results are displayed in bold.

[b] Panel B displays the prediction results for the cost of sales series. The table displays the number of firm years in each SIC group, the *MAPE* from the peer model (that is based on Model 5 ($COGS_t = \alpha + \beta_1 COGS_{t-12} + \beta_2 IND_t + \beta_3 AP_t + \varepsilon_t$)) and from the benchmark model (based on Model 9 ($COGS_t = \alpha + \beta_1 COGS_{t-12} + \beta_2 AP_t + \varepsilon_t$)). The table also displays the *MAPE*s that are based on Models 5 and 9 with material error seeded into the AP account. Additionally, the results for two one-sided Wilcoxon rank-sum tests probabilities for the differences in the monthly *MAPE* values (between Models 5 and 9) are presented, and the average Adj-R^2 for each model is displayed in the last two columns. The left side probabilities are for testing the hypothesis that *MAPE*s from the peer models are lower than the *MAPE*s from the benchmark models, while the right side probabilities are for testing the hypothesis that the benchmark models perform better. Significant results are displayed in bold.

the right side of Panel B, we examine the impact of an error in the contemporaneous firm-specific independent variable on the prediction accuracy and observe that peer models somewhat mitigate the impact of the independent variable that is in error on the overall prediction accuracy.[17]

In aggregate, research questions 1 and 2 establish that the prediction performance as measured by the MAPE is generally better when more variables are included in the prediction model. We first observe that peer data contributes significantly to the prediction performance in Table 4 when the benchmark model relies primarily on historical data. When we include additional independent contemporaneous predictors such as AR and AP, we observe that the overall MAPEs in Table 5 are lower than the MAPEs in Table 4 indicating that the out-of-sample predictions benefit from the inclusion of additional independent contemporaneous variables. We also observe in Table 5 that when the independent predictors are materially incorrect, our predictions become less accurate. We also find that in the better-specified models the contribution of the peer data to the prediction accuracy is diluted. Yet, we find that if these contemporaneous variables contain errors, it becomes valuable to include peer data in the prediction models. These results are consistent with the fact that peer data are imperfect substitutions for contemporaneous firm-specific variables when such variables are not in error. Nonetheless, they emphasize that the risk from using account balances with coordinated errors could be reduced by using peer data.

Panels A and B of Table 6 present the results for the third research question examining the error detection performance using simulated errors. We evaluate the error detection using predictions from Models 3 and 7 (5 and 9 for the cost of sales) when coordinated errors are present. Table 6 presents the results for simulated coordinated errors seeded into more than one account simultaneously, i.e., an account that is in error is used to predict another account that is also in error. In Panel A (Panel B) of Table 6 both the revenue (*COGS*) and *AR* (*AP*) accounts are seeded with material errors. These tables present the percentages of Type I and Type II errors that occur when a statistical investigation rule is used. The error detection performance is evaluated using different risk levels to show the tradeoffs between Type I and Type II errors under each risk level. Generally, lower α levels lead to wider confidence intervals and consequently lead to fewer Type I errors and a larger number of Type II errors.

Both Panels in Table 6 indicate that peer models are often superior in preventing Type II errors but inferior in preventing Type I errors. These results indicate that due to the tighter confidence interval of the peer models, an auditor using the peer model would experience more frequent Type I errors, requiring the auditor to perform additional inquiries resulting in more expensive and less efficient audits. On the other hand, the peer model outperforms the benchmark model in detecting material deviations yielding lower Type II errors. Potentially, more precise predictions coupled with tighter confidence intervals contributed to these results. Hence, peer models are better at signaling the presence of material errors and are inferior at signaling the absence of material errors. However, since auditors are often more concerned with Type II errors than with Type I errors, examining the tradeoffs between the two is important.

Assessing the error detection performance can be easily achieved when a particular model leads to similar results for both error types. Thus, if the peer model leads to a smaller

[17] A supplementary analysis of the goodness of fit of the models in Tables 4 and 5 shows that in all cases, the adjusted R^2 is significantly higher for the peer models in comparison to the benchmark models. This implies that peer models are consistently superior to the benchmark models during the estimation period but not always during the hold-out period.

TABLE 6
Type I and Type II Error Rates

Panel A: Total Revenue Series (Based on Models 3 and 7)

SIC Code	Number of Firm-Years	Percentage of Errors Alpha = 0.05		Peer Model Percentage of Errors Alpha = 0.05		Benchmark Minus Peer	Benchmark Minus Peer	Type II Divided by Type I	Superior Model
		Type I	Type II	Type I	Type II	Difference (Type I)	Difference (Type II)	Difference (Type II) / Difference (Type I)	
1381	132	38.19%	44.48%	41.21%	46.73%	−3.02%	−2.25%	0.75	Benchmark
2821	98	33.36%	45.00%	35.42%	46.54%	−2.06%	−1.54%	0.75	Benchmark
2834	391	36.77%	37.56%	39.91%	38.22%	−3.14%	−0.66%	0.21	Benchmark
2911	250	43.13%	50.63%	44.09%	35.90%	−0.96%	14.73%	−15.34	Peer Model**
3089	155	35.70%	37.40%	37.89%	38.33%	−2.19%	−0.93%	0.42	Benchmark
3312	192	35.19%	47.13%	38.76%	40.94%	−3.57%	6.19%	−1.73	Peer Model**
3576	159	32.24%	56.06%	36.58%	51.19%	−4.34%	4.87%	−1.12	Peer Model**
3661	214	37.35%	45.93%	39.66%	43.70%	−2.31%	2.23%	−0.97	Peer Model**
3663	272	34.87%	50.46%	36.12%	48.74%	−1.25%	1.72%	−1.38	Peer Model**
3674	327	40.53%	44.02%	44.69%	40.48%	−4.16%	3.54%	−0.85	Peer Model**
3714	213	35.36%	44.99%	39.19%	40.83%	−3.83%	4.16%	−1.09	Peer Model**
4911	562	33.48%	33.65%	35.86%	28.56%	−2.38%	5.09%	−2.14	Peer Model**
5812	433	37.69%	39.96%	36.67%	39.43%	1.02%	0.53%	0.52	Peer Model
6021	567	35.98%	21.54%	39.03%	20.02%	−3.05%	1.52%	−0.50	Peer Model**
6022	563	38.31%	20.68%	37.73%	17.87%	0.58%	2.81%	4.84	Peer Model
6331	403	32.70%	25.39%	34.37%	25.97%	−1.67%	−0.58%	0.35	Benchmark
7370	262	39.49%	47.55%	41.63%	41.99%	−2.14%	5.56%	−2.60	Peer Model**
7372	554	36.30%	46.48%	36.81%	43.23%	−0.51%	3.25%	−6.37	Peer Model**

(continued on next page)

TABLE 6 (continued)

Panel A: Total Revenue Series (Based on Models 3 and 7) (continued)

SIC Code	Number of Firm-Years	Percentage of Errors Alpha = 0.05 Type I	Type II	Peer Model Percentage of Errors Alpha = 0.05 Type I	Type II	Benchmark Minus Peer Difference (Type I)	Benchmark Minus Peer Difference (Type II)	Type II Divided by Type I Difference (Type II) / Difference (Type I)	Superior Model
1381	132	63.67%	23.56%	67.65%	25.29%	−3.98%	−1.73%	0.43	Benchmark
2821	98	57.83%	24.38%	60.05%	24.72%	−2.22%	−0.34%	0.15	Benchmark
2834	391	63.16%	20.00%	65.55%	21.71%	−2.39%	−1.71%	0.72	Benchmark
2911	250	65.70%	25.67%	64.00%	17.42%	1.70%	8.25%	4.85	Peer Model
3089	155	61.95%	16.29%	64.63%	18.70%	−2.68%	−2.41%	0.90	Benchmark
3312	192	61.72%	23.13%	65.11%	21.04%	−3.39%	2.09%	−0.62	Peer Model**
3576	159	59.19%	30.65%	60.72%	26.52%	−1.53%	4.13%	−2.70	Peer Model**
3661	214	62.14%	24.04%	64.76%	23.38%	−2.62%	0.66%	−0.25	**
3663	272	60.38%	28.56%	62.45%	27.89%	−2.07%	0.67%	−0.32	**
3674	327	67.95%	24.62%	70.23%	23.56%	−2.28%	1.06%	−0.46	**
3714	213	60.73%	25.17%	65.64%	21.25%	−4.91%	3.92%	−0.80	Peer Model**
4911	562	60.25%	15.05%	62.19%	14.36%	−1.94%	0.69%	−0.36	**
5812	433	64.06%	22.10%	64.87%	22.34%	−0.81%	−0.24%	0.30	Peer Model
6021	567	60.71%	11.22%	63.70%	10.08%	−2.99%	1.14%	−0.38	**
6022	563	64.54%	10.36%	63.80%	8.79%	0.74%	1.57%	2.12	Peer Model
6331	403	59.76%	13.15%	60.13%	14.16%	−0.37%	−1.01%	2.73	Benchmark
7370	262	66.10%	27.20%	68.40%	22.20%	−2.30%	5.00%	−2.17	Peer Model**
7372	554	61.78%	24.51%	63.64%	22.88%	−1.86%	1.63%	−0.88	Peer Model**

(continued on next page)

TABLE 6 (continued)

Panel B: Cost of Sales Series (Based on Models 5 and 9)

SIC Code	Number of Firm-Years	Percentage of Errors Alpha = 0.05		Peer Model Percentage of Errors Alpha = 0.05		Benchmark Minus Peer	Benchmark Minus Peer	Type II Divided by Type I	Superior Model
		Type I	Type II	Type I	Type II	Difference (Type I)	Difference (Type II)	Difference (Type II) / Difference (Type I)	
1381	132	38.48%	42.60%	40.97%	34.14%	−2.49%	8.46%	−3.40	Peer Model**
2821	98	24.98%	39.08%	30.85%	32.52%	−5.88%	6.56%	−1.12	Peer Model**
2834	391	28.35%	42.86%	33.68%	40.01%	−5.34%	2.85%	−0.53	Peer Model**
2911	250	37.80%	34.11%	41.33%	21.32%	−3.53%	12.79%	−3.62	Peer Model**
3089	155	27.73%	26.23%	28.97%	24.74%	−1.25%	1.50%	−1.20	Peer Model**
3312	192	27.19%	22.18%	31.42%	22.81%	−4.24%	−0.64%	0.15	Benchmark
3576	159	30.23%	42.53%	31.67%	43.33%	−1.44%	−0.80%	0.56	Benchmark
3661	214	33.27%	39.35%	38.63%	37.33%	−5.36%	2.03%	−0.38	**
3663	272	32.94%	45.69%	33.55%	45.15%	−0.61%	0.54%	−0.88	Peer Model**
3674	327	34.83%	43.77%	38.38%	39.46%	−3.55%	4.31%	−1.21	Peer Model**
3714	213	33.95%	25.21%	39.41%	21.32%	−5.46%	3.89%	−0.71	Peer Model**
4911	562	26.36%	30.13%	30.13%	25.84%	−3.77%	4.29%	−1.14	Peer Model**
5812	433	32.99%	19.62%	34.36%	19.57%	−1.38%	0.04%	−0.03	**
6021	567	40.78%	37.80%	41.56%	29.61%	−0.78%	8.19%	−10.45	Peer Model**
6022	563	42.86%	36.22%	36.38%	30.56%	6.48%	5.66%	0.87	Peer Model
6331	403	32.99%	39.98%	34.82%	34.03%	−1.83%	5.95%	−3.25	Peer Model**
7370	262	33.33%	35.01%	35.51%	37.14%	−2.18%	−2.14%	0.98	Benchmark
7372	554	31.41%	48.69%	34.23%	47.99%	−2.82%	0.70%	−0.25	**

(continued on next page)

TABLE 6 (continued)

Panel B: Cost of Sales Series (Based on Models 5 and 9)

SIC Code	Number of Firm-Years	Percentage of Errors Alpha = 0.05		Peer Model Percentage of Errors Alpha = 0.05		Benchmark Minus Peer	Benchmark Minus Peer	Type II Divided by Type I	Superior Model
		Type I	Type II	Type I	Type II	Difference (Type I)	Difference (Type II)	Difference (Type II) / Difference (Type I)	
1381	132	64.65%	19.84%	64.23%	17.43%	0.43%	2.42%	5.68	Peer Model
2821	98	46.72%	19.79%	53.97%	17.92%	−7.25%	1.86%	−0.26	**
2834	391	55.60%	23.06%	59.59%	21.49%	−4.00%	1.56%	−0.39	**
2911	250	62.77%	17.05%	66.10%	9.30%	−3.33%	7.75%	−2.33	Peer Model**
3089	155	53.08%	9.60%	56.01%	9.91%	−2.93%	−0.31%	0.11	Benchmark
3312	192	53.51%	7.20%	56.24%	8.79%	−2.73%	−1.59%	0.58	Benchmark
3576	159	57.93%	20.52%	58.16%	20.86%	−0.23%	−0.35%	1.50	Benchmark
3661	214	60.26%	18.45%	63.21%	18.30%	−2.95%	0.15%	−0.05	**
3663	272	58.16%	21.84%	62.20%	23.68%	−4.04%	−1.84%	0.46	Benchmark
3674	327	60.98%	20.44%	65.06%	20.15%	−4.08%	0.29%	−0.07	**
3714	213	60.16%	10.64%	65.01%	9.92%	−4.84%	0.71%	−0.15	**
4911	562	52.11%	12.43%	57.03%	10.77%	−4.92%	1.66%	−0.34	**
5812	433	60.39%	8.65%	61.33%	8.29%	−0.94%	0.36%	−0.38	**
6021	567	63.56%	18.99%	63.53%	14.07%	0.03%	4.92%	158.68	Peer Model
6022	563	67.12%	20.44%	61.65%	15.48%	5.48%	4.96%	0.91	Peer Model
6331	403	60.02%	20.99%	62.28%	17.52%	−2.26%	3.47%	−1.54	Peer Model**
7370	262	62.81%	17.25%	61.68%	20.02%	1.13%	−2.76%	−2.44	Peer Model**
7372	554	57.12%	24.68%	58.97%	25.83%	−1.85%	−1.15%	0.62	Benchmark

** Depends on ratio of costs of Type I and Type II errors.

This table presents the error detection performance of Type I and Type II errors after material errors are seeded into the total revenues *and* AR accounts (cost of sales *and* the AP accounts in Panel B) (i.e., coordinated errors). Each error is seeded separately into each monthly observation and statistical investigation rule is used to detect these errors. Two different risk levels $\alpha = 0.05$, and $\alpha = 0.33$ are used and investigated separately for the peer model and the benchmark model. Predictions and confidence intervals are calculated using estimations from Models 3 and 7 (5 and 9 in Panel B). The differences between the two models are assessed by calculating the differences between the error rates of the two models. If both differences have an equal sign, we can unambiguously determine which model is superior.

percentage of both Type I and Type II errors, we can conclude that for that industry the peer model is superior. However, when one error type is increasing while the other is decreasing, an additional analysis is needed to resolve this ambiguity. We define *BenchmarkTypeI* and *II* be the percentages of errors generated using the benchmark model, *PeerTypeI* and *II* be the percentages of errors generated using the peer model, and *W(TypeI)* as the additional work performed by auditors when an error is detected by the procedure but no error actually exists, while *W(TypeII)* represent the cost to auditors from not detecting material error when such error exists. In order to conclude that the peer model is better than the benchmark model, the following inequality between the total error costs must hold:

$$\begin{aligned} &W(TypeI)^*(BenchmarkTypeI) + W(TypeII)^*(BenchmarkTypeII) \\ &\quad > W(TypeI)^*(PeerTypeI) + W(TypeII)^*(PerrTypeII) \end{aligned}$$

Rearranging this inequality for the case when $BenchmarkTypeI < PeerTypeI$, we obtain the following condition for the peer model to be better than the benchmark model:

$$\frac{-(BenchmarkTypeII - PeerTypeII)}{(BenchmarkTypeI - PeerTypeI)} > \frac{W(TypeI)}{W(TypeII)}.$$

Thus, the ratio of the two cost differences should determine the tradeoffs between the two models in the case of ambiguity. More specifically, we need to calculate the differences in the error rates for each error type between the two models and use the absolute value of the ratio of the differences (compared with *W(TypeI)/W(TypeII)*) as the determinant of which model yields better results.

It is documented in Table 6 Panel A that for $\alpha = 0.05$ the benchmark model is superior for five industries while the peer model is superior in two of the 18 industries. The above results represent cases in which both Type I and Type II error rates were superior. We proceed by analyzing the cases in which the overall benefits are ambiguous. We observe that in Panel A of Table 6, the absolute value of the ratio of the two cost differences is always greater than or equal to 0.5 indicating that even with a conservative assessment of Type II error being only twice as expensive as Type I error, the error detection performance of peer models is superior for the remaining 11 industries. Thus, for the revenue series, 13 of the 18 industries experience overall lower cost associated with error detection. Panel B of Table 6 yields similar results for the cost of sales series. We observe that the peer model is superior in one case while the benchmark model is superior in three cases. Additionally, we observe that the absolute value of the cost ratio is greater than 0.5 in 12 additional industries. Thus, 13 of the 18 industries experience superior error detection performance when the peer model is used. Panels A and B of Table 6 also display the error detection performance for $\alpha = 0.33$ levels. We find that differences between the peer and the benchmark models remain constant across risk conditions.[18]

ADDITIONAL ANALYSIS

The results of our study rely on a particular peer selection approach. Alternative procedures such as using the entire industry as peers or a manual selection of peers might be

[18] We perform the analyses in Tables 4, 5, and 6 using quarterly rather then interpolated monthly data and find that these results are consistent with those tabulated in the paper. Examining the differences between monthly and quarterly predictions, we find results that generally support the findings of previous studies that documented that higher frequency data generally result in superior predictions.

superior to our approach. Thus, we examine whether our proposed peers selection approach is superior to using the entire industry as peers. We estimate Models 3 and 5 using the entire industry as peers and compare the estimated results to our original peer-based predictions. We test the differences in the predictions together for the sales and cost of sales accounts by industry. We find that MAPEs generated by the peer model are smaller for 11 industries out of which five are statistically significant. However, we also observe that using the entire industry as peers leads to statistically significant improvement over our proposed approach for three industries. These results suggest that learning from specific peer companies rather than the entire industry often leads to superior results. We believe that practicing auditors should be able to use their knowledge and experience to select peers more accurately. Nonetheless, more evidence is needed in order to examine competing peer selection criteria for empirical academic research.

Previous research did not examine the impact of analytical procedures on companies experiencing rapid change. In fact, Allen (1992) suggests that many prior studies represent "independent success stories" by conducting case studies within the context of the retail and the manufacturing industries. These studies mainly involved companies that are members of a fairly stable group of industries that do not experience frequent structural changes. Chen and Leitch (1998) studied analytical procedures using simulated data and found that all models performed better for companies that have a greater degree of stability in their business and economic activities. This suggests that there is a need to investigate ways in which analytical procedures can be better applied to companies that are less stable and experience a greater degree of performance fluctuation. The use of time series data in regression models can capture systematic changes in account balances over time. However, companies may grow over time in a nonsystematic manner. Consequently, predictions for such companies are potentially more complex and more sensitive to longitudinal changes. Generally, it is harder to forecast account balances for fast growing or contracting companies. It is therefore expected that fast changing companies would experience inferior prediction and error detection performance.

Since this study investigates the performance of analytical procedures for multiple industries, we use this opportunity to examine whether company specific characteristics such as size are associated with the prediction performance. We conjecture that large companies are on average more stable and therefore their accounts are easier to predict. Hence, we expect that larger companies will experience greater prediction accuracy and superior error detection performance. We also expect that peer data will be especially beneficial for rapidly changing companies in comparison to stable companies

Models 10 and 11 (in Table 3) are used to test the association between the prediction accuracy and error detection performance to proxies that are associated with the rate of change and size of a company. The dependent variables are the *MAPE*, which is the Mean Absolute Percentage Error for each company year and *ERROR*, which is the sum of the monthly Type I or Type II errors for each company year. A surrogate for the level of change of a company is constructed by using the absolute average change in sales (Compustat quarterly data item 2) and *EPS* (Compustat quarterly data item 19) of a company during the estimation period, and the difference between the average rate of change in the estimation years to the average rate of change in the prediction year. The latter measures the changes between the period in which the model is developed and the period in which the model is tested. We use the log of the quarterly total assets (Compustat quarterly data item 44) to proxy for size. The dependent variable in Model 12, *MAPE-DIFF*, is the difference in the MAPE between the benchmark model and the peer model. Similarly, the dependent variable in Model 13, *ERROR-DIFF*, is the difference in the number of Type I or Type II errors between the benchmark model and the peer model. In both cases, high difference value indicates that the peer model is superior.

We use the prediction performance and the error detection performance for both the revenue and cost of revenue accounts. Thus, we run our analyses using a pooled model that contains predictions for both accounts. Panel A in Table 7 and Panels A and B in Table 8 present the results from Models 10 and 11, respectively. Results in Panel A of Table 7 are consistent with Chen and Leitch (1998) and overwhelmingly indicate that the level of change, as measured by our four growth proxies, is positively associated with inferior prediction accuracy (higher MAPE). It is also apparent that larger companies experience significantly lower MAPE indicating higher prediction accuracy. This is the case for both the peer and the benchmark models. Panels A and B in Table 8 document similar analyses for the Type I and Type II errors respectively. We observe that larger companies experience higher Type I errors and lower Type II errors for both the peer and the benchmark models. We also observe that both rate of change proxies are positively associated with higher occurrence of Type II errors. However, while we do not observe a consistent pattern in the association between the *EPS* rate of change and the occurrence of Type I errors, we do observe that a faster change in sales is positively associated with Type I errors. Thus, our analyses indicate that larger companies experience more accurate predictions and Type II error detection coupled with inferior Type I error detection. Additionally, companies with faster revenue change experience an increase in both Type I and Type II errors.

We continue by examining the association between company size and rate of change to the potential improvements attributed to using models with peer data. We use Models 12 and 13 and report our results in Panel B of Table 7 and Panel C of Table 8. While both *MAPE-DIFF* and *ERROR-DIFF* lack a consistent pattern for association with the rate of change, we observe that company size is positively associated with the improvement in MAPE and Type II error detection. Thus, we conclude that when auditing larger companies, auditors that use peer data could potentially improve their prediction accuracy and Type II error detection.

DISCUSSION AND LIMITATIONS

The study examines the potential benefits of using contemporaneous peer data in performing analytical procedures. We introduce an approach for selecting peers for each client and perform a number of tests to examine peers' contribution to the performance of analytical procedures. We use peer models in various ways and observe that peer data is extremely useful when no other contemporaneous variables are included. We also observe that when other contemporaneous variables are included, peer data are still useful when coordinated errors exist. Our results strongly indicate that using peer data is especially beneficial for improving the overall error detection performance. Furthermore, the results indicate that fast changing companies experience inferior prediction and error detection accuracy, and that large companies experience more accurate predictions, lower Type II errors, and higher Type I errors. Moreover, we find that significant improvements in the performance of analytical procedures are associated with larger clients suggesting that auditors of larger companies can potentially benefit more from the use of peer data.

Since audits are performed *ex post* (at the end of the period), auditors can currently include contemporaneous information from peer companies in the expectation models. Given the increasing industry/auditor concentration (Hogan and Jeter 1999), and the current consolidation among the large public accounting companies, auditors that specialize in certain industries can transfer information from one audit to the next and consequently improve the effectiveness of their analytical procedures. Additionally, practicing auditors

TABLE 7

Regression Models of Prediction Performance on Size and Various Change Proxies

Panel A: Regression of the Benchmark *MAPE* and Peer *MAPE* on Size and Various Change Proxies

Independent Variables	Dependent Variable *MAPE* from the Benchmark Model				Dependent Variable *MAPE* from the Peer Model			
Intercept	0.237 46.95	0.204 43.83***	0.177 34.12***	0.242 51.27***	0.263 53.49***	0.203 40.09***	0.2319 51.07***	0.27 58.53***
PRE-SALES[a]						0.1825 25.13***		
DIFF-SALES[b]		0.185 26.15***	0.176 23.61***				0.185 26.86***	
PRE-EPS[c]	0.0018 2.68***				0.0025 3.7***			
DIFF-EPS[d]				0.000 0.36				0.0003 1.72*
SIZE	−0.0068 −9.78***	−0.006 −9.27***	−0.0039 −6.00***	−0.007 −10.51***	−0.0119 17.43***	−0.0089 −13.84***	−0.011 −17.57***	−0.012 −18.41***
Adjusted R^2	0.011	0.0711	0.0605	0.011	0.0342	0.0875	0.0947	0.0337

(continued on next page)

TABLE 7 (continued)

Panel B: Regression of Differences in the *MAPE* between the Benchmark and the Peer Model on Size and Various Change Proxies

Independent Variables	Dependent Variable *MAPE-DIFF*			
Intercept	−0.027 −8.48***	−0.027 −8.36***	−0.030 −8.99***	−0.027 −8.44***
PRE-SALES[a]			0.010 2.28**	
DIFF-SALES[b]		0.000 0.00		
PRE-EPS[c]				0.000 −2.06**
DIFF-EPS[d]	0.000 −1.93*			0.000 0.36
SIZE	0.005 10.86***	0.005 11.13***	0.005 11.34***	0.005 10.84***
Adjusted R^2	0.0122	0.0114	0.0119	0.0123

*, **, *** Significant at the .10, .05, and .01 levels, respectively.
[a] *PRE-SALES* = average of the absolute change rate in sales during the preceding three years;
[b] *DIFF-SALES* = difference between the averages of the absolute rate of change in the estimation years to the prediction year;
[c] *PRE-EPS* = average of the absolute change rate in the *EPS* during the preceding three years; and
[d] *DIFF-EPS* = difference between the averages of the absolute *EPS* rate of change in the estimation years to the prediction year.

TABLE 8
Regression Models of the Sum of Errors on Size and Various Change Proxies

Panel A: Type I Errors

Independent Variables	Dependent Variable Type I Error from the Benchmark Model				Dependent Variable Type I Error from the Peer Model			
Intercept	3.556 27.75***	2.814 20.68***	3.327 27.51***	2.674 22.15***	3.798 29.44***	3.032 22.09***	3.562 29.26***	2.979 24.38***
PRE-SALES[a]		1.450 7.27***				1.539 7.65***		
DIFF-SALES[b]				3.328 18.24***				3.040 16.47***
PRE-EPS[c]	−0.105 −6.04***				−0.098 −5.58***			
DIFF-EPS[d]			−0.011 −2.34**				−0.001 −0.25	
SIZE	0.124 6.99***	0.153 8.83***	0.139 7.87***	0.149 8.9***	0.118 6.59***	0.151 8.61***	0.133 7.48***	0.143 8.42***
Adjusted R^2	0.0111	0.0109	0.007	0.0386	0.009	0.011	0.006	0.0322

Panel B: Type II Errors

Independent Variables	Dependent Variable Type II Error from the Benchmark Model				Dependent Variable Type II Error from the Peer Model			
Intercept	4.594 34.08***	4.248 29.78***	4.834 37.97***	4.439 34.66***	5.074 39.57***	4.576 33.77***	5.249 43.39***	4.797 39.46***
PRE-SALES[a]		1.675 8.01***				1.896 9.54***		
DIFF-SALES[b]				2.017 10.42***				2.257 12.28***

(continued on next page)

TABLE 8 (continued)

PRE-EPS[c]	0.100				0.075			
	5.47***				4.29***			
DIFF-EPS[d]			0.001				0.001	
			0.18				0.26	
SIZE	−0.197	−0.181	−0.210	−0.198	−0.317	−0.292	−0.327	−0.311
	−10.53***	−9.93***	−11.35***	−11.11***	−17.86***	−16.89***	−18.55***	−18.43***
Adjusted R^2	0.0178	0.0206	0.0140	0.0251	0.0397	0.0457	0.0368	0.0515

Panel C: Differences in Errors between the Peer and the Benchmark Model

Independent Variables	Dependent Variable Differences in Type I Errors				Dependent Variable Differences in Type II Errors			
Intercept	−0.243	−0.218	−0.234	−0.304	−0.481	−0.328	−0.415	−0.358
	−2.30**	−1.94*	−2.36**	−3.02***	−5.18***	−3.34***	−4.73***	−4.06***
PRE-SALES[a]		−0.089				−0.221		
		−0.54				−1.54		
DIFF-SALES[b]				0.288				−0.239
				1.90*				−1.80*
PRE-EPS[c]	−0.007				0.025			
	−0.52				2.01**			
DIFF-EPS[d]			−0.010				0.000	
			−2.54**				−0.11	
SIZE	0.006	0.003	0.006	0.006	0.121	0.111	0.116	0.114
	0.43	0.19	0.42	0.45	9.37***	8.89***	9.13***	9.26***
Adjusted R^2	0.000	0.000	0.000	0.000	0.0094	0.0091	0.009	0.0092

*, **, *** Significant at the .10, .05, and .01 levels, respectively.

PRE-SALES[a] = the average of the absolute change rate in sales during the preceding three years;
DIFF-SALES[b] = the difference between the averages of the absolute rate of change in the estimation years to the prediction year;
PRE-EPS[c] = the average of the absolute change rate in the EPS during the preceding three years; and
DIFF-EPS[d] = the difference between the averages of the absolute EPS rate of change in the estimation years to the prediction year.

with multiple clients could use their knowledge to better match peer companies and potentially utilize more granular data such as business and geographic segment data to achieve further improvements.

The current legal regime requires auditors to protect clients' data but does not forbid auditors from using these proprietary data for their own analyses. Specifically, Rule 301 of the American Institute of Certified Public Accountants (AICPA) Code of Professional Conduct (1996) states that "a member in a public practice shall not disclose any confidential client information without specific consent of the client." However, current rules do not restrict auditors from using clients' data to improve their audits. In fact, Guy and Carmichael (2002) interpreted the Statement of Auditing Standard (SAS) No. 56 and stated that "In circumstances where the auditor specializes in a specific industry, the auditor may use clients' data to develop plausible expectations (for example, gross margin percentage, other income statement ratios, and receivable and inventory turnover ratios)." Additionally, there is anecdotal evidence suggesting that national offices of large audit firms use data from a pool of companies in the same industry as a benchmark for other companies. The national office distributes such aggregated data to engagement auditors but does not disclose the exact profile of the companies that are used in the pool. Nonetheless, auditors need to be cautious not to disclose any confidential information when they ask management to provide plausible explanations for account balances that deviate from their expected values. Moreover, auditors need to make sure that private information about one client is not disclosed in the workpapers of another client because of the risk that this information could be subpoenaed.

Despite the fact that sharing information across auditing firms is only theoretically feasible, it is of great value to understand whether such an approach would result in improved performance of analytical procedures. If sharing client information among auditors is proven to significantly improve the performance of analytical procedures, then the auditing profession ought to consider advocating for a new regulation, under which peer information could be made available through a trusted third party. There is a precedent for this type of information sharing within the credit industry. In many countries, the credit industry shares information among credit checking companies through information brokers (credit bureaus). These brokers work on the principle of reciprocity, under which lenders who do not provide data to the bureau are denied access. Information sharing in the credit industry results in increased competitiveness in the credit markets, increased efficiency in asset allocation, and increased lending volume. Future regulatory initiatives within the auditing profession, using the credit industry as a model, could consider providing cross firm access to contemporaneous data through a trusted third party. Moreover, with the recent advancements in enterprise systems and networking there is now a suitable automated solution for employing peer-based analytics.

The technological ground for adopting the proposed peer-based approach for analytical procedures is widely available. XBRL (eXtensible Business Reporting Language) is a language for electronic communication of business and financial data. XBRL provides an identifying tag for each individual item of data (such as total sales), thus creating an unambiguous way to identify financial facts and making it seamless to compare the sales (or any other financial item) of one company to the next. Therefore, auditors can potentially automate the process of XBRL data collection from multiple companies and securely transmit this information to a central data repository. Hence, XBRL data from multiple companies can be collected automatically and securely allow auditors to perform analytical procedures that incorporate peer data shortly after the data is received. The technology for incorporating XBRL is currently supported by major ERP vendors (i.e., SAP R/3 and Oracle Financials), thus in many cases no additional software is needed. Because there is one uniform conventional way for digitally exchanging data, no human intervention is required. Therefore, using automated procedures, data from individual companies can be included in the regression models without becoming available to any individual audit team member, thus, helping auditors protect the confidentiality of their clients' data.

It is important to note that the reported results suffer from a number of limitations. Companies with no peers are dropped from the sample but still need to be audited. The reason for the elimination is the inherent limitation of automatically assigning peers to audit clients. It is likely that practicing auditors will be able to identify peers for most of the eliminated companies or otherwise charge a premium over the standard audit fees to potentially account for the additional risk and effort.

The results in this study should be carefully applied. The findings in this study are based on interpolated data points and not on real data. While these results support our conjectures of general prediction improvements and superior error detection, the accuracy of these predictions may not be sufficient for the practicing auditor. The purpose of this research was to study, in isolation, the prediction improvements that result from the incorporation of contemporaneous peer data. Any additional information that might be obtained through inquiry by an experienced auditor is not captured in our analysis. Thus, it is likely that the reported effectiveness of our models in detecting errors is estimated conservatively. Yet, our results strongly support the premise of information sharing across and within auditing firms and warrant further investigation.

REFERENCES

Allen, R. D. 1992. Statistical analytical procedures using industry specific information: An empirical study. Dissertation, Michigan State University.

———, M. S. Beasley, and B. C. Branson 1999. Improving analytical procedures: A case of using disaggregate multilocation data. *Auditing: A Journal of Practice & Theory* 18 (Fall): 128–142.

American Institute of Certified Public Accountants (AICPA). 1988. Analytical Procedures. *Statement of Auditing Standards* No. 56. New York, NY: AICPA.

———. 1996. Code of Professional Conduct. *Professional Standards*. Vol. II. New York, NY: AICPA.

Chen, Y., and R. A. Leitch. 1998. The error detection of structural analytical procedures: A simulation study. *Auditing: A Journal of Practice & Theory* 17 (Fall): 36–70.

———, and R. A. Leitch. 1999. An analysis of the relative power characteristics of analytical procedures. *Auditing: A Journal of Practice & Theory* 18 (Fall): 35–69.

Cogger, K. O. 1981. A time-series analytical approach to aggregation issues in accounting data. *Journal of Accounting Research* 19 (Autumn): 285–298.

DeFond, M., and J. Jiambalvo. 1994. Debt covenant violation and manipulation of accruals: Accounting choice in troubled companies. *Journal of Accounting & Economics* 17: 145–176.

Dugan, T., J. Gentry, and K. Shriver. 1985. The X-11 model: A new analytical review technique for the auditor. *Auditing: A Journal of Practice & Theory* 4 (Spring): 11–22.

Dzeng, S. C. 1994. A comparison of analytical procedures expectation models using both aggregate and disaggregate data. *Auditing: A Journal of Practice & Theory* 13 (Fall): 1–24.

Efron, B., and C. Morris. 1973. Stein's estimation rule and its competitors: An empirical Bayes approach. *Journal of the American Statistical Association* 68 (341): 117–130.

———. 1977. Stein's paradox in statistics. *Scientific American* 119–127.

Foster, G. 1981. Intra-industry information transfers associated with earnings releases. *Journal of Accounting and Economics* 3 (December): 201–232.

Guy, D. M., and D. R. Carmichael. 2002. *Practitioner's Guide to GAAS*. New York, NY: John Wiley and Sons.

Hawawini, G., V. Subramanian, and P. Verdin. 2003. Is performance driven by industry or firm specific factors? A look at the evidence. *Strategic Management Journal* 241: 16.

Hogan, C. E., and D. C. Jeter. 1999. Industry specialization by auditors. *Auditing: A Journal of Practice & Theory* 18 (Spring): 1–17.

Ijiri, Y., and R. A. Leitch, 1980. Stein's paradox and audit sampling. *Journal of Accounting Research* 18 (1): 91–108.

Knechel, R. W. 1988. The effectiveness of statistical analytical review as a substantive auditing procedure: A simulation analysis. *The Accounting Review* (January): 74–95.

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

Kwon, S. Y. 1996. The impact of competition within the client's industry on the auditor selection decision. *Auditing: A Journal of Practice & Theory* 15 (Spring): 53–69.

Leitch, R. A., and Y. Chen. 1999. Simulation of controlled financial statements. *Review of Quantitative Finance and Accounting* 13: 189–207.

———. 2003. The effectiveness of expectation models in recognizing error patterns and generating and eliminating hypotheses while conducting analytical procedures. *Auditing: A Journal of Practice & Theory* 22 (Fall): 147–170.

Lev, B. 1980. On the use of index models in analytical reviews by auditors. *Journal of Accounting Research* (Autumn): 524–550.

Loebbecke, J. K., and P. J. Steinbart. 1987. An investigation of the use of preliminary analytical review to provide substantive audit evidence. *Auditing: A Journal of Practice & Theory* (Spring): 74–89.

Perry, S., and T. Williams. 1994. Earnings management preceding management buyout offers. *Journal of Accounting & Economics* 18: 157–179.

Taylor, M. A. 2000. The effects of industry specialization on auditors' inherent risk assessments and confidence judgments. *Contemporary Accounting Research* 17 (Winter): 693–712.

U.S. General Accounting Office (GAO). 2002. *Financial Statement Restatements: Trends, Market Impacts, Regulatory Responses, and Remaining Challenges*. GAO-03-138. Washington, D.C.: Government Printing Office.

Wheeler, S., and K. Pany. 1990. Assessing the performance of analytical procedures: A best case scenario. *The Accounting Review* 65: 557–577.

Wild, J. J. 1987. The prediction performance of a structural model of accounting numbers. *Journal of Accounting Research* (Spring): 139–160.

———, and F. S. Biggs. 1990. Strategic consideration for unaudited account values in analytical review. *The Accounting Review* 65 (January): 227–241.

Wright, S., and A. Wright 1997. The effect of industry experience on hypothesis generation and audit planning decisions. *Behavioral Research in Accounting* (Spring): 273–294.