Prepublication version of the paper in *Journal of Information Systems*, Vol. 16, No. 2, Fall 2002, pp. 165-182.

Does the Year 2000 XBRL Taxonomy Accommodate Current Business Financial Reporting Practice?

Matthew Bovee [1]
Michael Ettredge [2]
Rajendra P. Srivastava [3]
Miklos Vasarhelyi [4]

February 21, 2001

[1] Doctoral Student, School of Business, The University of Kansas, Lawrence, KS 66045, email: mbovee@ku.edu
[2] Corresponding author. Associate Professor, Accounting and Information Systems, The University of Kansas, Lawrence, KS 66045, email: mettredge@ku.edu
[3] Ernst & Young Professor and Director, Ernst & Young Center for Auditing Research and Advanced Technology, School of Business, The University of Kansas, Lawrence, KS 66045, email: rajendra@ku.edu
[4] Professor, Faculty of Management, Rutgers University, email: miklosv@andromeda.rutgers.edu

Acknowledgements: We thank our student research assistants M. Gao, N. Govind, J. Mehta, and D. Jyoti. This study was funded by several firms that are members of XBRL.Org, and is intended to provide the XBRL taxonomy task force, and other interested parties, with feedback about the actual fit between the existing taxonomy and recent current practice as represented by the selected financial reports. We thank the firms that funded this effort: EDGAR Online, Ernst & Young, KPMG, PricewaterhouseCoopers, and XBRL Solutions. Special thanks are due to Liv Watson, who made the project possible. We received helpful comments from Zack Coffin, Eric Cohen, David Hardidge, Charlie Hoffman, David Vun Kannon, and Mark Schnitzer.

Does the Year 2000 XBRL Taxonomy Accommodate Current Business Financial Reporting Practice?

Abstract

XBRL (Extensible Business Reporting Language) is an application of XML (eXtensible Markup Language) intended for use in digital business reporting. Observers predict XBRL will provide benefits to firms that adopt it, such as enhancing information use, facilitating comparability and consistency, and providing technological capability for near-continuous financial reporting through the Web. However, questions arise regarding how well the proposed taxonomy for financial statements corresponds to firms' preferred reporting practices. We argue that a poor fit may lead to information loss and thus to resistance to use or general adoption of the taxonomy. A lack of fit could therefore negate anticipated firm or information-user benefits.

To address this issue, we assess how well the year 2000 version, for financial reporting by commercial and industrial (C&I) firms under U.S. GAAP, accommodates current financial reporting practices of public companies. We attempt to match each line item in the 1999 annual financial statements of 67 companies with an XBRL taxonomy 'tag,' employing two measures of the frequency of 'special attention' (difficult-to-match) items as indicators of the goodness of 'fit' between the taxonomy and firms' reporting practices. Analyses show a good fit on average, but with significant differences in the number and proportion of exceptions across financial statements and industries. In light of these results, we suggest modifications to the taxonomy, and discuss the need for industry-specific taxonomies.

Introduction

XBRL (Extensible Business Reporting Language) is an application of XML (Extensible Markup Language) intended for use in digital business reporting. Hill (2001, 55) describes XML as: "enabling data on the Web or any large network to be readily swapped between any kind of device...any kind of application, regardless of what programming language the application was originally written in."[1] To accomplish this, XML 'tags' enclose (delimit) each fact or item of data generated. The data item and tag together constitute a string of plain text that can be digitally transmitted. XML tags provide enabled software with context information to aid interpretation of the data, with multiple-nested tags providing additional context. XML-consistent tags allow a firm to aggregate financial reporting data across its sub-units, even if they use incompatible hardware and software, through the consistent use of like tags for comparable data.

Anticipating the value of XML, the AICPA and major international firms are supporting the efforts of the XBRL.Org, an international consortium of firms, to develop XBRL. This effort includes development of a taxonomy for financial reporting under U.S. generally accepted accounting principles (GAAP). That taxonomy is intended to provide a set of XML-consistent tags that identify various items of financial and non-financial information relevant to business decisions. The goal of the developers of XBRL is to tag (in an XML-based framework) every piece of information relevant to business reporting and decision-making. A taxonomy of such tags would provide a stable, semantically consistent system, enabling not only efficient and effective search and reporting of such information, but continuous monitoring and auditing of such information as well.

In addition to the task force that is developing a taxonomy for financial reporting under U.S. GAAP, several others are developing taxonomies for different purposes, such as a taxonomy for management reporting of balanced scorecard information, and a taxonomy to be used with IRS tax filings. Taxonomies also are being developed for financial reporting by governments. The ebXML task force (not affiliated with the XBRL consortium) is developing a taxonomy for tagging information at the transaction (versus account) level. This system is intended ultimately to provide a digital language for continuous monitoring and auditing.

For these benefits to be realized, such taxonomies should capture information reflected in current reporting practices sufficiently well to be adopted by the communities they are intended to serve. Otherwise, the taxonomies are likely to be met with resistance, to be modified and thus lose semantic consistency, or to be disregarded in favor of other (probably customized) solutions. The goal of this article is to assess how well the year 2000 version of the XBRL taxonomy for financial reporting by commercial and industrial (C&I) firms under U.S. GAAP accommodates current financial reporting practices of public companies. We argue below that this issue is interesting because lack of fit between the taxonomy and firms' preferred reporting practices leads to loss of information and, potentially, to resistance to adoption of the taxonomy.

In the next section of the paper we provide more background on XBRL. Then we discuss the potential for loss of information when applying the XBRL taxonomy to firms' financial reports. This leads to a section that specifies the research issues and describes the research methods. We then discuss the sample and the tagging process in which the taxonomy was applied to sample firms' financial statements. After a section that describes some of the difficulties encountered in tagging, we provide results followed by conclusions.

Background on XBRL

This section provides an overview of extensible business reporting language (XBRL). XBRL is intended for business reporting, both internal and external to a firm. An initial application will be to publish digital corporate financial statements for external users. According to the XBRL consortium,[2] XBRL enables a dramatic improvement in the processing of financial reports. XBRL documents can be prepared efficiently, exchanged reliably, published easily, analyzed quickly, and retrieved by investors simply.[3]

Observers predict that XBRL soon will give many firms the technological *capability* for providing near-continuous financial reporting through the Web or other mediums.[4] *Motives* for more frequent financial reporting also exist, particularly since the Security and Exchange Commission (SEC) issued Regulation FD in October 2000. Regulation FD (for "fair disclosure") essentially requires that companies deliver material information simultaneously to all audiences (Parish 2000). It is intended to prevent corporate managers from favoring some analysts over others, or over the public as a whole, at least with respect to material information. While Regulation FD still requires companies to disseminate material information via press releases, it specifically encourages them also to use the Internet to accomplish broad and rapid dissemination (Baker 2000).

One observer predicts that "FD should jump-start the move to real-time reporting" (Seligman 2000, 148). The rationale is that firms can best comply with FD through continuous release of financial data to the public rather than quarterly financial reports supplemented by occasional news releases. In addition, continuous release of financial data arguably should prevent the major stock price declines that often accompany occasional news releases, and that generate lawsuits against managers. XBRL should facilitate real-time financial reporting via the

Internet, although we emphasize that XML and XBRL provide agreed-upon methods for exchanging information via any digital medium, not just the Internet.

To implement XBRL requires the XBRL Specification (Hampton and vun Kannon 2001), the XBRL Schema, an XBRL Taxonomy, and an XBRL Instance document.[5] The specification provides the technical grammatical rules for creating taxonomies and instance documents; the schema defines the core building blocks for building them; a taxonomy defines the tags for things of interest and their interrelationships; and the instance document is the actual business report containing the information of interest embedded in tags defined by the taxonomy plus address information pointing to the definitional documents. We focus here on the XBRL taxonomy.

Accounting software vendors are likely to design XBRL capability into their accounting systems to allow users to cross-reference their accounts to this framework, enabling a more efficient reporting process and a more efficient experience for the users of financial information. Consider the following example. A firm's Balance Sheet contains the following data item for 1999:

Net Receivables $153

In the electronic file containing the financial statements, XBRL coding represents this information as follows:[6]

```
< ci:currentAssets.receivablesNet numericContext="c1">153
</ci:currentAssets.receivablesNet>
```

Additional coding specifies the grouping of this element with related elements (i.e. the items belonging to the balance sheet), the other elements in the grouping, the definition of numeric context of the value (eg. the reporting organization, precision, period, unit of reporting, reporting

scenario and assumptions), and the reference locations and documents indicated by namespace abbreviations such as "ci". This coding comprises 'meta-data' about the content and context of the data. Net receivables are part of current assets; current assets are part of total assets that, in turn, forms a portion of the balance sheet. Typical XML nesting of tags would require this item to be enclosed by *at least* four sets of tags to capture this relationship. However, XBRL uses shorthand: the "ci" part of the type tag above points to the external file containing the C&I taxonomy structure, where these relationships are specified. Both the dollar amount and its identifying information are a plain text string not specific to any software application. The group type tag includes the identity of the element or account, and of its 'parent' account. If an investor downloads the XBRL-tagged financial statements from the firm's Web site, XBRL-enabled software recognizes the amount, $153. Furthermore, the amount is labeled as 'Receivables, Net,' (the standard label for the item as specified in the referenced taxonomy) and located among 'Current Assets' under 'Assets' in a correctly formatted Balance Sheet.

The phrase 'correctly formatted Balance Sheet' above implies a standard that specifies what elements should appear in a Balance Sheet and how they should be arranged. A financial reporting 'taxonomy' can be defined as a chart of accounts with agreed-upon meanings, together with specified relations between the accounts. Agreed-upon account tags are useful for comparability because firms frequently employ different account titles that mean essentially the same thing. One firm might have an account titled 'Trade Receivables' on its Balance Sheet. Another firm might have an account titled 'Accounts Receivable.' If each firm identifies its account with the taxonomy tag of the type:

ci:accountsReceivableTradeNet.accountsReceiveableTradeGross,

XBRL-enabled software will recognize that the two balance sheet items refer to the same asset. A financial reporting taxonomy therefore provides a means for retrieving and comparing digital financial information across firms.

The XBRL taxonomy for financial reports consists of hierarchies of accounts. For example, in the 'Balance Sheet' taxonomy, the element 'Accounts Payable' has two sub-elements or children: 'Trade Accounts Payable' and 'Other Accounts Payable.'

Trade Accounts Payable
<u>Other Accounts Payable</u>
Accounts Payable

As represented above, the 'children' of a 'parent' account are located immediately above that account, with the last child being underlined. When dollar amounts are assigned to each account, the amount assigned to a parent should equal the sum of the amounts for its children. This representation of a hierarchy inverts the tree structure view employed by the XBRL taxonomy task force. However, it resembles the presentation of financial statement data and related totals in financial reports, and we use it in subsequent examples.

It is worth noting here that the taxonomy does not represent a 'universal chart of accounts,' since its use is not mandatory. As envisioned, a firm can maintain any internal chart of accounts it wishes. However, firms that want to distribute the data in digital form, and that want the meaning of the data to be understood by XBRL-enabled software, will need to identify each of their accounts with a tag from the relevant taxonomy, or to create a custom tag as discussed subsequently.

As figure 1 indicates, the year 2000 XBRL taxonomy is largely derived from procedures employed by large accounting firms when conducting audits of clients' financial statements.[7] These procedures include checking the financial statement contents against lists of financial

statement elements required by U.S. GAAP.[8] The auditors' procedures affect managers' reporting decisions, since financial statement presentations must comply with GAAP. However, managers' reporting decisions are influenced by additional factors. For example, the desires of information users, or managers' self-interest, might cause managers to provide more financial information than is required by GAAP. Additional reasons why firms' current reporting practices do not map perfectly to the XBRL taxonomy are as follow. GAAP defines the elements of financial statements, and provides limited guidance regarding the relations between the various elements. However, the components of U.S. GAAP were issued piecemeal over a period of decades and alone do not provide a well-specified taxonomy of elements. As mentioned above, GAAP frequently allows firms to use various terms when referring to the same construct.[9] For another example, although various GAAP pronouncements identify individual accounts to be treated as current assets, GAAP does not provide a complete and official list of current assets. For all these reasons, firms' financial reports that comply with GAAP will not necessarily map perfectly to elements of a taxonomy that also is consistent with GAAP. In figure 1 these differences in 'fit' are represented by the broad, two-headed arrow. The creators of the XBRL taxonomy attempted to close this gap, to some extent, by referring to various issues of *Accounting Trends & Techniques* for information about firms' actual reporting practices.

The next section discusses how lack of fit between the XBRL taxonomy and firms' preferred reporting practices can affect the quality of financial information in digital financial reports created using XBRL.

Potential for Loss of Information

We begin with a brief coverage of the characteristics of useful information as stated by the Financial Accounting Standards Board (FASB) in *Statements on Financial Accounting Concepts No.2* (FASB 1980). This provides a widely accepted terminology for discussion of the concept of 'useful information.' Then we discuss how lack of fit between the taxonomy and current reporting practices can cause loss of information.

SFAC No. 2 describes the important qualitative characteristics of financial statement information that is useful for decision making, in particular by external users such as investors, creditors, and regulators. For financial reporting, the primary decision-specific qualities are relevance and reliability. Essentially, relevance means that the information (if valid) is useful to the decision maker, and reliability means that the information is free from error and bias, and faithfully represents what it purports to represent. In order to clarify the meaning of these terms, *SFAC No. 2* further subdivides these major characteristics into the following components: *value for feedback and prediction, timeliness, comparability, consistency, verifiability, neutrality, and representational faithfulness*.

Examining the nature and purpose of the XBRL taxonomy, it is apparent that the design of the taxonomy potentially affects the quality of information in XBRL-generated reports via some, but not all, of these dimensions. For example, the nature of the taxonomy should not affect the *timeliness* dimension.[10] Two other characteristics of information unlikely to be affected by the nature of the taxonomy are *verifiability* and *neutrality.* In contrast, however, the nature of the taxonomy could affect several other characteristics of information identified by *SFAC No. 2* and discussed below, including *feedback and predictive value*, *comparability*, and *representational faithfulness*.

For cost and benefit reasons, similar data items frequently are combined for presentation in financial reports. For example, GAAP allow firms to present a total for 'Accounts Receivable' rather than listing each customer's account individually with its amount. The latter would produce financial reports of excessive length, and would provide competing firms with information that could damage the reporting firm. Thus financial reporting involves aggregation of similar data. Aggregation of accounting data involves loss of detail, but the result may or may not be less useful to an investor. Accountants and managers typically apply knowledge and judgment when aggregating accounting data. This might compensate in some cases for the lost detail (Sunder 1997). In general, however, aggregation involves loss of information.

The aggregation of data does not cease at the level where individual data items (such as individual receivables) are 'rolled up' into a master or control account total (such as Accounts Receivable). Within limits imposed by GAAP and the Securities Exchange Commission (SEC), firms' managers can decide how far to aggregate control accounts when designing financial statements. For example, a firm might disclose on its balance sheet an amount for 'Property and Plant, Net' (i.e., net of accumulated depreciation), and an amount for 'Equipment, Net.' Or, it could combine the two into an account titled 'Property, Plant and Equipment, Net.' The latter choice provides the user with data that are more aggregated than does the former choice. The level of aggregation desired by managers likely varies with their incentives and with the firms' circumstances.[11]

Differences between the taxonomy and firms' desired reporting practices, represented by the broad, two-headed arrow in figure 1, potentially cause loss of information. Thus it is desirable for the taxonomy to capture information at the same level of detail provided by the preparing firms, to the extent that is possible. However, the taxonomy must embody *one*

particular collection of control accounts. It is impossible to create a taxonomy that will accommodate the reporting practices of every firm. Firms that desire the same level of dis-aggregation embodied in the taxonomy (or less) can achieve complete *comparability* for their data, by using the taxonomy as written.[12] Firms that desire greater dis-aggregation than the taxonomy provides can preserve *predictive and feedback* value by creating 'custom' tags that extend the taxonomy.[13] Firm-generated custom tags, however, involve a potential loss of *comparability* across firms.

Consider, for example, a firm's published balance sheet that discloses an account titled 'Investment in ABC Company,' and another titled 'Investment in XYZ Company,' among its non-current assets. Each of these accounts can be mapped to the tag titled 'Long Term Investments.' Thus one option the firm has is to use that tag twice: once for each investment stated. Use of that tag preserves *comparability* with amounts tagged as 'Long Term Investments' by other firms, but does not preserve all the information in the firm's Balance Sheet. If the example firm wants to preserve the information that one 'Long Term Investment' amount is associated with ABC Company, and the other with XYZ Company, it can create custom tags as children (extensions) of the parent tag for 'Long Term Investments' or as completely customized, idiosyncratic tags. Extensions to the taxonomy preserve *predictive and feedback* value while permitting firm- or industry-specificity. Custom tags thus allow a firm to preserve more detail, but at the possible cost of *comparability*. XBRL-enabled software will not recognize custom tags such as 'Long Term Investment in ABC' and 'Long Term Investment in XYZ' as equivalent to the taxonomy's 'Long Term Investment.' However, the amounts associated with the custom tags should roll up into the sum associated with their parent account, and *comparability* is restored at the level of the parent account (and, in turn, of its parents).

Finally, the design of a taxonomy potentially affects the *representational faithfulness* of information. The XBRL taxonomy might employ a single, standardized tag for an account that many firms prefer to describe with a variety of near-synonyms. A situation such as this raises a question whether the XBRL tag in question faithfully represents what it purports to represent.

In summary, differences exist between the taxonomy and firms' preferred reporting practices. Differences in aggregation of data and accounts are particularly important. If the taxonomy is *more* aggregated than managers prefer, firms might employ the taxonomy as written, resulting in loss of information. Or, they can accomplish desired dis-aggregation using firm-specific, custom tags. Such tags, however, do not represent agreed-upon concepts. Different firms can devise different custom tags for the same account. This leads to loss of information comparability across firms. If the taxonomy is *less* aggregated than managers prefer, problems sometimes can be avoided. For example, a firm might present data for a particular 'parent' account while not presenting information at a more dis-aggregated level, i.e., for 'children' of that parent account. Unfortunately, firms' preferred methods of aggregating information for presentation often result in violations of the taxonomy hierarchy. That is, a firm might desire to present an account that is equivalent to the aggregation of children from two or more *different* parent accounts. This requires a custom tag and, again, leads to loss of information when comparing financial data across firms. These problems are discussed in more detail, and examples are provided, in a subsequent section. First, however, we specify the research issues that we address.

Research Issues and Methods

The primary research question of interest is: "How good is the 'fit' between the 2000 taxonomy and firms' current reporting practices?" We use two variables to address the question. First, we identify and count instances where a sample firm's reporting practices do not agree with the taxonomy. We refer to such instances as a 'special attention items' (SAIs), since each such instance is investigated. Let SAIi denote the number of SAIs for firm i. A second metric converts the number of exceptions, SAIi, into a proportion for each firm. Let LINEi equal the total number of line items in a sample firm's financial statements. We define PROPi equal to SAIi divided by LINEi. Variable PROPi is the proportion of line items (in a firm's financial statements) that do not map to the taxonomy. We present descriptive data for SAIi per firm and for PROPi per firm for sample firms. The descriptive statistics and associated tests of central tendency provide evidence regarding the taxonomy's goodness of 'fit.'

In order to assess whether 'fit' differs by financial statement, we present descriptive statistics for SAIi per firm by financial statement. The financial statements examined are the 'Balance Sheet,' (BS) 'Income Statement,' (IS) and 'Statement of Cash Flows' (SCF). We also present descriptive statistics for PROPi per firm by financial statement. We test whether mean SAIi is equal across the three financial statements, and whether mean PROPi is equal across statements. In order to assess whether 'fit' differs by industry, we present descriptive statistics for PROPi per firm by industry. We test whether mean PROPi is equal across the industries investigated.

Finally we examine whether the 'fit' between firms' reporting practices and the taxonomy differ with firm size, and with a proxy for Web-based financial reporting practices. We assume that firms currently using their Web sites to disseminate extensive financial

information will be more interested than others in using XBRL for this purpose. We investigate whether financial reports of firms currently providing more financial information at their corporate Web sites tend to have better or worse 'fit' with the taxonomy. A research assistant visited each sample firm's Web site and determined the amount of financial information provided. A scoring instrument developed by Ettredge, Richardson and Scholz (2002) was used to code the extent of financial information at the site. Sites with none of the scored items received a zero score, while companies with no Web site were not included in this portion of the analysis. Let firm i's Web reporting score total be designated as 'WEBi'. To determine the relations between firm size, Web reporting, and 'fit,' we estimate the correlations between SALEi (natural log of 1999 revenues for firm i), WEBi, SAIi, and PROPi.

We turn now to a description of the sample and of the 'tagging' process used to gather the data underlying the variables described above.

The Sample and the Tagging Process

To accomplish the goals of the study we obtained a representative sample of public firms and industries. From the industry groups available at the Fortune website (www.fortune.com) we selected ten that represent a variety of businesses. In order to obtain variance in sample firm size, we selected the largest five and smallest two firms (by revenue ranking) within each industry chosen. From the initial sample of seventy firms we deleted one firm that had initiated Chapter 11 filings, and two other firms (mentioned below), leaving a final sample of 67 firms from 10 industries.[14] Sample firms and industries are shown in the Appendix. The sample spans nearly the entire range of the Fortune 1000 ranking by 1999 revenue (2^{nd} to 999^{th}). The annual SEC filing (10K, 10K405, 10K405A, DEF14A) containing the balance sheet, income statement and

cash flow statement for each company was retrieved from the SEC EDGAR database (www.sec.gov) and these financials and source files were individually archived. The financial statements were printed individually, preserving original formatting, for comparison with the taxonomy.

Research assistants compared individual line items from sample firms' financial statements with the taxonomy and assigned a taxonomy tag to each item when possible. Line items that could not easily be mapped to a tag by the research assistants were designated as 'special attention' items. A two-person team consisting of a faculty member and a doctoral student reviewed all special attention items and attempted to assign tags. The approach employed was to identify each special attention item with an existing tag when possible, and to use each tag no more than once per firm. Multiple uses of the same tag arguably results in lost information. The remaining number of special attention items equals SAIi for each firm i. One of the authors scanned all sample firms' financial statements, and counted the number of line items, LINEi, for each of a firm's three financial statements, as well as total LINEi summed across the three statements. PROPi then was computed for each firm's individual financial statements, and for each firm's three statements combined.

We wish to emphasize that our approach used to compute SAIi somewhat overstates the problems involved in assigning tags to line items for three reasons. First, we used each tag only once although the taxonomy allows multiple uses. Thus we exhausted the available tags and generated SAIs more often than strictly required. However, this permits identifying cases where additional detail might be desired by firms. Second, nearly every section of the taxonomy provides an 'Other' tag that could also be used multiple times for otherwise untaggable data.[15] Third, we did not refer to financial statement footnotes to clarify the nature of ambiguous

account titles employed by sample firms. Many of the SAIs likely could be assigned to tags using footnote information.

Difficulties encountered in tagging are described and categorized in the next section.

Difficulties Encountered in Tagging

When investigating individual SAIs, it became apparent to us that there are three general categories. We assigned each SAI to one of these categories. First, line items for which a taxonomy tag does not exist, but arguably should, were designated as potential 'new tag' items. Second, line items that mapped to several taxonomy tags were labeled 'grouped' items. Third, remaining SAIs were classified as 'firm specific.' Generally these consisted of idiosyncratic line items containing firm- or industry-specific detail. Each category of SAI is more thoroughly described in the sub-sections below with accompanying examples drawn from sample firms' major financial statements: BS, IS, and SCF. Then we provide descriptive statistics for the SAIs.

Proposed New Tags

Our comparison of firms' published financial statements with the taxonomy revealed numerous instances where the firms' accounts are more dis-aggregated than the (unmodified) taxonomy permits. In such instances we tried to decide whether the accounts in question represent firm-specific items that do not justify a proposed new tag, or represent new tags that perhaps should be added to the taxonomy. At a general level, it is unclear to us how to judge whether a proposed financial reporting taxonomy provides an 'optimal' level of dis-aggregation. The admittedly *ad hoc* procedure we employed was based on frequency of occurrence. If a particular account title could not be mapped to an XBRL tag, and that account title appeared in

the financial statements of several sample firms, we identified that account as a potential 'new tag'.

A number of sample firms disclosed purchases of intangible assets in the Investing section of their SCFs. The year 2000 version of the XBRL taxonomy provides for purchases of various types of property, plant and equipment, and for purchases of financial securities, but does not provide a tag for purchases of intangibles such as patents, copyrights, etc. Based on sample firms' reporting practices we recommend the following new tags:

Software Purchased or Capitalized
Trademarks, Brand Names, Copyrights
Patents
Franchise Rights and Licenses
Goodwill
Media Content
Finance Contracts
Other Contractual Rights
Customer Lists
Other Intangible Assets
Payments for Intangible Assets and Software

We recommend addition of the same new tags (excepting the 'Payments' tag) to the BS taxonomy, as well as to the SCF taxonomy.[16]

Firm-Specific Tags

If only one or two sample firms reported a particular account title, we identified that account as 'firm-specific'. It is worth noting that some items we identified as 'firm specific' might actually be industry-specific.[17] We return to this point when examining the distribution of firm-specific SAIs across sample industries.

Ford and General Motors provide examples of firm-specific reporting. Each provides separate information in its BS for its financing activities and for its manufacturing and other activities. For example, Ford discloses 'Cash and Cash Equivalents' in its list of Automotive

Assets, and again in its list of Financial Services Assets, rather than a single 'Cash and Cash Equivalents' line item for the entire firm. The two firms could tag their balance sheets by business sector, if desired, using either custom tags or extended tags. This approach reduces *comparability* with other firms, but preserves any *predictive* and *feedback value*.[18]

Grouped Items

Grouped items represent situations in which sample firms aggregate accounts in a way that does not comply with the parent/child relations embedded in the taxonomy. Often a sample firm will disclose an account that can be mapped to two or more tags (i.e., the firm's account maps to a 'group' of tags). However, these tags do not constitute the only children of the same parent tag. If they did, the firms' account could simply be mapped to the parent tag.

Clearly, firms can create custom tags that combine two or more tags if desired. Items that are custom-tagged, however, might not be comparable with the *same* items custom-tagged by other firms if the various firms use differing terminology. In the simplest case, as illustrated in figure 2, *comparability* is lost up to the second level above the point of grouping. When items A.1.2 and A.2.1 are combined into grouped item B, items A.1 and A.2 are no longer comparable across users of the original taxonomy. Comparability is preserved only for existing siblings of A.1.2 and A.2.1, and for the 'grandparent' item A. Such coarsening of information and loss of *comparability* is worse when grouping spans levels of a hierarchy as well as branches.

Xerox Corp.'s (12/31/1999) IS discloses:

> Restructuring charge and asset impairments

This maps to two different taxonomy tags: 'Restructuring Charges' and 'Impairment of Long Lived Assets.' Both tags are children of the parent tag 'Operating Expenses.' The firm's account

could be mapped to the parent tag, but only if Xerox did not disclose other accounts that also are children of 'Operating Expenses'.

Descriptive Statistics

Table 1 shows SAIs by financial statement and by the types (categories) to which they were assigned. Viewed by type, grouped items occurred least often (n = 230), new tag items more often (n = 295), and firm specific items most often (n = 340). Viewed by statement type, balance sheet and income statement items (n = 214 and 219, respectively) occur least often and are dominated by cash flow statement exceptions (n = 433). When viewed by statement *and* type it is clear that 'new tag' and 'firm specific' items outnumber 'grouped' items for the IS and SCF. The opposite is observed for the BS, with grouped exceptions predominating. Inspection of the year 2000 taxonomy confirms that the balance sheet is the most dis-aggregated portion of the taxonomy, and the prevalence of grouped items for that statement suggests that sample firms frequently employ less dis-aggregation than the taxonomy permits. Conversely, the IS and SCF taxonomies are less dis-aggregated. This results in sample firm IS and SCF statements with numerous proposed new tags or firm specific tags.

Table 2 present SAIs by category of SAI and industry. The entertainment and the petroleum refining industries stand out from the others, particularly with regard to numbers of firm-specific SAIs. The large number of firm-specific items for these two industries indicate that the SAIs we identified as firm-specific likely include a number of items that actually are industry-specific. The beverages and pharmaceuticals industries have considerably lower numbers of total SAIs (motor vehicles and parts are low due to deletion of two sample firms). These results suggest the taxonomy is a somewhat better 'fit' for some industries than others. In

particular, it suggests that efforts underway and planned to develop industry-specific taxonomies are justified.

Table 3 shows SAIs by financial statement and by quintiles sorted on the Web financial disclosure score, WEBi. The WEBi quintiles capture firms' use of corporate Web sites to disseminate financial information to investors (a proxy for propensity for early adoption XBRL). The lowest-scoring companies are in quintile WEBQ1 and successively higher-scoring companies are in quintiles WEBQ2, etc. In general, the number of SAIs for each financial statement is lower for high scoring firms (WEBQ4 and WEBQ5) than for low scoring firms (WEBQ1 and WEBQ2). In other words, the number of SAIs in all three financial statements generally decreases as Web-based financial dissemination scores increase. The data suggest that firms most likely to be early adopters of XBRL have financial statements that fit the taxonomy well.

Table 4 provides numbers of SAIs by type of SAI and by Web score. For all three types of SAIs, firms currently using their Web sites most extensively to disseminate financial data have smaller numbers of exceptions.

Analyses by firm size (revenue rank) did not show any associations of interest. We turn now to the research questions of most interest.

Research Results

Table 5 indicates that the median number of SAIs per firm is 12. About half of the Sais are generated by the SCF (with median = 6) versus the IS and BS (median = 3 each). The hypothesis that the number of SAIs is the same across financial statements can be rejected (F-stat = 27.7, $p<0.001$). The mean SAIi per firm, in the population from which sample firms are drawn,

differs from zero at a high level of significance (t-stat = 18.2, p<0.001). The median of PROPi per firm, expressed as a percentage, is 14%. The largest PROPs are generated by the SCF (with median = 18%) versus the IS (median = 13%) and BS (median = 9%). The hypothesis that PROPi is the same across financial statements can be rejected (F-stat = 16.0, p<0.001). The mean PROPi per firm, in the population from which sample firms are drawn, differs from zero at a high level of significance (t-stat = 20.9, p<0.001). The statistics suggest the XBRL taxonomy provides a good fit with firms' current reporting practices. For example, 75% of sample firms have PROPi of 18% or less. However, the mean numbers and proportions of exceptions, per firm and per statement, are not negligible. Thus there is room for substantial improvement.

Table 6 presents descriptive statistics for PROPi per firm by industry. Median PROPi per firm varies from 8% for Food & Drug to 21% for petroleum refining. A test rejects the hypothesis that mean proportion of exceptions per firm is equal across industries (F-stat = 2.8, p=0.004). Differences in mean SAI by industry (not tabulated) also are significant (F-stat = 2.5, p=0.011). Similar to table 2, these results indicate the C&I taxonomy is a better fit for firms in some industries than in others, and suggests that industry-specific taxonomies could be desirable.

Table 7 provides Pearson correlations among variables SAIi, PROPi, WEBi, and a proxy for size, sales (SALEi), all measured on a per firm basis. SAIi and PROPi are positively correlated as might be expected. WEBi is negatively associated with SAIi, consistent with the Totals columns of tables 3 and 4. Other correlations are not significant.

Pair-wise comparisons in table 8 show that statements of cash flow (SCF) had significantly greater mean proportions of exceptions than income statements (IS), and also greater proportions than balance sheets (BS). Income statements (IS) had significantly greater mean proportions of exceptions than did balance sheets (BS).

Table 9 dis-aggregates the table 8 results by industry. An F-test (not tabulated) rejects the hypothesis that mean PROPi per firm is equal across industries ($F = 2.8$, $df=9$, $p = 0.004$). All significant differences in pair-wise comparisons (indicated by bold font style and asterisks) involve either the entertainment industry or the petroleum refining industry, with one exception. The exception is that mean PROPi per firm for Food & Drug differs from that of Computers & Office Equipment industries.

In general, the analyses by industry indicate the XBRL C&I taxonomy does not accommodate reporting practices in two industries, Entertainment and Petroleum Refining, as well as it does the other industries studied. This suggests a need to develop industry-specific extensions to the taxonomy to increase the likelihood of adoption, and to maximize the potential benefits of XBRL as a shared system. These extensions need to be consistent with the existing taxonomy to maintain semantic consistency across its levels and across contexts of intended use.

Analysis also confirms that significant differences exist among the three main financial statements in the ability of the XBRL taxonomy to accommodate firms' reporting practices. As discussed previously, the *type* of exceptions also differ by financial statement. The taxonomy might better fit current reporting practices if it provided a less disaggregated BS branch, and more disaggregated IS and SCF branches.

Summary and Conclusions

We have presented a descriptive analysis of the fit between the year 2000 taxonomy for XBRL financial reporting under U.S. GAAP and the balance sheet, income statement and cash flow statement line items for 67 companies from 10 industries for 1999. Future work should examine a broader sample of industries and a larger number of companies within industries to

fully assess the implications of industry practices for taxonomies. We recommend that at least 1,000 companies should be sampled and analyzed by SIC code to provide information relevant to industry-specific taxonomies. Such taxonomies should largely be grounded in firms' current reporting practices, so long as they comply with GAAP.

Line items difficult to map to taxonomy tags were designated 'special attention' items (SAIs) and assigned to one of three types: suggested new taxonomy tags, two or more tags grouped together, or firm-specific items. There is a difference in the degree of information aggregation provided in the balance sheet, income statement and cash flow statement sections of the taxonomy, and this is reflected in the significantly differing numbers of SAIs by financial statement.

The results indicate that the C&I taxonomy provides a good fit overall, and a significantly better fit for some industries than for others, suggesting the need for current and expected efforts to develop industry-specific taxonomies. Overall, the number of SAIs is modest relative to the total of line items tagged, even before such ameliorating steps are taken (i.e., typically less than 15%). We view as encouraging the finding that firms scoring higher on a measure of likely interest in adopting XBRL (WEBi) also generated fewer SAIs and therefore are better fits with the current taxonomy.

XBRL.org are considering the findings of this study in their update to the C&I taxonomy, currently in progress. Regarding the findings, David Vun Kannon, co-chair of the XBRL.org Specification Group has said: "I think that the information provided by this study to the US Domain working group will be very valuable for updating the C&I taxonomy." Additional comments about specific findings, in personal communications with other XBRL.org members have been similarly positive.[19]

The study raises some questions that should be addressed by future work. Rather than being extensions of this study, the questions mostly concern how XBRL will function in the future. First, can business information taxonomies generally be laid out in hierarchical or 'parent/child' format? Our priors are that if this is possible anywhere in business, it should be possible in creating taxonomies for financial statements. Yet we identified at least one situation (not discussed) where the hierarchical nature of the taxonomy proved an impediment to preserving information firms desire to disclose. Another obvious and important question is, 'Who is going to create, who is going to maintain, and who is going to coordinate among, the various taxonomies current or prospective (i.e., International, general U.S. GAAP, industry-specific U.S. GAAP, firm-specific (managerial), U.S. federal tax reporting, etc.)?' At present it appears the AICPA will create and support an organization to maintain and disseminate XBRL for U.S. GAAP. It is not yet clear how other issues raised above will be resolved.

One promise of XML and of XBRL is the ability to port the tagged data into a multitude of other formats and uses. The ideal result should be the same in all cases: an XBRL-tagged value from the financial reporting domain that is interpretable and usable (comparable and consistent) when imported into another domain (for example, in tax reporting). If substantially different information is required in the new domain than is provided through the financial reporting taxonomy, then the end result will be sub-optimal. This points to other areas of future work: evaluating semantic consistency of XBRL taxonomies with other domain- or purpose-specific uses of XBRL-tagged information, reconciling the XBRL taxonomies with other taxonomies, and developing intelligent software tools to act at the interface between domains defined by these taxonomies. Of the many potential areas where XBRL-tagged data may be useful, we feel one in particular needs attention: taxonomies or models of reporting at the

transaction level need to be combined with XBRL to maximize the internal benefits of the XBRL taxonomy to the firm, and to enable XBRL for purposes of continuous auditing[20]. Lastly, XBRL holds great promise for international reporting purposes, but differences in language, practice and process complicate the issues raised above. Their resolution, whether through regulatory groups, standards bodies or software tools, will be necessary to obtain the full value of XBRL.

References

Baker. A. 2000. Virtual Discussion on Virtual Disclosure. Investor Relations (October): 55-57.

Ettredge, M., S.Y. Kwon, and D. Smith. 2002. Competitive Harm and Companies' Positions on SFAS No. 131, Journal of Accounting Auditing and Finance (forthcoming).

Ettredge, M., V. Richardson, and S. Scholz. 2002. Determinants of Voluntary Dissemination of Financial Data at Corporate Web Sites, Journal of Accounting and Public Policy (forthcoming).

Financial Accounting Standards Board (FASB). 1980. Statement of Financial Accounting Concepts No. 2.

Hampton, L, and vun Kannon, D. 2001. Extensible Business Reporting Language (XBRL) 2.0 Specification. http://www.xbrl.org/TR/2001/xbrl-2002-02-04.doc

Hartman, A., J. Sifonis, and J. Kador. 2000. Net Ready: Strategies for Success in the E-conomy. MacGraw-Hill, New York.

Hill, G. 2001. Slouching Tiger, Hidden Dragon. E-Company (May): 52-62.

Parish, A. 2000. Breaking the Ice. Investor Relations (October): 26-28.

Patel, K., and M. P. McCarthy. 2000. Digital Transformation: The Essentials of e-Business Leadership. McGraw-Hill, New York.

Seligman, D. 2000. 24-7 Accounting. Forbes (October 30): 147-148.

Sunder, S. 1997. Theory of Accounting and Control, Cincinnati, OH. South-Western Publishing.

Tapscott, D., D. Ticoll, and A. Lowy. 2000. Digital Capital: Harnessing the Power of Business Webs. Boston, MA. Harvard Business School Press.

Appendix
Sample Firms by Industry

Industry	Company
Beverages	Anheuser-Busch
	Brown-Forman
	Canandaiqua Brands
	Coca-Cola
	Coca-Cola Enterprises
	PepsiCo
	Whitman
Computer Software	Computer Associates Intl.
	Compuware
	Electronic Arts
	Microsoft
	Novell
	Oracle
	PeopleSoft
Computers and Office Equipment	Compaq Computer
	Dell Computer
	Hewlett-Packard
	International Business Machines
	Silicon Graphics
	Western Digital
	Xerox
Entertainment	AMFM
	CBS
	SFX Entertainment
	Time Warner
	USA Networks
	Viacom
	Walt Disney
Food and Drug	Albertson's
	CVS
	Kroger
	Safeway
	Walgreen
	Whole Foods Market

Appendix (continued)
Sample Firms by Industry

Food Services	CBRL Group
	Darden Restaurants
	Jack in the Box
	McDonald's
	Tricon Global Restaurants
	Wendy's International
General Merchandisers	J.C. Penney
	Kmart
	Sears Roebuck
	Target
	Value City
	Wal-Mart Stores
Motor Vehicles and Parts	Dana
	Johnson Controls
	Oshkosh Truck
	TRW
	Wabash National
Petroleum Refining	Chevron
	Conoco
	Crown Central Petroleum
	Exxon Mobil
	Murphy Oil
	Texaco
	USX
Pharmaceuticals	Allergan
	American Home Products
	Amgen
	Bristol-Myers Squibb
	Johnson & Johnson
	Merck
	Pfizer
	Phar-Mor

Figure 1
Influence Diagram

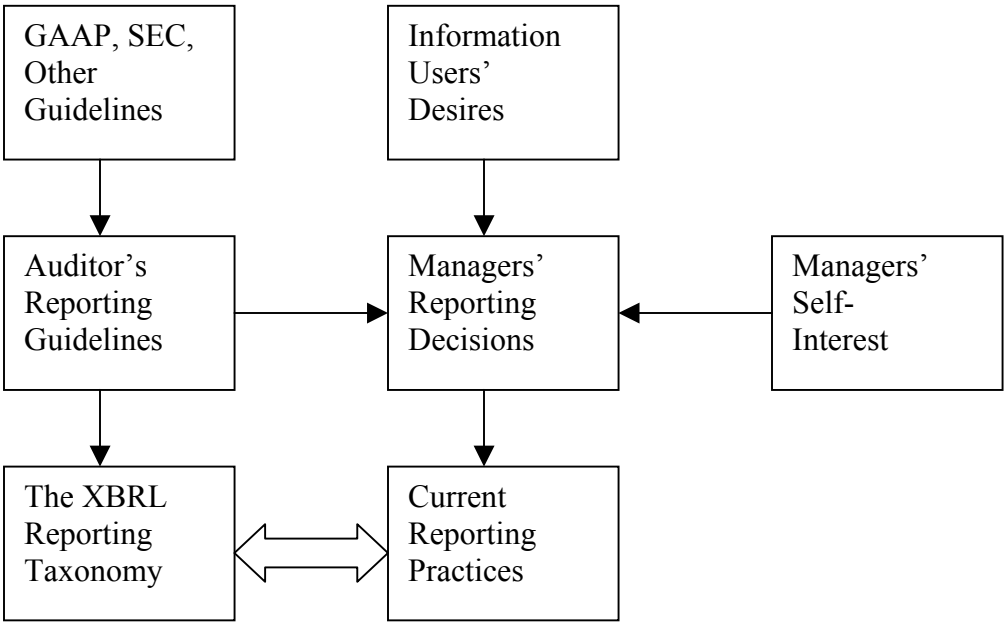

Figure 2
Fracturing of Taxonomic Parent-Child Relationship by Grouping Items

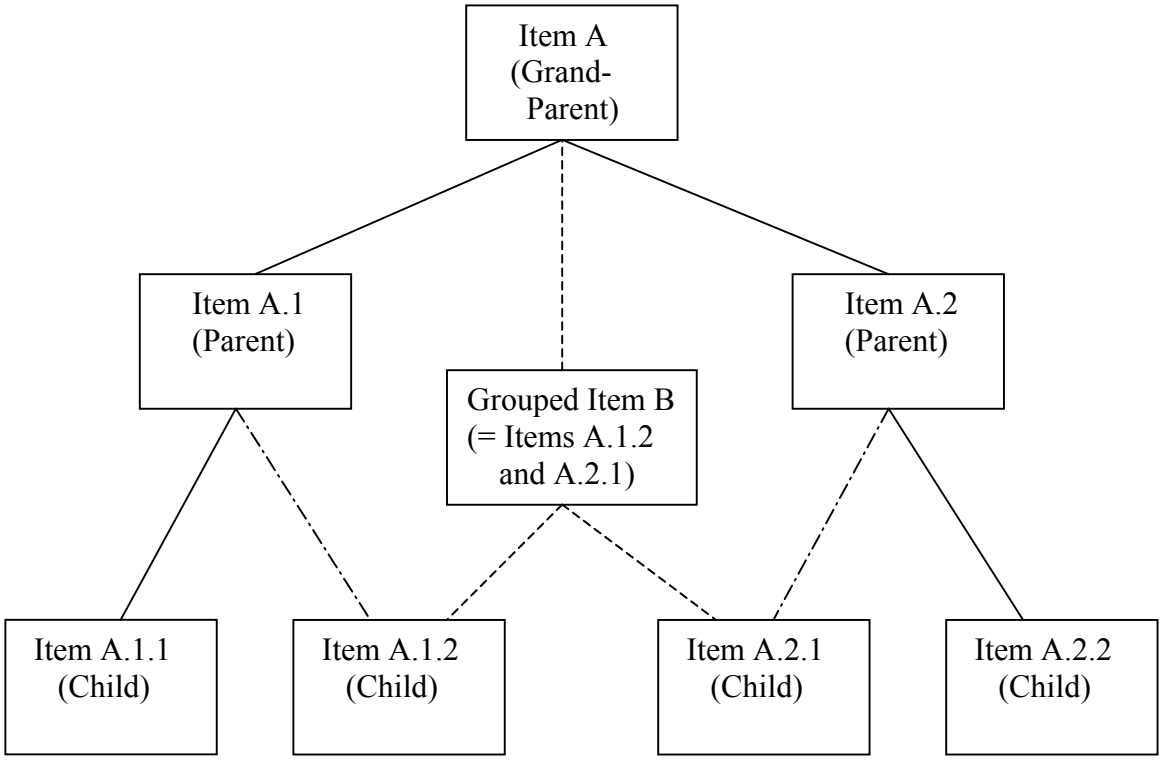

Table 1
Number of SAIs by Financial Statement and Type

Type	BS	IS	SCF	Totals
Grouped	110	47	73	230
New Tag	51	62	183	295
Firm Specific	58	105	177	340
Totals	219	214	433	866

Table 2
Number of SAIs by Type and Industry

	Grouped	New Tag	Firm Specific	Totals
Beverages	22	26	30	78
Computer & Office Equipt	31	17	33	81
Computer Software	24	40	31	95
Entertainment	12	29	85	126
Food & Drug	13	36	27	76
Food Services	15	28	31	74
General Merchandisers	29	34	33	96
Motor Vehicles & Parts	16	29	3	48
Petroleum Refining	41	32	52	125
Pharmaceuticals	27	25	15	67
Totals	230	296	340	866

Table 3
Number of SAIs by Financial Statement and Web Score

Web Score Quintile	BS	IS	SCF	Totals
WEBQ5	30	20	62	112
WEBQ4	40	31	72	143
WEBQ3	49	38	91	177
WEBQ2	44	74	101	218
WEBQ1	51	43	82	176
Totals	214	206	408	826

Table 4
Number of SAIs by Type and Web Score

Web Score Quintile	Grouped	New Tag	Firm Specific	Totals
WEBQ5	43	36	33	112
WEBQ4	41	62	40	143
WEBQ3	49	60	68	177
WEBQ2	57	54	107	218
WEBQ1	38	68	70	176
Totals	228	280	318	826

Table 5
Distribution of SAIi and PROPi per Firm and per Financial Statement

	Mean	S.Dev.	Min.	25th %	Median	75th %	Max.
SAIi per Firm	12.93	6.36	3	8.5	12	16	33
SAIi per BS	3.27	2.51	0	2	3	4	12
SAIi per IS	3.19	2.79	0	1	3	4	13
SAIi per SCF	6.46	3.62	1	3	6	9	15
PROPi per Firm	14%	2%	4%	11%	14%	18%	29%
PROPi per BS	10%	8%	0%	6%	9%	12%	43%
PROPi per IS	14%	11%	0%	5%	13%	18%	59%
PROPi per SCF	19%	8%	4%	12%	18%	26%	37%

Test of: mean SAIi per firm equals zero: $t = 18.2$, sig. $p < 0.001$

Test of: mean PROPi per firm equals zero: $t = 20.9$, sig. $p < 0.001$

Test of: mean SAIi per statement is equal across financial statements:
$F = 27.7$, $(df = 2)$, $p < .001$

Test of: mean PROPi per statement is equal across financial statements:
$F = 16.0$ $(df = 2)$, $p < .001$

Table 6
Distribution of PROPi per Firm by Industry

	Mean	S.Dev.	Min.	25th %	Median	75th %	Max.
Beverages	12%	7%	5%	8%	11%	16%	28%
Computer & Office Equipt	13%	11%	0%	6%	10%	16%	36%
Computer Software	16%	7%	5%	10%	15%	24%	26%
Entertainment	18%	12%	0%	11%	15%	21%	59%
Food & Drug	11%	10%	0%	4%	8%	14%	33%
Food Services	14%	10%	0%	7%	12%	17%	33%
General Merchandisers	15%	8%	0%	11%	15%	18%	37%
Motor Vehicle & Parts	12%	3%	8%	9%	12%	13%	18%
Petroleum Refining	20%	12%	0%	13%	21%	27%	45%
Pharmaceut-icals	11%	7%	0%	7%	10%	17%	23%

Test of: mean PROPi is equal across industries:
F = 2.8 (df = 9), p = .004

Test of: mean SAIi is equal across industries (data for SAIs are not tabulated):
F = 2.5 (df = 9), p = .011

Test of: mean SAIi is equal across industries with LINEi as covariate (data for SAIs are not tabulated):
F = 2.1 (df = 9), p = .028

Note: with LINEi as a covariate, there is a significant industry x statement interaction:
F = 1.7 (df = 18), p = .036

Table 7
Correlation Matrix

	SAI per Firm	PROPi per Firm	SALEi per Firm
PROPi per Firm	0.91**	-	
SALEi per Firm	-0.04	-0.03	-
WEBi per Firm	-0.16*	-0.12	0.01

Each cell contains estimated Pearson correlation coefficient. Non-parametric correlations display the same pattern with stronger significance.

*** Differs from zero at p < 0.05 (2-tailed).**

**** Differs from zero at p < 0.01 (2-tailed).**

Table 8.
Pair-wise Comparisons of mean PROPi by Statement

STATEMENTS	**Column J**		
Row I	BS	IS	SCF
BS	-		
IS	3.8%*	-	
SCF	8.3%*	4.5%*	-

Each cell contains the mean differences between row I and column J for PROPi (e.g. mean PROPi for IS - mean PROPi for BS = 3.8%).

***Significant difference at p≤ 0.05**

Table 9.
Pair-wise comparisons of mean PROPi per Firm by industry

	Column J								
Row I	Beverages	Computer Software	Computers & Office Equipt.	Entertainment	Food & Drug	Food Services	General Merchandisers	Motor Vehicles & Parts	Petroleum Refining
Computer Software	1.0%	-							
Computers & Office Equipt.	3.7%	2.7%	-						
Entertainment	**5.3%***	4.3%	1.6%	-					
Food & Drug	-1.6%	-2.6%	**-5.3%***	**-6.9%***	-				
Food Services	1.2%	0.2%	-2.5%	-4.1%	2.8%	-			
General Merchandisers	3.1%	2.1%	-0.5%	-2.1%	4.8%	2.0%	-		
Motor Vehicles & Parts	-1.4%	-2.4%	-5.0%	**-6.7%***	-0.2%	-2.6%	-4.5%	-	
Petroleum Refining	**8.0%***	**7.0%***	4.4%	2.8%	**9.7%***	**6.9%***	4.9%	**9.4%***	-
Pharmaceuticals	-1.0%	-2.0%	-4.7%	**-6.3%***	0.6%	-2.2%	-4.1%	0.4%	**-9.0%***

Each cell contains the mean difference between row I and column J
(e.g. mean PROPi for Computer Software - mean PROPi for Beverages = 1.0%).

***Significant at $p \leq 0.05$.**

End Notes

[1] While many technologies exist for the purpose of information exchange, XML is platform independent, relatively simple, and thus widely applicable.

[2] Visit the XBRL Web site at www.xbrl.org for information about the consortium of firms and organizations that are supporting the development and implementation of XBRL.

[3] Visit the XBRL Web site at www.xbrl.org for descriptions of numerous other potential uses.

[4] In order to obtain this capability currently, a firm might be required to restructure its organization, re-engineer its processes, and employ uniform firm-wide ERP software.

[5] See http://www.xbrl.org/, http://www.xbrl.org/TR/TechnicalOverview.htm, and Hampton and vun Kannon (2001) for further detail.

[6] The nomenclature used here is from the current XBRL C&I Taxonomy, using the dot-delimited parent.child, lower camel case descriptive naming to uniquely specify taxonomic elements. Recent changes to the XBRL specification have elaborated the hierarchic business rules defining parent-child relationships between elements in the taxonomy into three separate elements. In addition, these rules and the label and reference information for each element, previously embedded in element definitions found in the XBRL Schema, have been separated into Xlink-compliant linkbases. See XBRL Specification 2.0, Hampton and vun Kannon (2001) for further detail and officially sanctioned examples.

[7] This information was provided to the authors by one of the original developers of the XBRL taxonomy.

[8] U.S. GAAP consists primarily of authoritative pronouncements issued by the Financial Accounting Standards Board. They publish these *Original Pronouncements* in several volumes annually.

[9] Financial software will not recognize the synonyms unless specifically written to do so. XBRL-enabled software will, however, recognize as similar each account associated with the same tag from a given taxonomy. Thus XBRL potentially provides semantic consistency for electronic business reporting.

[10] If firms keep their accounting records up-to-date and 'close their books' promptly after each fiscal period, digital documents created from these files using XBRL will be timely.

[11] See Ettredge, Kwon and Smith (2002) for a study of managers' incentives to aggregate and dis-aggregate information about business segments.

[12] The FASB states (FASB, 1980) that comparability enables users to identify real similarities in economic phenomena. Comparability also supports other useful attributes of information such as decision usefulness and reliability.

[13] Information has *predictive* value if it improves decision-makers' abilities to predict. It has *feedback* value if it enables them to confirm or correct earlier predictions.

[14] Whereas most industry categories contain seven firms, Motor Vehicles and Parts represents five and Food Services contains six due to companies being dropped from the sample.

[15] Use of the 'Other' tag potentially results in loss of information.

[16] The XBRL taxonomy task force has adopted these recommendations for post-year-2000 versions of the C&I taxonomy.

[17] If so, these items are candidates for addition to future industry-specific taxonomies.

[18] These two firms are generally non-comparable to all other sample firms so we dropped them from the sample.

[19] Positive comments have been received regarding, for example, the identification of missing tags, the examination of special attention items in the balance sheet and income statement in particular, the suggestions for new tags for the statement of cash flow, and the identification of problems with grouped items.

[20] XBRL General Ledger (XBRL GL, http://www.xbrl.org/gl/gl.htm), released for public comment on January 16, 2002, and ebXML (http://www.ebxml.org/) are examples of two applications addressing this gap.