Data differences —XBRL versus Compustat

Pavani Tallapally
Slippery Rock University

Michael S. Luehlfing
Louisiana Tech University

Madhu Motha
Butler County Community College

Abstract

Given their proprietary data standardization process, Compustat (www.compustat.com) provides accounting data which may differ, to some degree, from accounting data provided by some individual companies via XBRL (eXtensible Business Reporting Language).   In this regard, the purpose of this study is to analyze the extent of such differences, if any, between accounting data provided by Compustat and accounting data provided by (via) XBRL. The results suggest that differences exist and the reconciliation of such differences is not obvious.

Keywords: Compustat, Data Aggregators, Data Intermediaries, EDGAR/IDEA, Taxonomies, XBRL
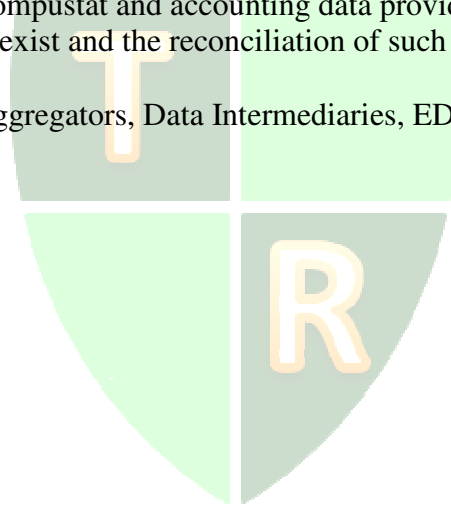
INTRODUCTION

Third-party data aggregators (a.k.a. data intermediaries) such as Compustat (www.compustat.com) provide accounting data to financial professionals as well as academic researchers. Miguel, 1977, Kinney and Swanson, 1993, and, Yang et al. 2003, suggest that accounting data provided by Compustat differs, to some degree, from accounting data provided by some individual companies. A possible competitor to Compustat is XBRL (eXtensible Business Reporting Language) which is a worldwide standard for the publishing, exchange, and analysis of financial reports and data (www.xbrl.org). Quoting Charles Hoffman, CPA (a.k.a., "The father of XBRL"), Tie (2005) reports that "XBRL will significantly improve the ability of CPA financial managers to distribute information to stakeholders precisely as reported, rather than as condensed or otherwise modified by third-party data aggregators to facilitate distribution." In contrast to XBRL, Compustat edits the "as reported" accounting data of an individual company by using a proprietary data standardization process. Thus the accounting data (of a specific company) provided by Compustat may differ, to varying degrees, from the accounting data provided by that same specific company (via XBRL or via a traditional mechanism such as annual reports). In this regard, Kern and Morris, 1994 suggest that "[a]nalysts and researchers need to exercise great care when selecting databases and variables from those databases. These choices can affect the results of and the inferences drawn from empirical research in ways more than is anticipated by researchers." Motivated by the concerns of Kern and Morris, 1994 (as well as the assertions of Charles Hoffman, CPA), this study seeks to gain insights into the issue (i.e., modified versus "as reported" accounting data) by analyzing the extent of differences, if any, between accounting data provided by Compustat and accounting data provided by XBRL.

BACKGROUND AND MOTIVATION

Traditionally, individual companies provide accounting data in a non-standardized manner. Stated otherwise, while accounting data must comply with Generally Accepted Accounting Standards (GAAP) and/or some other guidelines, there remains a certain amount of flexibility in the reporting of such data. To enhance comparability, Compustat edits the "as reported" accounting data of a company by using a proprietary data standardization process. This process is described below (www.compustat.com).

> Our internal research team rigorously examines original company sources by carefully extracting financial information, removing reporting biases and reconciling data discrepancies. After collecting data from diverse sources, we standardize it by financial statement and by specific data item definition, preparing information that is comparable across companies, industries, time periods and sectors. This standardized presentation makes it easier to identify companies with similar characteristics, such as capital structure and operating performance and is designed to complement how the data [are] used. Additionally we analyze financial statement notes to provide detailed breakouts to gain additional insight overlooked by other companies.

In essence, Compustat transforms data extracted from 10-K's or annual financial statements by using a proprietary data standardization process which includes the use of proprietary data item definitions. These proprietary data item definitions are fully disclosed on Compustat's website (www.compustat.com).

Unlike Compustat, XBRL "is an open standard, free of license fees, being developed by a non-profit making international consortium (www.xbrl.org)." Additionally, rather than editing "as reported" data as is done by Compustat, XBRL underpins the actual development of the "as reported" accounting data in conjunction with a GAAP-based taxonomy (Bovee et al. 2002, Bovee et al. 2005, Piechocki et al. 2009, Zhu and Wu 2010). This GAAP-based taxonomy is described in the following excerpt (www.xbrl.org):

> The U.S. GAAP Financial Reporting Taxonomy is a list of computer-readable tags in XBRL that allows companies to tag precisely the thousands of pieces of financial data that are included in typical long-form financial statements and related footnote disclosures. The tags allow computers to automatically search for, assemble, and process data so it can be readily accessed and analyzed by investors, analysts, journalists, and regulators.

As suggested below, it is the aforementioned tagging process that is critical to XBRL's functionality (www.xbrl.org):

> The idea behind XBRL, eXtensible Business Reporting Language, is simple. Instead of treating financial information as a block of text - as in a standard internet page or a printed document - it provides an identifying tag for each individual item of data. This is computer readable. For example, company net profit has its own unique tag.

Recognizing the potential of XBRL to enhance financial reporting, the Securities and Exchange Commission (SEC) mandated the use of XBRL. Specifically, WebCPAStaff (2009) reports that…

> [t]he XBRL mandate is part of the SEC's planned move to an Interactive Data Electronic Applications, or IDEA, system to replace its old EDGAR [Electronic Data Gathering, Analysis, and Retrieval] system for storing and accessing financial filings.

> The first required submissions will be for quarterly reports containing financial statements for a fiscal period ending on or after June 15, 2009. For calendar-year companies, this requirement will first apply to their June 30, 2009, quarterly reports filed on Form 10–Q.

Intuitively, since XBRL is used to develop the "as reported" accounting data provided by a specific company to the SEC, then XBRL appears positioned to fully support the EDGAR/IDEA activities of the SEC. In contrast, it appears that Compustat is not necessarily fully positioned to support the EDGAR/IDEA activities of the SEC since Compustat provides "standardized" financial statement data (which is not necessarily the equivalent of the XBRL

financial statement data). With respect to these intuitions, the following question is asked: To what extent, if any, is Compustat accounting data different from XBRL accounting data?

METHODOLOGY AND RESULTS

The analysis focused on the sales/revenue amounts reported by both Compustat (via Research Insight) and XBRL (via the EDGAR/IDEA website) to gain insights as to the extent, if any, that Compustat accounting data differs from XBRL accounting data. Specifically, the sales/revenue amounts of twenty-seven non-financial companies included in the DOW 30 companies were compared. The three financial companies were excluded from the analysis since these companies used the "Banking and Savings Institutions" XBRL Taxonomy not the "Commercial and Industrial" XBRL Taxonomy as did the other twenty-seven companies in the sample. Compustat and XBRL accounting data were collected for each company in the sample for the years ending after June 30, 2009 but before July 1, 2010. The results of the initial comparisons are reported in Table 1 (Appendix).

As indicated in Table 1, Compustat reported lower sales/revenue amounts (than XBRL) for the following six companies: Chevron, Dupont, Exxon, Home Depot, Pfizer and United Technologies. In contrast, Compustat reported higher sales/revenue amounts for one company—Wal-Mart. Equal sales/revenue amounts were reported (by both Compustat and XBRL) for the other twenty companies included in the sample. An analysis of aggregate data differences when compared with XBRL data is provided in Table 2 (Appendix).

As indicated in Table 2, the aggregate average difference was 5.19% for the six companies where the XBRL reported sales/revenue amounts were higher than the Compustat reported amounts. In contrast, the difference for the one company where Compustat reported a higher sales/revenue amount than XBRL was 0.26%. Overall, the aggregate average difference for the entire sample of twenty-seven companies was only 1.47%. Admittedly, while 1.47% may appear rather low and thus rather appealing from a portfolio perspective, the 1.47% resulted in large part to the fact that no differences between the Compustat accounting data and the XBRL accounting data were noted for twenty of the twenty-seven companies in the sample. Accordingly, an individual analysis of difference on a company by company basis was performed (since investors and creditors normally make company specific decisions). The results of the analysis of individual company differences are provided in Table 3 (Appendix).

As indicated in Table 3, differences between Compustat and XBRL reported sales/revenue amounts existed with respect to seven of the twenty-seven companies (approximately 26%) included in the sample. Fortunately, the (absolute dollar value) differences noted were not material (assuming a 1% hypothetical level of materiality based on individual company XBRL sales/revenue threshold) for five of the seven companies. Unfortunately, differences with respect to two of the seven companies (both oil companies) were well in excess of the hypothetical one percent materiality threshold. For all seven companies where differences were noted, attempts were made to reconcile the differences noted between sales/revenue amounts reported by XBRL and Compustat using other "same year" income statement line items related to each company—as well as by reference to the relevant data definitions provided by XBRL (http://xbrl.us/taxonomies/Pages/US-GAAP2009.aspx) as well as Compustat (www. compustat.com). Unfortunately, with one exception (Exxon Mobil), reconciliations were not obvious (possible) relating to the differences noted between the Compustat and XBRL accounting data. Fortunately, the net income amounts reported by both

Compustat and XBRL were the same (i.e., were equal) for each company in the sample. Thus it is assumed that the differences noted in reported sales/revenue amounts were offset by the cumulative (but unobservable differences) related to components of other income statement debit and/or credit balances. Given that differences exist, and given that the differences (are assumed to) relate to the proprietary accounting data standardization process employed by Compustat (since the EDGAR/IDEA data were developed for each individual company using XBRL technology), then it appears that Compustat is not fully positioned to support the EDGAR/IDEA activities of the SEC.

IMPLICATIONS AND CONCLUSIONS

Given that the results of the study indicate that differences exist between the sale/revenue data reported by XBRL and Compustat, the concerns of Kern and Morris (1994) as well as the assertions of Charles Hoffman, CPA are supported. Accordingly, researchers as well as practitioners should use caution when selecting data sources. Admittedly, sample in the study was limited in nature. However, there is no reason to believe that an expanded sample would yield significantly different results. Future researchers may want to focus on the "costs versus benefits" of using "purchased databases" (such as Compustat) versus "free databases" (such as EDGAR/IDEA via XBRL).

REFERENCES

Bovee, M., Ettredge M., Srivastava, R. P., and Vasarhelyi M, A. (2002). Does the Year 2000 XBRL Taxonomy Accommodate Current Business Financial-Reporting Practice? *Journal of Information Systems.* 16(2), 165-182.

Bovee, M., Ettredge, M., Srivastava, R. P., and Vasarhelyi M, A (2005). Financial Reporting and Auditing Agent with Net Knowledge (FRAANK) and eXtensible Business Reporting Language (XBRL) . *Journal of Information Systems.* 19(1), 19-41.

Kern, B. B, and Morris, M. H. (1994). Differences in the COMPUSTAT and Expanded Value Line Databases and the Potential Impact on Empirical Research. *Accounting Review*. 69(1), 274-284.

Kinney, M. R. and Swanson, E.P. (1993). The Accuracy and Adequacy of Tax Data in COMPUSTAT. *Journal of the American Taxation Association*. 31(3), 309-35.

Miguel, J.G. (1977). The Reliability of R&D Data in Compustat and 10-K Reports. *The Accounting Review*. (July), 638-641.

Piechocki, M., Felden, C., Graning, A., and Debreceny, R. (2009). Design and Standardisation of XBRL Solutions for Government and Transparency. *International Journal of Disclosure and Governance*. 6(3), 224-240.

Tie, R. (2005). XBRL: It's Unstoppable. *Journal of Accountancy*. 200(2), 32-35.

WebCPAStaff (2009.) SEC Begins XBRL Mandate for First 500 Companies. (February 11, 2009). Available at http://www.accountingtoday.com/news/30734-1.html

Yang, D. C., Vasarhelyi, M. A, & Liu, C. (2003). A Note on the Using of Accounting Databases. *Industrial Management & Data Systems*. 103 (3/4), 140-145.

Zhu, H., and Wu, H. (2010). XBRL and Interoperability of Financial Statements in the U.S. *Working Paper.* (March 25, 2010). Available at SSRN: http://ssrn.com/abstract=1581511

Table 1 - Initial Comparisons (Compustat Versus XBRL)

Company	Year End	COMPUSTAT	XBRL	Difference
3M Company (MMM)	12/31/09	$23,123.00	$23,123.00	$0
Alcoa Inc. (AA)	12/31/09	$18,439.00	$18,439.00	$0
American Express Co. (AXP)	12/31/09	#	#	#
AT&T Inc (T)	12/31/09	$123,018.00	$123,018.00	$0
Bank of America Corp. (BAC)	12/31/09	#	#	#
Boeing Co. (BA)	12/31/09	$64,306.00	$64,306.00	$0
Caterpillar, Inc. (CAT)	12/31/09	$32,396.00	$32,396.00	$0
Chevron Corp. (CVX)	12/31/09	$159,293.00	$167,402.00	-$8,109
Cisco Systems Inc. (CSCO)	7/31/09	$40,040.00	$40,040.00	$0
Coca-Cola Co. (KO)	12/31/09	$30,990.00	$30,990.00	$0
DuPont (E.I.) deNemours (DD)	12/31/09	$27,268.00	$27,328.00	-$60
Exxon Mobil Corp. (XOM)	12/31/09	$275,564.00	$301,500.00	-$25,936
General Electric Co. (GE)	12/31/09	$155,777.00	$155,777.00	$0
Hewlett-Packard Co. (HPQ)	10/31/09	$114,552.00	$114,552.00	$0
Home Depot, Inc. (HD)	1/31/10	$65,955.00	$66,176.00	-$221
Intel Corp. (INTC)	12/31/09	$35,127.00	$35,127.00	$0
International Bus. Mach. (IBM)	12/31/09	$95,758.00	$95,758.00	$0
J.P. Morgan Chase & Co. (JPM)	12/31/09	#	#	#
Johnson & Johnson (JNJ)	12/31/09	$61,897.00	$61,897.00	$0
Kraft Foods Inc. Cl A (KFT)	12/31/09	$40,386.00	$40,386.00	$0
McDonalds Corp. (MCD)	12/31/09	$22,744.70	$22,744.70	$0
Merck & Co. Inc. (MRK)	12/31/09	$27,428.30	$27,428.30	$0
Microsoft Corp. (MSFT)	6/30/10	$62,484.00	$62,484.00	$0
Pfizer Inc. (PFE)	12/31/09	$49,934.00	$50,009.00	-$75
Procter & Gamble Co. (PG)	6/30/10	$78,938.00	$78,938.00	$0
Travelers Cos. Inc (TRV)	12/31/09	$24,680.00	$24,680.00	$0
United Technologies Corp. (UTX)	12/31/09	$52,810.00	$52,920.00	-$110
Verizon Communications Inc. (VZ)	12/31/09	$107,808.00	$107,808.00	$0
Wal-Mart Stores, Inc. (WMT)	1/31/10	$406,103.00	$405,046.00	$1,057
Walt Disney Co. (DIS)	9/30/09	$38,063.00	$38,063.00	$0

Note: # implies that this company was excluded from the analysis since this company used the "Banking and Savings Institutions" XBRL Taxonomy not the "Commercial and Industrial" XBRL Taxonomy as did the twenty-seven companies included in the sample.

Table 2 - Analysis of Aggregate (Average) Differences between Compustat and XBRL

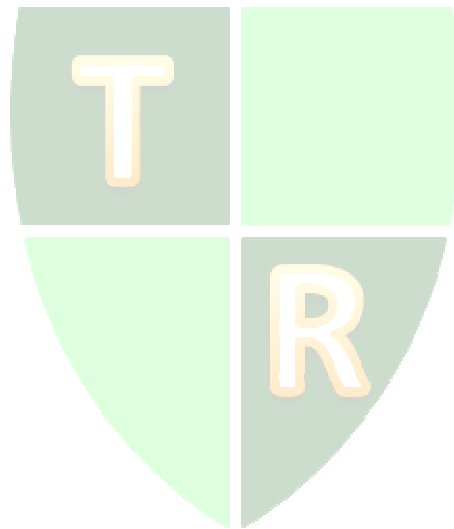| Sample/Subsample Description	Average (or Actual) Dollar Difference	Percentage of Relevant XBRL Sales/Revenue Total
All 27 Companies	-$33,454*	1.47%*
All 7 Companies with Differences	-$33,454*	3.13%*
6 Companies—Compustat < XBRL	-$34,511*	5.19%*
1 Company—Compustat > XBRL	$1,057	0.26%

Note: * indicates net differences

Table 3 - Analysis of Individual Company Differences between Compustat and XBRL

Individual Company Name	Dollar Difference	Percentage of XBRL Sales/Revenue Amount
Chevron Corp. (CVX)	-$8,109	-4.84%
DuPont (E.I.) deNemours (DD)	-$60	-0.22%
Exxon Mobil Corp. (XOM)	-$25,936	-8.60%
Home Depot, Inc. (HD)	-$221	-0.33%
Pfizer Inc. (PFE)	-$75	-0.15%
United Technologies Corp. (UTX)	-$110	-0.21%
Wal-Mart Stores, Inc. (WMT)	$1,057	0.26%