Spring 5-2-2019

The Quality of XBRL Structured Financial Statements: An Empirical Examination of Custom Tags

Aidana Razhap kyzy
CUNY Hunter College

The quality of XBRL structured financial statements: an empirical examination of custom tags.

By

Aidana Razhap kyzy

Submitted in partial fulfillment
of the requirements for the degree of
Master of Arts in Economics, Hunter College
The City University of New York

2019

Thesis Sponsor:

May 2, 2019	Christine Tan
Date	Signature
May 2, 2019	Devra Golbe
Date	Signature of Second Reader

Abstract

In 2009 the US Securities and Exchange Commission (SEC) adopted the eXtensible Business Reporting Language (XBRL) system to improve the process by which financial statements can be used. Interactive financial data filed with the SEC using XBRL provides easily readable and comparable financial data, thereby improving transparency and efficiency in the corporate market. SEC rules permit companies to use custom tags in their financial reports in cases when an appropriate element cannot be found in the Financial Accounting Standards Board (FASB) standard XBRL taxonomy. The inordinate use of custom tags may result in a reduction of financial report quality by diminishing the comparability and usability of filings by investors and analysts. Using XBRL-based empirical data from 2015 to 2017 fiscal years, this paper explores the inordinate use of custom tags. Do high uses of custom tags result from the complexity of a company's operational structure or are they used deliberately by managers attempting to manipulate their financial disclosures? I find that the use of custom tags is positively related to variables indicating the lower quality of financial reports.

Contents

1. Introduction

The mission of the U.S. Securities and Exchange Commission (SEC) entails the enforcement of transparency within the US corporate sector by maintaining fair, orderly and efficient markets, facilitating capital formation and protecting its investors (SEC, 2019). The SEC's role in regulating US stock markets and providing accurate and consistent information regarding financial disclosure and profitability is essential to the strong functioning of the US economy. Since 2009, the SEC has focused on expanding its data organization technology and has transitioned to the use of structured data through eXtensible Business Reporting Language (XBRL). XBRL is a data standard used for structuring information contained in unstructured financial statements. The SEC implementation of XBRL, however, allows for the potential of compromised data collection by permitting the use of unstandardized data input, thereby reducing the comparability of financial data analysis. This paper disambiguates and identifies factors that correlate with inordinate use of custom tags within firms that file with the SEC, thus contributing to the goal of greater transparency in US markets.

XBRL is implemented to facilitate automated production and consumption of high-volume data measuring business performances (Liu, Luo, Sia, O'farrell, Teo, 2014). XBRL's structure allows for immediate processing and web access of corporate information, offering a computerized language for financial reporting that can be easily read, extracted and compared by computers. This system improves transparency of corporate filings increasing financial market efficiency and supplying important data for construction of decision models (SEC, 2006). Transitioning from traditional financial reports to XBRL is reliant on the US Generally Accepted Accounting Principles (US GAAP) taxonomy maintained and updated annually by the Financial Accounting

Standards Board (FASB). The FASB taxonomy provides a comprehensive dictionary of business reporting elements (tags) that represent accounting concepts necessary for the disclosure of financial information. SEC rules, however, permit companies to deviate from the FASB Taxonomy in cases when an appropriate element cannot be found to tag a particular disclosure in the company's filing. The company's taxonomy can be created using firm-specific custom tags also known as 'extensions'. Extensions reduce the comparability of a company's financial information vis a vis other companies, because custom tags confound the interpretation of data by their computer programs.

Previous research finds that a divergence from the standard taxonomy can reduce comparability, creating complications for digesting financial information (Dhole, Lobo, Mishra, and Pal, 2015). Taxonomy extensions complicate interpretation of business data, since custom tags do not correspond with the standard taxonomy used by other companies in the same industry. Excessive use of custom tags may indicate two different things (perhaps simultaneously): (1) complex operations; and/or (2) low-quality of financial reports. In the first instance, the company's appropriate implementation of custom tags to disclose values not yet defined in the US GAAP taxonomy indicate correct application of the SEC rules. Another instance, however, may indicate equivocal uses of custom tags to mask questionable activities of the company resulting in low-quality of financial reports. This paper explores whether the use of custom tags correlates with the complexity of the business or is an intentional manipulation of reported values by corporate managers.

A high-quality financial report is identified by its readily understandable and clear disclosure of relevant information of the financial statements. Reliable financial reports must be

free of material error, faithfully representing all activities utilized by the company (FASB, 2019). These reports must be comparable to the financial information presented for other accounting periods and firms in the same industry (Bragg, 2018). Consistency in financial statements is imperative for maintaining economic and business stability. Distorted reporting may lead to suboptimal deployments of capital, misallocation of resources and potentially create an ambiguous information environment for financial analysts to conduct their investment analyses. Investors who rely on accuracy of financial reports can be negatively affected by investing in companies with inflated or questionable values. This in turn may skew significant business and strategic decisions based on a flawed interpretation of economic reality (SEC, 2006). Inaccuracy of financial statements due to either complexity or low-quality of financial reports may result in an erosion of confidence. A low-quality financial report refers to the company's attempt to conceal its bona fide business performance, reducing its comparability. Analyzing XBRL-based data, I examine the relationship between a firm's inordinate use of extensions, which may be indicative of complexity and/or low-quality of financial reporting.

Low-quality financial disclosures can have significantly detrimental effects on financial markets and engender shocks to the economy (SEC, 2006). The SEC posits that XBRL increases market efficiency by providing access to all financial information, expanding the scope of research for analysts and investors, thereby increasing overall market liquidity (Cong, Hao, and Zou, 2014). This paper hypothesizes that companies with low financial reporting quality, as well as those companies with complex corporate structures will use more custom tags. The results of my analysis are consistent with the hypothesis that firms with low-quality financial reporting use more custom

tags. The coefficients on the measures for complexity, however, show that complex firms are not necessarily associated with more custom tags.

The study proceeds as follows: Section 2 provides in-depth background information and a subsequent literature review. Section 3 describes the methodology and key variables used to derive results for the regression. Section 4 explains the data collection and Section 5 discusses the results and implications of findings. The paper concludes with a summary of results, limitations and further prospects for the research.

2. Literature Review

Twenty first century society is increasingly reliant on the manipulation of complex data as a means to produce information. This creates a need to integrate rigorous data analytics in the study of socio-economics factors impacting every aspect of life. Cloud based structured data has become essential for quantitative research. Establishing an environment where data is readily available for interpretation, it provides us with accessible and discrete bytes for analysis by both humans and computers, communicating information that will provide building blocks for political and economic policy. Structured data for these purposes is created using standards including XBRL which leverages XML syntax.

The US Securities and Exchange Commission mandates that public companies file annual financial status reporting. In 2009 the SEC adopted an XBRL system as a means to improve the usefulness of 10Q|K financial filings. The Final Rule No. 33-9002 (Interactive Data to improve Financial Reporting Rule), enforced by the SEC in late 2009, required all public companies filing with the SEC to use XBRL as a standard for business reporting. The use of XBRL is meant to simplify the exchange and communication of financial statement information.

XBRL-based financial data facilitates effective retrieval and use of financial information while also contributing to an automated regulation process conducted by the SEC. XBRL technology allows for the structuring of quantitative financial and business data through tags, using relevant components of over 20,000 standard official elements approved by the SEC as of 2019, in the US GAAP Taxonomy (FASB, 2019). In addition, the SEC allows for customized extension elements. While standard tags structured by the SEC, and the FASB reflect the disclosure requirements of the US GAAP and common financial reporting practices of companies, custom tags are to be used only if a standard tag does not fully explain/map to the disclosed value in the financial statements and footnotes.

The use of company-specific custom tags in disclosing financial data reduces the comparability of data across firms. Dhole, Lobo, Mishra and Pal (2015) identified the importance of standard tags in decreasing the complexity of accessible data and improving the financial market's information environment. Several studies report that standard tags improve comparability, reducing information asymmetry and processing costs while customized extensions increase it (Dhole, Lobo, Mishra and Pal, 2015; Tylor and Dzuranin, 2010). Studies focusing on identifying fraudulence in managerial behavior, show that accruals-based earnings management declined more for firms that use more standard taxonomy tags, suggesting that the use of extensions renders financial statements less transparent and more subject to earnings manipulation. These findings are consistent with the argument that extensions are harder to interpret (Felo, Kim and Lim, 2018; Li and Nwaeze, 2016; Kim, Kim and Lim, 2013).

These findings suggest that the use of XBRL in analysis reduces error and maximizes accuracy of results in analyzing data relevant to the functioning of markets and the economy. The

SEC is the main regulator of business and financial drivers of the US economy. Since the adoption of interactive data consumption within the SEC, filings have been regulated to provide a structured basis for simplified analysis. Information disclosed by XBRL thus becomes more accessible to financial statement users. Arnold, Bedard, Phillips and Sutton (2012) find that both professional and non-professional investors, in fact, make better use of XBRL tagged information. The format consisting of this type of data makes it easier for users to evaluate statement information in a time sensitive manner by simplifying access and comparison of filed information. This in turn translates into improved decision making by investors, increasing firm disclosure and market efficiency. However, adoption of detailed tagging[1] results in an increased number of standard as well as custom XBRL tags. This may increase the opportunity costs of processing XBRL financial information, since a large number of custom tags may result in the ignoring of essential financial information that is difficult to compare or process. Other negative results of relatively high use of custom tags include an increase in processing costs (slower preparation time and accuracy of analyzing financial data) for investors and analysts and increased potential for manager's manipulation of financial information (Li and Nwaeze, 2015).

This research contributes to insights into the usage of custom tags. Is it the complexity of a company or possible managerial manipulation that drives up the percentage of custom tag usage in a firm? This present research contributes to the area of financial reporting investigation by showing that the use of custom tags may be an additional indicator of aberrant behavior in publicly traded companies. I interrogate the economic and empirical implications of measuring complexity and the low-quality of financial disclosures of public firms.

[1] Detailed tagging indicates the tagging of quantitative amounts disclosed in the footnotes of the financial statements.

3. Methodology

Utilizing linear regression model with clustered standard errors (CSE), I estimate the percentage of custom tags and the dollar value tagged by extensions on measures of complexity and low-quality financial disclosures. To validate the research hypothesis, t-statistics and p-values are used to examine the significance of the regression models and test their coefficients, indicating the relationship between the types of companies and the amount of extensions they use. To address possible heterogeneity in the data, the model specifies a robust clustering of standard errors.

Proxies for complexity of a firm include a fraction of intangible assets to total assets, and the number of segments in each company. Proxies for low-quality financial disclosures are: 'unqualified' auditor's opinion and two measures of material weaknesses. Variables are defined and discussed in the next section.

1. $Custom_{it} = \beta_0 + \beta_1 lnIntan\_ta_{it} + \beta_2 Unqualified_{it} + \beta_3 BigFour_{it} + \beta_4 Segment_{it} + \beta_5 Weakness_{it} + \varepsilon_{it}$
2. $Custom_{it} = \beta_0 + \beta_1 lnIntan\_ta_{it} + \beta_2 Unqualified_{it} + \beta_3 BigFour_{it} + \beta_4 Segment_{it} + \beta_5 CountWeakness_{it} + \varepsilon_{it}$

where *Custom* is the dependent variable representing a percentage of custom tags. The indices *i* and *t* indicate entity and time, respectively and β is the coefficient estimate for the association of independent variables with the outcome variable. *lnIntan_ta* is the independent variable indicating the log of intangible assets over the log of total assets. *Unqualified* is a dummy variable indicating whether a company received 'unqualified' opinion rating on a financial statement from the auditor, *BigFour* is a dummy indicating whether a company was audited by the Big Four auditing firms and *Segment* is a count variable representing the number of segments in a company. *Weakness* is a dummy signifying the material weaknesses issued by the auditor under Sarbanes-Oxley Act

Section 404 in the financial statement. Finally, *Count Weakness* is a count of all material weaknesses found in each financial statement. Since variables for *Weakness* and *Count Weakness* are interdependent and therefore, highly correlated, two separate regressions are used to estimate the effects of material weaknesses on the use of extensions.

The second linear model regresses the same independent variables on the dollar value of custom tags over total assets – *Custom Value*:

3. $Custom\ Value_{it} = \beta_0 + \beta_1 lnIntan\_ta_{it} + \beta_2 Unqualified_{it} + \beta_3 BigFour_{it} + \beta_4 Segment_{it} + \beta_5 Weakness_{it} + \varepsilon_{it}$

4. $Custom\ Value_{it} = \beta_0 + \beta_1 lnIntan\_ta_{it} + \beta_2 Unqualified_{it} + \beta_3 BigFour_{it} + \beta_4 Segment_{it} + \beta_5 CountWeakness_{it} + \varepsilon_{it}$

Additional testing in this study includes a fixed effects regression model with the company and time-specific fixed effects.

3.1 Dependent variables

The study considers two possible measures of the use of custom tags.

The percentage of custom tags relative to standard tags used by each company (*Custom*). This is calculated by dividing the count of XBRL extensions over the total count of tags for a single firm for a given filing. The value represents firms with a higher percentage use of extensions in their financial reports.

The second outcome variable is the dollar value tagged using custom tags (*Custom Value*). It is obtained by dividing a disclosed dollar amount of a custom tag over a firm's total assets. The value is then averaged per firm-year filing. The variable indicates whether a company is using extensions to mask significant dollar amounts.

3.2 Independent variables

On the right side of the regression model a number of variables were selected to characterize the complexity of a firm and low-quality financial disclosure.

a. Measures of complexity

1. Intangible Assets/Total Assets (*lnIntan_ta).*

To measure complexity, the study uses the natural logarithm of intangible assets (Compustat variable INTAN) over the total assets (Compustat variable TA) of a company. Intangible assets which represent the information uncertainty in the filing are scaled by the total assets which are, per se, a proxy for the company's size, the resulting independent variable is indicative of the level of uncertainty presented by intangible assets. Intangible assets are harder to evaluate and therefore insinuate complexity within the financial statement (Lev, 2000; Skroupa, 2017). Corporations increase their total market value by expanding the use of intangibles since they represent a company's potential revenue and growth. This proposed variable proxies for the complexity of a company; it accounts for intangible assets which represent a value of the business and its long-term prospects in proportion to the company's overall assets. I expect a positive significant correlation of *lnIntan_ta* and the use of custom tags.

2. Number of segments in a company (*Segments*)

Another measurement of complexity in the model is the count of segments (*Segments)* in each firm. This method is consistently used by the accounting literature to proxy for business complexity. Several studies find that a high count of segments for a company is associated with lower financial reporting quality (Doyle, Ge, and McVay 2007; Hoitash, Hoitash and Bedard, 2009). This paper follows previous study examples using segments data to proxy for a company's

complexity. I expect a positive significant correlation between the number of segments and the use of custom tags.

b. Measures of low quality in financial statement disclosures.

There is a possibility of company's intentional manipulation of financial data by increased use of extensions or through intentional obfuscation of reports by the incorrect use of custom tags (Debreceny, Farewell, Piechocki, Felden, Gräning, and d'Eri, 2011).

1. Material weakness *(Count Weakness and Weakness).*

Internal control at the firm-level refers to the assessment of a company's financial reporting reliability and the process of evaluating achievement of that company's strategic and operational goals, as well as assessing of the company's compliance with applicable laws and regulations. The following two variables used to control for the low-quality of financial reports are: a dummy for the existence of a material weakness (*Weakness*) and a count of material weakness for each firm (*Count Weakness)*. A material weakness is defined as a deficiency in internal controls of financial reports. A material weakness in internal control may result in future difficulties with the financial statements. (SEC, 2007). Lashgari, Gawradar and Bakhshayesh (2015) found that companies with high material weakness will present reduced financial information quality. A value of 1 is assigned to variable *Weakness* if a company has an existing material weakness in their filing and a value of 0 if a company has no detected material weakness. While *Weakness* represents the fact of simply having an identified material weakness, observations for *Count Weakness* indicate the number of material weaknesses identified through the assessment of internal controls. Two separate regressions show the association of each internal control variable on the use of custom tags. I

expect a positive significant correlation between both measures of material weakness and the use of custom tags.

2. Auditor's opinion (*Unqualified*).

In the United States, auditors provide assurance that financial statements comply with accounting regulations; and the SEC has enforcement authority with respect to low-quality financial statements. Dechow, Sloan and Sweeney (1996), and Beneish and Messod (1999) find that US companies face a significant stock price penalty if the SEC decides to pursue them for violating accounting standards. Auditors provide investors with independent assurance that the firm's financial statements conform to standard accounting principles. It is important to factor in the auditor's opinion to control for the quality of the financial disclosures.

As a last measure of low-quality financial reports, the auditor's opinion variable *Unqualified* (denoted by a code ranging from 0 to 5) is used to create a dummy specifying financial statements that were fairly presented and approved by the auditing firms. Further descriptive codification of the audit opinion is listed in Table 2. If a filing was 'unqualified' and therefore fairly presented and approved by the auditing firm, a value of 1 was assigned to the observation, value of 0 was assigned to a firm if an opinion was not 'unqualified' (code is greater than 1). These filings were accepted on the principles of consistency and adequacy of financial disclosures. I expect a negative significant correlation between audit opinion *Unqualified* and the use of custom tags.

This study uses the existence of a material weakness (*Weakness)*, count of material weakness (*Count Weakness*) and auditor's opinion (*Unqualified*), to proxy for lower quality financial reporting.

c. *Indeterminate: can measure for either complexity or quality of financial report.*

1. Companies audited by the four largest auditing and accounting firms (*Big Four*).

A variable (Big Four) represents companies audited by the four largest accounting and auditing firms. Prior literature has shown that the Big Four auditors are associated with high standard financial reporting, and therefore have consistently been used to proxy for the audit quality (Lawrence, Minutti-Meza and Zhang, 2011). Big Four will tend to resolve custom tag usage issues by aligning extension elements to the existing taxonomy tags of their client firms. This will result in negative significant relationship between the Big Four and the dependent variable. It is also important to consider that complex firms using extensions to disclose elements not defined by the US GAAP taxonomy will more likely utilize authenticated services of the Big Four, resulting in a positive relationship between the Big Four and use custom tags. Assuming the aforementioned reasoning, the expected sign on the variable *Big Four* is indeterminate. Using Compustat variable 'AU', a dummy variable (Big Four) was created to identify companies that were audited by the four largest accounting and auditing firms: Deloitte Touche Tohmatsu Limited, PricewaterhouseCoopers (PwC), Ernst & Young (E&Y), Klynveld Peat Marwick Goerdeler (KPMG).

4. Data

The 10K XBRL filings submitted to the SEC from Electronic Data Gathering, Analysis, and Retrieval (EDGAR) for the fiscal years from 2015 through 2017 provided the data for the tags. XBRL data for each fiscal year was extracted separately with a total count of 6.5 million observations. The final data for the regression analysis identified 4,329 public companies by their

Central Index Keys (CIK). Table 1 provides a descriptive statistics of initial data collection. The dependent variable is constructed from the EDGAR dataset; it is defined as the percentage of custom tags used by each firm for each filing. Independent variables were drawn from Compustat and Audit Analytics database through Wharton's Research and Data Services (WRDS). From WRDS, Compustat Fundamentals Annual and Segments (Non-Historical) data, the study selected segment observations for all active companies filing within the fiscal years of 2015 through 2017. The count of segments for each company was merged with variables from Fundamentals Annual to create an initial dataset of independent variables. Variables for internal controls were pulled from Auditors and SOX 404 Section of the WRDS Audit Analytics database. The four datasets of dependent and independent variables were then merged on columns for CIK and reporting fiscal year. Rows with missing data on the right-hand side of the model were removed. The resulting dataset yields a total of 11,430 observations for 4,329 firms with variables of company's characteristics. Table 1 further details the changes in the number of observations as the process of data cleaning progressed.

5. Results

5.1 Descriptive statistics

Table 5 presents descriptive statistics for the panel dataset containing 11,430 observations on 4,329 firms. Approximately 18% of all tags were custom in the sample of observations from 2015 to 2017 fiscal years. ± 62% of all observations were audited by the Big Four auditing firms. 75% of all observations were rated 'unqualified' by auditors on their financial reports; this indicates that approximately 25% of observations presented unsatisfactory financial reports as

examined by firm's auditors. 15% of the observations were identified to have some kind of material weakness. Mean value on variable *Custom Value* shows that on average, dollar amounts of custom tags are 47% of total assets. Some disclosed custom tags in the data have dollar values that are significantly greater than total assets (as indicated by Min and Max columns in Table 5). This could in part, be due to the scaling issues of disclosed elements resulting in the incorrect tagging of financial information.

5.2 Correlation matrix

Table 4 presents the correlation matrix. As expected, variables *lnIntan_ta* (p = 0.03) and *Segments* (p = 0.04) proxy for the firm's complexity, are positively correlated with the use of custom tags (*Custom)*. Proxies for internal controls *Weakness* (p = 0.00) and *Count Weakness* (p = 0.00), auditors *Big Four* (p = 0.00), and auditor's opinion *Unqualified* (p = 0.01), also show positive and significant correlations with the use of custom tags. Correlation matrix also shows high significant correlation of the *Custom Value* to independent variables. Variables *Segments* (p = 0.00), *Big Four* (p = 0.00) and auditor's opinion *Unqualified* (p = 0.00) are negatively correlated with *Custom Value*. Material weakness measures *Weakness* (p = 0.00) and *Count Weakness* (p = 0.00) show positive significant correlation to the outcome variable *Custom Value.*

5.3 Estimation Results

The model is estimated using two outcome variables (*Custom* and *Custom Value*) and two estimation strategies: clustered standard errors and fixed effects. The clustered standard error (CSE) approach allows for the mixture of both within-firm and between firm effects and allows for the inclusion of variables which are fixed within firms. The fixed effect approach helps account

for omitted characteristics of firms but does not allow estimation of effects for variables which are constant within a firm.

The regressions for the models are represented in Tables 6, 7 and 8. Table 6 and 7 show Clustered Standard Errors model estimation coefficients on the outcome variable *Custom and Custom Value* respectively, while Table 8 shows the results of the Fixed Effects regressions on both outcome variables *Custom* and *Custom Value*. The following analysis of the regression will discuss each model separately.

5.3.1 Clustered Standard Errors regression results

Columns 1 and 2 of Table 6, present the clustered standard errors method to examine the effects of independent variables on percent use of extensions (*Custom*). The difference in the two columns are due to the different estimate variables used for internal controls. Column 1 estimates the effects of *Weakness* and Column 2 estimates the effects of *Count Weakness*. Sign and significance of the results for the two regressions are similar. The CSE model corrects for the time invariant variables, standard errors, heteroskedasticity and serial correlation of the variables. My hypothesis suggests that the effect on the use of extensions will be different for firms with different characteristics. This model is better for estimating the research hypothesis since it allows for the mixture of both within-firm and between firm effects. Variables used to proxy for a company's complexity, including number of segments (*Segments)* and the ratio of intangible assets to total assets (*lnIntan_ta),* are not significantly correlated with the percent use of extensions. These findings do not support the hypothesis that more complex companies may be involved with more activities that are unique to the company and/or its sectors and thus require the use of custom tags not defined in the standard US GAAP taxonomy.

Big Four, *Unqualified* and material weakness indicators: *Weakness* and *Count Weakness* are all significant in the CSE regressions Coefficients on the *Big Four* (CSE 1: t = 5.38 and CSE 2: t = 4.86) show that companies engaging the services of the Big Four auditing firms are more likely to use extensions having on average 2.1 (CSE 1) and 1.9 (CSE 2) percentage points greater use of extensions per company over the 3 years of the research period. It is likely that the positive correlation between the outcome variable and the Big Four indicates that complex companies (with their complex accounting structure) prefer the attested services of the top auditing firms, thereby increasing the number of custom tags associated with Big Four audits. The positive coefficient on the *Big Four* may be associated with bona fide custom tags provided by the complex firm. Big Four auditors, hired by their respective companies, verify extensions in financial reports: their reputation, high standards and best practices validate financial statements and market valuation of these companies. Future research on the quality of custom tags legitimated by the Big Four could further explain the established results.

Coefficient on *Unqualified* (CSE 1: t = -4.18 and CSE 2: t = -4.55) show that companies validated to have adequate financial statements (ie. had 'unqualified' ratings for their financial reports) used fewer custom tags. The coefficient is statistically significant at the 1% level. These findings are consistent with the expectations since, receiving an 'unqualified' rating in the audit reports indicate conformity with auditing standards in the financial statements. In such cases, the quality of the financial report is reliable. It is important to note however, that the causal effects of the negative coefficient on 'unqualified' auditor's opinion may be reversed: it is possible that firms using fewer custom tags are in conformity with the US GAAP and therefore, more likely to get 'unqualified' ratings.

Companies with material weaknesses (*Weakness,* t = 4.03) were 2.4 percentage points more likely to use custom tags. For each additional count of material weaknesses (*Count Weakness*, t = 3.16) extensions in a filing increased by 0.5 percentage points. The above stated results are statistically significant at the 1% and 5% levels respectively, in the CSE model. The results are consistent with the expectations and are similar to Hoitash and Hoitash (2018), who find that material weakness in a company's filings correlate with the lower quality of financial reports. Higher *Weakness* symbolizing the inadequacy of the financial statement reliability, correlates with higher uses of custom tags, while an auditor's opinion (*Unqualified*) validated and accepted, decrease it.

In Table 7, the dependent variable is the average of the absolute value of the dollar value of custom tags over a firm's total assets (*Custom Value*). The model, estimating the relationship between a firm's characteristics and the value of a disclosed extension shows statistical significance for all variables except for the variable *lnIntan_at*.

The coefficients on *Segments* (CSE 1: t= -3.23 and CSE 2: t=-3.30) are statistically significant at 5% and 1% levels respectively. The results indicate that for each additional segment in a company, the value of a disclosed extension decreases by 0.01 percentage points. This could indicate that firms with more segments might disclose smaller dollar values through more uses of custom tags. This could also indicate that complex firms may not necessarily use custom tags to mask significant dollar values through their extensions.

Coefficients on the *Big Four* (CSE 1: t= -2.74 and CSE 2: t=-3.18) indicate that being audited by the Big Four reduces the firm's likelihood of reporting high dollar amounts through their custom tags by 0.2 percentage points. Results are significant at 5% and 1% levels

respectively. This further implies that Big Four auditing firms legitimate the accurate custom tag values that are not questionable with respect to a firm's operating structure. This means that Big Four will not certify high values of custom tags that could indicate masking of certain corporate operations. However, further research identifying the quality of custom tags approved by the Big Four, can provide further explanations to these findings.

Coefficients on *Unqualified* (CSE 1: t= -10.69 and CSE 2: t=-10.39) also show negative significant correlation between values of disclosed custom tags and the auditor's opinion. Companies with approved and adequate financial statements, report lower dollar values of custom tags. Results are significant at the 1% level and are consistent with the expectation that high-quality financial reports will have lower values disclosed with custom tags.

Variables *Weakness* (CSE 2: t=4.74*)* and *Count Weakness* (CSE 1: t= 4.46) are statistically significant at the 1% level. Coefficients on the measures of material weakness show that material weakness is associated with higher dollar values of custom tags. This suggests that low-quality financial reports proxied by material weakness may use extensions to mask significant dollar amounts in their financial statements.

5.3.2 Fixed Effects regression results

Table 8 displays the results from the linear regression of model 1. Columns 1 - 4 demonstrate the relationship of all independent variables on the outcome variables *Custom* and *Custom Value* using firm and time-varying fixed effects. The fixed effects model estimates *within company effects*, prioritizing the changes of variables over time. The coefficients that are statistically significant in the fixed effects model are the number of *Segments* on the outcome variables *Custom* (FE 1: t = 4.89 and FE 2: t = 4.92) and *Custom Value* (FE 3: t = -2.66 and

FE 4: t = -2.67). An explanation for the insignificance of remaining variables may emerge from the model specification itself. Fixed effects regression specifically estimates changes within a firm. The fixed effects model shows the effect on a firm's use of extensions over time when independent variables also change over time. Since the data sample is only for 3 years, the model that examines the changes within firms and between time periods, may not yield appropriate and expected results. Coefficient*s* on *Segments* for both outcome variables *Custom* and *Custom Value* are significant at the 1% level and 5% level respectively. These results may simply indicate that observed firms had high variations in their segment development. The fixed effects model remains relevant however, for the purpose of this research, since it estimates the effects of changes within a firm's characteristics on the use of custom tags. For the fixed effects model, a dataset with additional years of observation could result in the statistical significance of coefficients. These findings may indicate that the model is not entirely representative for explaining the relationship of complexity and/or company quality of the financial report to uses of custom tags.

6. Conclusion

In 2009 the US SEC made it mandatory for firms to file interactive data using XBRL. The mandate created a new body of literature examining the effects of the new system. Studies since have shown that XBRL improved information efficiency and accessibility expanding the use value as well as the scope of research relevant to financial statements. XBRL data taxonomy uses standard tags to provide easily accessible and machine-readable structured financial data. The SEC, however, also allows for creation of custom tags in cases where a disclosed value cannot be explained by the existing standard tag. The overuse of custom tags results in a reduction of financial report quality, diminishing its comparability and the usability of the filing (Dhole, Lobo,

Mishra, and Pal, 2015). I attempt to further build on XBRL literature by examining use of custom tags and their association with complex and/or low-quality financial statements.

The hypothesis states that there is a higher level of custom tag implementation among firms identifying with complex systems of operation as well as firm's low-quality financial statements. My analysis reveals that a higher percentage of custom tags with respect to the total tags, have a positive correlation with low-quality financial reports as proxied by the 'unqualified' auditor's opinion, and measures for material weakness. I also find that companies identified with lower quality financial disclosures report higher dollar values under custom tags. The results show that companies identified with proxies for complex operational structure including fraction of intangible assets, and the number of segments, are not necessarily associated with inordinate use of extensions.

Research subsequent to this study could implement a string-matching method to identify companies that have lowest differences in descriptive sequences between their use of extensions and the FASB's standard taxonomy. Research beyond the scope of this paper can provide further insights into the results of this study and their potential implications. String matching technique projects the intended uses of extensions by identifying 'legitimate' extensions that were properly placed to describe unique line items, and 'incorrect' extensions that could have been explained by the already existing standard tags. Utilization of string-matching technique can illuminate the relationship between the quality of financial reports and the type of extensions associated with which they are associated. This method can also establish Big Four's association with both 'legitimate' or 'incorrect' types of extensions. This additional metric provides a more accurate representation of low-quality financial statements.

The significance of this study is its contribution to ongoing efforts to maximize transparency in financial statement reports and financial markets, including the refinement of policies that provide for comparable and reliable data for analysis. This research provides a platform for further investigation; its initial findings must be interpreted with caution, since the post-XBRL sample period (3 years) as well as chosen descriptive variables are not strong enough to draw definitive conclusions about the increased uses of XBRL extensions.

7. Tables and Figures

Table 1: Sample Selection

Details	Observations
Initial observations sample of tags from EDGAR from period 2015 - 2017	6,561,207
# of firms in 2015	8,264
# of firms in 2016	7,504
# of firms in 2017	6,923
New transformed observations sample with a sum of tags for each CIK (Previous sample of observations included individual string names of tags in the statement)	46,982
Initial observations sample of financial and segment variables from Compustat	17,249
Initial observations sample of auditor and internal control variables from Audit Analytic	33,614
Merged sample (after removing observations for CIK with no data for all 3 years)	36,163
Observations lost as a result of merging and removing missing values of independent variables	(21,733)
Final # of observed firms from 2015 - 2017	4,329
Final Sample	11,430

The table above is based on a sample of total observations drawn from the EDGAR, Audit Analytics and Compustat North America Industrial Annual data file. The sample covers the period from 2015 to 2017.

Table 2: Audit Opinion Codification for the variable *Unqualified*

Code	Name	Description
0	Unaudited	The financial statements were not audited because they represent consolidated accounts and the auditor's letter refers only to parent accounts.
1	Unqualified	The financial statements are presented fairly, and the auditing firm approves of the accounting principles reflected in the financial statements, the consistency of their application and the adequacy of financial disclosure.
2	Qualified	The financial statements are presented fairly, but the auditing firm is concerned about either limitation on the scope of the examination or unsatisfactory financial statement presentations.
3	No opinion	The auditing firm does not express an opinion regarding the company's capability to continue business operations.
4	Unqualified with additional language	The auditing firm's opinion is unqualified, but explanatory language has been added to the standard report.
5	Adverse opinion	This code indicates that the financial statements are not presented fairly, and the auditing firm does not approve of the accounting principles reflected in the financial statements, the consistency of their applications or the adequacy of financial disclosure.

Compustat description of audit opinion codification.
This study drops all observations for unaudited companies.

Table 3: Variable Definitions.

Description and the definitions of all variables used in the analysis of this paper.
Each variable observation represents a company-year combination.

Name	Definition
Custom	Percentage of tags that are custom out of the total number of tags in each company-year.
Custom Value	Dollar value of custom tags divided by firm's Total Assets and is averaged per firm-year filing.
lnIntan_ta	Natural Log r atio for of Intangible Assets to Total Assets
Segments	Number of segments in a company
Big Four	1 if the auditor is a part of the Big Four and 0 otherwise
Unqualified	An audit opinion: 1 if a company has been given an 'unqualified' rating and 0 otherwise
Count Weakness	Numerical count of material weaknesses in a company.
Weakness	Material weakness identified in assessment of internal controls: 1 if a company has a material weakness and 0 otherwise

Table 4: Correlation Coefficients

Variables	Custom	lnIntan_ta	Custom Value	Segment	Unqualified	Big Four	Count Weakness	Weakness
Custom	1							
lnIntan_ta	0.023* (0.036)	1						
Custom Value	0.182* (0.000)	0.012 (0.270)	1					
Segment	0.019* (0.046)	0.02 (0.074)	-0.129* (0.000)	1				
Unqualified	-0.069* (0.000)	-0.001 (0.919)	-0.267* (0.000)	0.131* (0.000)	1			
Big Four	0.055* (0.000)	0.045* (0.000)	-0.175* (0.000)	0.277* (0.000)	0.263* (0.000)	1		
Count weakness	0.032* (0.001)	0.01 (0.352)	0.173* (0.000)	-0.137* (0.000)	-0.287* (0.000)	-0.286* (0.000)	1	
Weakness	0.044* (0.000)	0.012 (0.286)	0.044* (0.000)	-0.166* (0.000)	-0.345* (0.000)	-0.359* (0.000)	0.790* (0.000)	1

P values are in parenthesis
* Two tailed statistical significance at 5 percent
Variables are defined in Table 3

Table 5: Summary Statistics

This table shows summary statistics for a sample of 4,329 firms for the period 2015–2017. Initial data from EDGAR was combined with additional data from Compustat and Audit Analytics.

Name	Obs	Mean	Std.Dev.	Min	Max
Number of Companies: 4,329					
Custom	11430	18.18	10.74	.488	100
lnIntan_ta	8272	.602	2.982	-134.75	132.41
Custom Value	11,337	.4786	2.547	-29.65	24.56
Segments	11248	9.412	8.832	1	162
Count Weakness	10769	.372	1.081	0	20
Weakness	10769	.159	.365	0	1
Big Four	11430	.625	.484	0	1
Unqualified	11430	.756	.429	0	1

Variables are defined in Table 3

Table 6: Clustered Standard Errors: Custom

1) $Custom_{it} = \beta_0 + \beta_1 lnIntan\_ta_{it} + \beta_2 Unqualified_{it} + \beta_3 BigFour_{it} + \beta_4 Segment_{it} + \beta_5 Weakness_{it} + \varepsilon_{it}$

2) $Custom_{it} = \beta_0 + \beta_1 lnIntan\_ta_{it} + \beta_2 Unqualified_{it} + \beta_3 BigFour_{it} + \beta_4 Segment_{it} + \beta_5 CountWeakness_{it\ i} + \varepsilon_{it}$

	Expected sign	CSE 1: Custom	Expected sign	CSE 2: Custom
lnIntan_ta	+	0.056 (1.02)	+	0.0616 (1.12)
Segments	+	0.0268 (1.4)	+	0.025 (1.31)
Big Four	**?**	2.051*** (5.38)	**?**	1.853*** (4.86)
Unqualified	-	-1.597*** (-4.18)	-	-1.793*** (-4.55)
Weakness	+	2.393*** (4.03)		
Count Weakness			+	0.561** (3.16)
Constant		17.14*** (36.44)		17.59*** (36.16)
Observations		7768		7768
Adjusted R^2		0.016		0.014

T-statistics in parenthesis.

***, **, * Indicate two-tailed statistical significance at the 1 percent, 5 percent, and 10 percent, respectively

Variables are defined in Table 3

Table 7: Clustered Standard Errors: Custom Value

3) $Custom\ Value_{it} = \beta_0 + \beta_1 lnIntan\_ta_{it} + \beta_2 Unqualified_{it} + \beta_3 BigFour_{it} + \beta_4 Segment_{it} + \beta_5 Weakness_{it\ i} + \varepsilon_{it}$

4) $Custom\ Value_{it} = \beta_0 + \beta_1 lnIntan\_ta_{it} + \beta_2 Unqualified_{it} + \beta_3 BigFour_{it} + \beta_4 Segment_{it} + \beta_5 CountWeakness_{it} + \varepsilon_{it}$

	Expected sign	CSE 1: Custom Value	Expected sign	CSE 2: Custom Value
lnIntan_ta	+	0.0177 (-1.54)	+	0.0186 (-1.54)
Segments	+	-0.0116 (-3.23)	+	-0.0118 (-3.30)
Big Four	**?**	-0.192** (-2.74)	**?**	-0.221 (-3.18)
Unqualified	**-**	-0.713*** (-10.39)	**-**	-0.741 (-10.69)
Weakness	+	0.0899*** (4.74)		
Count Weakness			+	0.126*** (4.46)
Constant		0.961*** (10.39)		1.026*** (11.18)
Observations		7643		7750
Adjusted R^2		0.58		0.037

T-statistics in parenthesis.

***, **, * Indicate two-tailed statistical significance at the 1 percent, 5 percent, and 10 percent, respectively

Variables are defined in Table 3

Table 8: Fixed Effects

1a|3a) $Y_{it} = \beta_0 + \beta_1 lnIntan\_ta_{it} + \beta_2 Unqualified_{it} + \beta_3 BigFour_{it} + \beta_4 Segment_{it} + \beta_5 Weakness_{it} + \gamma_i + \delta t + \varepsilon_{it}$

2a|4a) $Y_{it} = \beta_0 + \beta_1 lnIntan\_ta_{it} + \beta_2 Unqualified_{it} + \beta_3 BigFour_{it} + \beta_4 Segment_{it} + \beta_5 CountWeakness_{it} + \gamma_i + \delta t + \varepsilon_{it}$

Y represents the dependent variables *Custom (FE 1 and FE 2)* and *Custom Value (FE 3 and FE 4)* in the Fixed Effects regression model.

	FE 1 Custom	FE 2 Custom	FE 3 Custom Value	FE 4 Custom Value
lnIntan_ta	-0.0107 (-0.99)	-0.0118 (-1.09)	2.103** (-0.68)	2.105** (-0.69)
Segments	0.0764*** (4.89)	0.0769*** (4.92)	-0.00882** (-2.67)	-0.0088** (-2.66)
Big Four	-0.119 (-0.19)	-0.131 (-0.21)	0.195 (1.05)	0.194 (1.106)
Unqualified	-0.0418 (-0.24)	-0.0395 (-0.24)	-0.127* (-1.98)	-0.126* (-2.00)
Weakness	0.389 (1.1)		-0.0137 (-0.77)	
Count Weakness		0.194 (1.68)		-0.0196 (-0.63)
Year Fixed Effects	Yes	Yes	Yes	Yes
Firm Fixed Effects	Yes	Yes	Yes	Yes
Constant	17.12*** (35.02)	17.12*** (35.16)	0.27 (1.9)	0.268 (1.91)
Observations	7768	7768	7750	7750
Adjusted R^2	0.01	0.01	0.002	0.003

T-statistics in parenthesis.

***, **, * Indicate two-tailed statistical significance at the 1 percent, 5 percent, and 10 percent, respectively

Variables are defined in Table 3

8. References

1. Beatty, A., & Liao, S. (2014). Financial accounting in the banking industry: A review of the empirical literature. *Journal of Accounting and Economics*, *58*(2-3), 339-383.

2. Beneish, M. D. (1999). Incentives and penalties related to earnings overstatements that violate GAAP. *The Accounting Review*, *74*(4), 425-457.

3. Bragg, S. (2018). *The Qualitative Characteristics of Financial Statements*. Retrieved from Accounting Tools Accounting CPE Courses & Books: Available at: https://www.accountingtools.com/articles/what-are-the-qualitative-characteristics-of-financial-statem.html

4. Chidambaran, N. K., Kedia, S., & Prabhala, N. (2011). CEO director connections and corporate fraud. *Fordham University Schools of Business Research Paper*, (1787500).

5. Debreceny, R., Farewell, S., Piechocki, M., Felden, C., & Gräning, A. (2010). Does it add up? Early evidence on the data quality of XBRL filings to the SEC. *Journal of Accounting and Public Policy*, *29*(3), 296-306.

6. Debreceny, R. S., Farewell, S. M., Piechocki, M., Felden, C., Gräning, A., & d'Eri, A. (2011). Flex or break? Extensions in XBRL disclosures to the SEC. *Accounting Horizons*, *25*(4), 631-657.

7. Dechow, P. M., Sloan, R. G., & Sweeney, A. P. (1996). Causes and consequences of earnings manipulation: An analysis of firms subject to enforcement actions by the SEC. *Contemporary accounting research*, *13*(1), 1-36.

8. Dhole, S., Lobo, G. J., Mishra, S., & Pal, A. M. (2015). Effects of the SEC's XBRL mandate on financial reporting comparability. *International Journal of Accounting Information Systems*, *19*, 29-44.

9. Elder, R., Zhang, Y., Zhou, J., & Zhou, N. (2009). Internal control weaknesses and client risk management. *Journal of Accounting, Auditing & Finance*, *24*(4), 543-579.

10. Financial Accounting Standards Board (FASB). Modifications to the 2019 US GAAP Taxonomy. [online] Available at: https://www.fasb.org/cs/ContentServer?c=Document_C&cid=1176171813451&d=&pagename=FASB%2FDocument_C%2FDocumentPage

11. Financial Accounting Standards Board (FASB). (2015). *Conceptual Framework—Project Update*. [online] Available at: http://www.fasb.org/jsp/FASB/FASBContent_C/ProjectUpdatePage&cid=900000011090 [Accessed Mar. 2015].

12. Francis, J. R., Pinnuck, M. L., & Watanabe, O. (2013). Auditor style and financial statement comparability. *The Accounting Review*, *89*(2), 605-633.

13. Ge, W., & McVay, S. (2005). The disclosure of material weaknesses in internal control after the Sarbanes-Oxley Act. *Accounting Horizons*, *19*(3), 137-158.

14. Healy, P. M., & Palepu, K. G. (2001). Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of accounting and economics*, *31*(1-3), 405-440.

15. Hoitash, U., Hoitash, R., & Bedard, J. C. (2009). Corporate governance and internal control over financial reporting: A comparison of regulatory regimes. *The accounting review*, *84*(3), 839-867.

16. Humpherys, S. L., Moffitt, K. C., Burns, M. B., Burgoon, J. K., & Felix, W. F. (2011). Identification of fraudulent financial statements using linguistic credibility analysis. *Decision Support Systems*, *50*(3), 585-594.

17. Kile, C. O., & Phillips, M. E. (2009). Using industry classification codes to sample high-technology firms: Analysis and recommendations. *Journal of Accounting, Auditing & Finance*, *24*(1), 35-58.

18. Lashgari, Z., & Gawradar, A. (2015). Internal Control Weakness and Accruals Quality in Companies Listed on Tehran Stock Exchange. *Procedia-Social and Behavioral Sciences*, *207*, 454-461.

19. Lawrence, A., Minutti-Meza, M., & Zhang, P. (2011). Can Big 4 versus non-Big 4 differences in audit-quality proxies be attributed to client characteristics?. *The accounting review*, *86*(1), 259-286.

20. Skroupa, Christopher P. "How Intangible Assets Are Affecting Company Value In The Stock Market." *Forbes*, Forbes Magazine, 2 Nov. 2017. [online] Available at: www.forbes.com/sites/christopherskroupa/2017/11/01/how-intangible-assets-are-affecting-company-value-in-the-stock-market/.

21. Securities and Exchange Commission (SEC). (2009). Interactive Data to Improve Financial Reporting. [online] Available at: http://www.sec.gov/rules/final/2009/33-9002.pdf

22. Securities and Exchange Commission (SEC). (2006). Complexity in financial reporting and disclosure regulation. Speech by SEC Commissioner Cynthia A. Glassman. Remarks before the 25th Annual USC Leventhal School of Accounting SEC and Financial Reporting Institute Conference. [online] Available at: http://www.sec.gov/news/speech/2006/spch060806cag.htm

23. Securities and Exchange Commission (SEC). (2008). Final Report of the Advisory Committee on Improvements to Financial Reporting to the United States Securities and Exchange Commission. [online] Available at: http://www.sec.gov/about/offices/oca/acifr/acifrfinalreport.pdf

24. Securities and Exchange Commission (SEC). 2019. [online] Available at: https://www.sec.gov/Article/whatwedo.html

25. Bronson, S. N., Hogan, C. E., Johnson, M. F., & Ramesh, K. (2011). The unintended consequences of PCAOB auditing Standard Nos. 2 and 3 on the reliability of preliminary earnings releases. *Journal of Accounting and Economics*, *51*(1-2), 95-114.

26. Schneider, A., & Church, B. K. (2008). The effect of auditors' internal control opinions on loan decisions. *Journal of Accounting and Public Policy*, *27*(1), 1-18.

27. Courtis, J. K. (2004). Corporate report obfuscation: artefact or phenomenon?. *The British Accounting Review*, *36*(3), 291-312.

28. Taylor, E. Z., & Dzuranin, A. C. (2010). Interactive financial reporting: an introduction to eXtensible business reporting language (XBRL). *Issues in Accounting Education*, *25*(1), 71-83.

29. Yu Cong, Jia Hao, and Lin Zou (*2014*) The Impact of XBRL Reporting on Market Efficiency. Journal of Information Systems: Fall 2014, Vol. 28, No. 2, pp. 181-207.

30. Li, S., & Nwaeze, E. T. (2015). The association between extensions in XBRL disclosures and financial information environment. *Journal of Information Systems*, *29*(3), 73-99.

31. Kim, J. B., Kim, J. W., & Lim, J. H. (2013). Does XBRL adoption constrain managerial opportunism in financial reporting? Evidence from mandated US filers. In *CAAA Annual Conference*.

32. Arnold, V., Bedard, J. C., Phillips, J. R., & Sutton, S. G. (2012). The impact of tagging qualitative financial information on investor decision making: implications for XBRL. *International Journal of Accounting Information Systems*, *13*(1), 2-20.

33. Lev, B. (2000). *Intangibles: Management, measurement, and reporting*. Brookings institution press.

34. Doyle, J., Ge, W., & McVay, S. (2007). Determinants of weaknesses in internal control over financial reporting. *Journal of Accounting and Economics*, *44*(1-2), 193-223.

35. Liu, C., Luo, X. R., Sia, C. L., O'farrell, G., & Teo, H. H. (2014). The impact of XBRL adoption in PR China. *Decision Support Systems*, *59*, 242-249.