ERASMUS UNIVERSITY ROTTERDAM

ERASMUS SCHOOL OF ECONOMICS

MSc Accounting, Auditing and Control

Master Specialization Accounting & Auditing

The influence of XBRL adoption on analysts' behavior

ERASMUS UNIVERSITEIT ROTTERDAM
ERASMUS SCHOOL OF ECONOMICS

Author: Lotte Janssen
Student number: 447561
Thesis supervisor: drs. R. van der Wal RA
Second reader: dr. J. Pierk
Academic year: 2016-2017

Abstract

In 2008, the SEC made a mandate that firms have to report their 10-K filings in eXtensible Business Reporting (XBRL) format starting in 2009. This implementation occurred in different waves, depending on the public float of a firm. This thesis measures the effect of XBRL adoption on analysts' behavior in the United States market. To measure analysts' behavior, different proxies are used, such as the number of analyst following, forecast accuracy and dispersion, analyst forecast revision time and the information content of filings. In order to examine this effect, I use a difference-in-difference model to incorporate the influence of other economic factors as well. XBRL adoption positively influences the number of analyst following and the forecasts revision time, the forecast accuracy is affected negatively. No consistent evidence is found for the proxies dispersion and information content. Hence, results do not give evidence that the behavior of analysts moves in one direction.

Keywords: XBRL, analysts' behavior, 10-K filings, financial analysts, information quality

Table of Contents

List of tables and figures

1. Introduction

In 2004, the Dutch government started the Dutch Taxonomy Project to cut down the administrative burdens on administrative work. Since the Australian government followed this Dutch approach, the approach is internationally called the Standard Business Reporting Language (SBR) Program (The Dutch Taxonomy Project, 2008). The main goal of this program is to create ''a single set of definitions and language for the information reported by business government'' (Madden, 2009). Madden (2009) argues that another goal of the implementation of SBR is that documents could be sent directly and electronically from the business' system to the users. Sometimes the terms SBR and XBRL (eXtensible Business Reporting Language) are used interchangeable, however these concepts are different. The SBR taxonomy, the grouping of agreed reporting terms, is developed in a technical design. This design is called XBRL (Madden, 2009).

XBRL is an international standard that accommodates financial entities to communicate using a common digital language. This should help the financial entities to communicate their financial information, both internally and externally (Ib et al., 2015). XBRL is an application based on XML, which is the digital standard for the exchange of information between financial entities. The use of this common language has changed the way information is distributed, processed and analyzed. Furthermore, it has influenced the quality and cost of the financial information (Taylor and Dzuranin, 2010).

All information that is supplied by XBRL is standardized under a common language, using a taxonomy structure. The taxonomies define specific tags that ought to be used for certain data inputs. Because of these tags, interrelations between specific sets of data is showed. Furthermore, the taxonomy includes standards on mathematical relations between information and standards on the visual format of information (Bovee et al., 2002).

Many researchers have examined the potential advantages that add value to the financial reporting system. According to Vasarhelyi et al. (2012) XBRL enhances the transparency of financial reporting due to its objective of standardizing the structure and the information of financial statements. This view is supported by the fact that complex procedures of presentation and disclosure of financial statements triggers management to misrepresent firm financial information. Thus, reducing the complexity by standardizing the content of financial statements lowers data manipulation by management and improves transparency of information. In addition to transparency, another important benefit that XBRL provides to financial reporting system is the accuracy of data. XBRL is based on automating reporting processes, which implies minimization of human errors (Ahrendt, 2009). As such, users of financial reporting perceive that data consist of a high level of accuracy if these reports are in XBRL format (Ib et al., 2015).

As with implementation of any new technology, firms have to be careful for possible implications. With the implementation of XBRL, a possible implication is a mistake in the taxonomy. An error in the way tags are created is for example the tag for net profit before tax that by accident is named net profit after tax (Burnett et al., 2006) Burnett et al. (2006) also mention the mistake when the US GAAP taxonomy is used instead of the tag of IFRS. When such a mistake occurs, the whole process of mapping has to be redone. Trites (2006) discusses the possible error of using different taxonomies. Trites (2006) questions why certain taxonomies are used, and if this choice is logical. Bergeron (2004) mentions that new versions and updates appear every now and then, which increases the topic of integration and compatibility. Implementing these new versions and updates bring new costs, and working with it could be very time consuming.

As said before, XBRL structures the preparation of business and financial reports so that it can be communicated easier both internally and externally. For that reason, XBRL improves the capabilities of a CPA in such a way that a CPA is able to publish and direct financial statements more precisely to investors, lenders or other important stakeholders of a firm (AICPA, 2017). The American Institute of CPAs (AICPA) argues that XBRL causes the CPA profession to proactively fit its primary goal by protecting the public interest. They do this by making the investors access to the capital market better. The AICPA mentions the decrease in costs linked with covering a firm is another advantage of XBRL. Because of decreasing costs, a growth in the number of analyst following a firm arises for both small and large companies.

In 2008, the Securities and Exchange Commission (SEC) voted that filing in XBRL is mandatory as of June 15, 2009 for domestic and foreign large accelerated filers that use US GAAP and have a worldwide public common equity float that was at least $5 billion as of the end of the second fiscal quarter of their most recently completed fiscal year. The first wave would impact about 500 companies. One year later, a second wave of XBRL adopters followed, this wave exists of other large accelerated filers but with a public common equity float smaller than $5 billion. The last wave, again one year later, consisted of all firms reporting their financial statements in US GAAP (Efendi et al., 2014). For these filings in XBRL, the SEC does not require that firms get assurance from third parties. Srivastava and Kogan (2010) argue that the reason the SEC is not mandating such an assurance is because of the fact the SEC is afraid that filers are discouraged with incurring costs for assurance. Thereby, the SEC does not hold the filers liable for any mistakes in their filed XBRL documents as long as they look like the standard XBRL documents (Srivastava and Kogan, 2010). However, Boritz and No (2009) argue that assurance can be of importance. They use evidence from the Voluntary XBRL Filing Program (VFP) of the SEC to demonstrate that assurance is of importance. Assurance by a third party would be of best interest for the public who are relying on the XBRL filings (Srivastava and Kogan, 2010). Plumlee and Plumlee (2008) address that unaudited XBRL filings can contain material misstatements because of errors made in the tagging process and that this could be a real concern.

In this thesis, I examine the effect of the implementation of XBRL on the overall behavior of analysts. In this context, overall analysts' behavior means that I, in contrast to other studies, use several measures of analysts' behavior. Proxies used for analysts' behavior are analyst following, analyst forecast revision response time, information content of analyst reports and the accuracy and dispersion of forecasts made by analysts. So, the research question of this thesis is:

What is the effect of XBRL adoption on overall analysts' behavior?

I use the change in analysts' behavior to measure the impact of XBRL adoption because financial analysts in the capital market are a good proxy for informed traders and can be used as sign for information asymmetry (Liu et al., 2014a). Thereby, Liu et al. (2014a) mention that the process capabilities of financial analysts are outstanding. Rock et al. (2001) mention the importance of financial analysts as well, ''Financial analysts play important roles as information intermediaries, monitors of corporate behavior/performance and economic agents''. A lot of research is done to emphasize the importance of analysts' actions and behavior. Mikhail et al. (1999) find proof that forecast accuracy is valuable to analysts. They argue that when analysts suffer more turnover, they are less accurate than their competitors. This indicates that academics and investors rely on analyst forecasts. Whereas Yu (2009) indicates that ''financial analysts are important and influential users of financial reports''. The use of analysts' forecast to analyze market expectations is of such importance because investors' expectations are not observable (Yu, 2010). Therefore, I argue that the behavior of financial analysts can give good insights of XBRL adoption. There is little research done on the actual impact of the implementation of XBRL on the capital market (Liu et al., 2014a). Former research has examined the relation between XBRL-adopters and financial analyst following and forecast accuracy of XBRL-adopters. However, no consistent evidence is found between the relation of XBRL-adopters and overall analysts' behavior. To delve into the impact of XBRL adoption, I look into the change in overall analysts' behavior between adopters of XBRL and non-adopters.

Yoon et al. (2011) argue that the biggest advantage of XBRL adoption is the reduce in information asymmetry. All financial statements users get the same information under XBRL, so no party has different or better information. According to Healy and Palepu (1995), the fact that all users have the same information makes the market more efficient. Shaw (2003) agrees with this by indicating that higher quality information leads to a better functioning equity market. However, there is mixed evidence about the fact that XBRL improves the quality of information. For example, Debreceny et al. (2005) mention that because of flexibility in the tagging problems can arise. Another possibility is that other problems arise by making errors in the tagging process for example. These problems and errors can cause higher information asymmetry. Because of mixed results, more research is necessary. In this thesis, I narrow down this research gap by providing results about the impact of XBRL adoption on analysts' behavior.

As mentioned before, XBRL has as goal to improve the quality of financial information and make information more comparable among firms. Managers and regulators have the intention to improve quality and its usefulness (SEC, 2009). In the case of a negative impact of XBRL adoption on analysts' behavior, regulators and managers have to question if XBRL filings add value to the firm and regulators should think about the fact whether XBRL assurance could help to provide more reliable XBRL filings. For that reason, this research is interesting for CPAs as well, as they can be a third party who could offer assurance on XBRL filings. Thereby, XBRL helps CPAs as well to provide more accurate financial information to users of the financial statement for example for investors, regulators and analysts (AICPA, 2017). XBRL namely structures financial information and reports for both internally and externally decision making. Results from this research are also relevant for information consumers, filers and other stakeholders (Liu et al., 2014a).

This thesis scrutinizes the effect of implementation of XBRL on the behavior of analysts. This thesis uses a difference-in-difference test to examine the effect of XBRL adoption is. The treatment group exists of XBRL adopting firms and the control group exists of non-XBRL adopting firms. Analysts' behavior is measured in five different ways. First, I research whether there is a relation between implementation of XBRL and the number of analyst covering a firm. Following, this thesis examines if the forecasts made by analysts are more accurate and less dispersed. The fourth measure is the forecast revision response time. Finally, to measure the information content of XBRL filings, the effect of XBRL implementation on cumulative abnormal returns (CAR) is measured.

The sample used for the difference-in-difference test consists of two different groups, the treatment and control group. The treatment group consists of firms that adopted XBRL and the control group consists of firms that did not adopt XBRL. The treatment group is based on the SEC EDGAR Dashboard, which in an online platform where firms have to upload their 10-K filings in XBRL format. In the different regressions, I use data from 2000 until 2016. The time frame from 2011 on, is indicated as the post period. This because 2011 is the year that most of the firms were required by then to file their 10-K filings in XBRL format. Observations that do not have complete data of either the dependent variables or the control variables are dropped from the sample.

Section 5 provides results of the different analyses carried out. Throughout the different tests, the number of analyst is increasing and the average analyst forecast revision response time does decrease after XBRL adoption. For the other three variables, no consistent evidence is found. Indicating that the outcomes of XBRL adoption do differ using different regression models.

The remainder of this thesis is structured as following. The next section provides a literature review about XBRL and prior research done as well as the hypothesis development of the proxies of analysts' behavior. Section 3 provides more information about the sample and methodology used in this thesis. In this section, the variables used in the different regressions are explained as well as an elaboration about how the different variables are computed in this thesis. Section 4 provides results of the different regression models. In this section, I also carry out additional analyses to verify whether the results found in section 4.2 are robust. Finally, section 5 provides the conclusion of the literature. Besides that, I discuss the limitations of this thesis and provide suggestions for future research.

2. Literature review

This chapter starts with an introduction of XBRL and how it has developed over time. The second subsection elaborates about prior literature of XBRL and analysts' behavior. Results from previous research of XBRL adoption carried out in China and the United States do not have the same outcomes and for that reason it is still interesting to research the possible relation between XBRL and analysts' behavior. After the prior literature, I will introduce the variables used to measure analysts' behavior. Every variable section finishes with the formulation of a hypothesis.

2.1 XBRL

Peng and Chang (2010) mention that the AICPA argues that the development of XBRL is one of its ''top ten technologies'' for people engaged in the accounting profession. Nowadays, a lot of countries already made XBRL filings mandatory. For example, on December 17, 2008 the SEC determined that all publicly traded companies had to file their filings in XBRL format as of 2009 (Bizarro and Garcia, 2010). Mandating XBRL in the United States is in line with other rules of the SEC in order to replace the Electronic Data Gathering Analysis and Retrieval (EDGAR) system with the Interactive Data Electronic Applications (IDEA) system (SEC, 2009). Whereas the European Union voted for a new European Union (EU) Transparency Directive of higher transparency requirements for financial data a couple years later, namely in 2013. (Liu et al., 2017). The European Securities and Markets Authority (ESMA) requires that issuers in the EU have to report their financial statements in XBRL format as from January 1, 2020 (ESMA, 2017). However, several countries within the EU already make use of XBRL, for example the Netherlands, Belgium and Germany.

As with the adoption of XBRL, it is interesting to investigate whether the benefits mentioned by the SEC are indeed generated after implementation of XBRL. There are several studies that looked into the relation between XBRL adoption and information asymmetry (Blankespoor et al., 2014; Yoon et al., 2011; Tan and Shon, 2009). Groenewegen et al. (2010) define asymmetric information as that one party (principal) has less information than another party (agent) in a transaction. So, one party has superior information over the other party, a so-called informational advantage. Asymmetric information and incentive problems can lead to a capital market that is not functioning efficiently (Akerlof, 1970; Healy and Palepu, 2001; Jensen and Meckling (1976). The agency theory deals with this kind of problems and deals with the question how to reach an efficient outcome (Groenewegen et al., 2010). Adverse selection is a result of asymmetric information. With adverse selection, investors and analysts have less information about the performance of a firm than managers of the firm have and for that reason investors and analysts cannot value investment opportunities properly (Healy and Palepu, 2001). Adverse selection is also possible to occur among investors and analysts. Ravi and Hong (2014) argue that some investors have access to more information than other investors have. Therefore, informed traders have an advantage in investment opportunities in comparison with uninformed traders, indicating that there is adverse selection.

XBRL has as goal to provide all traders with the same kind of information. However, XBRL does not increase the quantity of information but it affects the quality of information (SEC, 2009). Under XBRL, data is standardized so that data can be processed faster and with more accuracy. Automated processing of data has as result that less human errors are made, which increases the quality of information. Yoon et al. (2011) state that implementation of XBRL improves the quality of information and thereby reduces asymmetric information. The SEC (2009) agrees that XBRL adoption also reduces information asymmetry among investors so that all investors have access to the same amount of financial information.

Earlier papers recognize the problem of the 'IT productivity paradox' as a cause for a poor relation between IT investments and performance (Liu et al., 2014b). Rai et al. (1997) argue that an explanation for this phenomenon is that it takes time to fully incorporate new IT investments in order to generate benefits of those investments. It is not only implementing the new IT investment, but the people within the business have also get acquainted with the new technology (Rai et al., (1997). Liu et al. (2014b) mention that the IT productivity paradox is reason that they find a negative effect of using XBRL on forecast accuracy. They used a sample of pre-adoption (2001 – 2003) and post-adoption (2004 – 2006) in China. Since China was the first country to implement XBRL in 2004, there is good reason to believe that indeed the IT productivity paradox caused the negative effect (Kernan, 2008). Blankespoor (2012) indicates that the long-term impact of XBRL implementation is therefore an interesting field for future research because when time elapse, firms, analysts and investors are learning more about the features of XBRL. The IT productivity paradox could therefore also be an explanation why some studies find a negative impact of XBRL on information asymmetry in the year(s) following the adoption of XBRL.

2.2 Prior Literature

As mentioned above, there are mixed results about XBRL adoption. Yoon et al. (2011) tested in the Korean market whether the implementation of XBRL reduces information asymmetry in the stock market. They found that there is a significant negative relation between XBRL implementation and inforamtion asymmetry. They also conclude that the magnitude of decrease in information asymmetry is even bigger for large-sized companies than for medium- and small-sized companies. Blankespoor et al. (2014) examined if information asymmetry increases or decreases around 10-K filings in XBRL in the initial year after the mandate in the United States. Their results indicate that, in contrast with Yoon et al. (2011), information asymmetry does increase after XBRL adoption. They find signfiicant results that after implementation, there is a higher abnormal bid-ask spread, a lower abnormal liquidity and a lower amount of abnormal trading volume. Liu et al., (2014a) scrutinize the relation between implementation of XBRL and the quantity and quality of information in capital markets in the United States. They measured the quantity and quality of the information content by analyst following and forecast accuracy. Their results indicate that after XBRL implementation the quantity and quality of information became better. Tan and Shon (2009) also looked into the number of analyst following after XBRL adoption. However, they used as sample firms that participated in the VFP in the United States. Their results suggest that after XBRL adopation analyst following and trading activity do increase. Liu et al. (2014c) investigate the assocation between XBRL adoption in PR China and uncertainty, such as information errors. They measure uncertainty as the costs of capital and transaction costs of a firm. They find that early adoption in China is associated with higher cost of capital and transaction costs. Liu et al. (2014b) also examine the early adoption of XBRL in China. They scrutinze the relation between XBRL and analysts' forecast accuracy. Their results are that forecast accruacy of analysts did decrease because of XBRL adoption due to a higher degree of uncertainty. The higher degree of uncertainty is caused by for example information errors. Efendi et al. (2014) analyze whether filing in XBRL leads to a higher information content. They investigate if the share price of firms changes on the day that a firm files the 10-K filing in XBRL format. They use data from the VFP and indicate that XRL provides incremental information content. Since the use of XBRL filing is increased, it interesting to scrutinze again what the effect of XBRL filings on the information content of analysts is.

Prior literature shows that the evidence between the United States and Asian countries are different. So, the impact of the XBRL mandate in Asian countries and the United States is not the same. Explanations for this could be differences in institutional and economic conditions (Liu et al., (2014a).

This thesis contributes in several ways. Where other studies did focus on just one or two measures of analysts' behaviour, I incorporate more measures which some om of them has not been tested before. By using several measures, I am looking to both the quantity and the quality of financial information after implementation of XBRL whereas other studies are only looking to only quality or only quanity. Analyst following is for example a proxy of the quantity of information available whereas forecast accuracy and information dispersion are proxies for the quality of information after XBRL (Yu, 2010). Hence, I look into the effect of XBRL adoption on overall analysts' behavior.

Thereby, it interesting to see how financial analysts react on XBRL adoption in the long-term. Several studies did research in PR China or on the VFP but because of the IT productivity paradox, it interesting to see how financial analysts' perceive XBRL adoption over a longer time frame. Since academics and investors do rely on the forecasts of analysts, it is useful to look to the change in analysts' behavior to indicate what the effects are after the adoption of XBRL. The outcomes of this study is useful for analysts, regulators and managers but also for CPAs.

2.3 Analysts' behavior

Bloomfield (2002) argues that when information in financial statements is complex, users of financial statements have more difficulties with interpreting and processing relevant information and that they need more time and effort to find the right information. This thesis focuses on disclosures in XBRL format to investigate whether firm disclosures in XBRL are easier and better to interpret for financial analysts.

2.3.1 Analyst following

I start this thesis with the effect of XBRL on analyst following. Prior research has examined the effect of XBRL on analyst following e.g. Li et al. (2012); Liu et al. (2014a); Tan and Shon (2009). However, there are no consistent findings about the effect of XBRL on the number of analyst following. Because of this and because of the IT productivity paradox effect, it is interesting to test if the number of analyst following is or is not affected by XBRL adoption. Bhushan (1989) investigated which factors are influencing the number of analysts following a firm. He found that several factors are influencing the number of analyst following, either influencing the demand side but also the supply of analyst services. Liu et al. (2014a) measured the quantity and the quality of information under mandatory XBRL by analyst following. They argue that firms with better quality of information attract more analyst following. Since the benefits of the adoption of XBRL, reduced processing costs, increased transparency and more accurate data, it is expected that adoption of XBRL leads to greater analyst following. However, when analysts find it difficult to make use of XBRL because of its complexity it could also have negative effects on analyst following. Li (2008) argues that increasing the complexity in firms' communication could reduce analyst following. So, once XBRL indeed has the capacity to make it easier for analysts to interpret financial information, I expect that the number of analysts increases after XBRL adoption because of decreased costs for the analyst.

Several factors influence the total costs of analysts following a firm. Lehavy et al. (2011) argue that financial analysts are facing costs to process the information provided by firms and that they are facing even higher costs when analysts have more difficulties with interpreting the information. A lower degree of understandability of information leads to higher private search costs for the analysts because they need more time and more information to correctly interpret the information provided by firms' management. XBRL has as goal to standardize financial information so that all users understand the information. Li (2008) states that firms are trying to make financial information more difficult to interpret when there is bad news. Under XBRL this would not be possible anymore, since firms have to use one single format to file their information.

Financial analysts benefit from the fact that not everyone has the same information about a firm. By making their own analysis about the performance of a firm, they can sell their opinions to investors. However, under XBRL everyone has access to the same kind of information and the question arises whether investors are still interested in buying the opinions of financial analysts. Because there are both positive and negative consequences for financial analysts, I test the following null hypothesis:

H_0: XBRL adoption has no association with analyst following.

2.3.2 Accuracy

XBRL has as benefit that accuracy of financial data is improved. As mentioned earlier, Liu et al. (2014a) indicate that financial analysts are important users of financial statements. For that reason, they argue that the forecast accuracy of financial analysts makes a good proxy for the quality of financial information. Hong and Kubik (2003) mention that analysts care about the accuracy of their forecasts as their forecast accuracy is an important factor of the success of an analyst's career. Marshall et al. (2010) argue that implementing XBRL causes that data is obtained quickly and effectively and it therefore increases the accessibility of data. Baldwin and Brand (2011) mention that when analysts have easier and more effectively access to financial data, they are able to include more data in their analysis. Including more data in the analysis leads to a more accurate analysis. Thereby, XBRL is based on automating reporting processes, which implies minimization of human errors (Ahrendt, 2009). As such, users of financial reporting indicate that data is of a higher level of accuracy if these reports are in XBRL format (Ib et al., 2015). For the expected benefits of XBRL mentioned above, I argue that XBRL adoption makes the analyst forecast accuracy of higher quality. Therefore, the hypothesis is:

H_2: XBRL adoption improves analysts' forecast accuracy.

2.3.3 Dispersion

Taylor and Dzuranin (2010) state that XBRL allows its users to report their financial information in one single format. Users can use a wide variety of programs to prepare their financial information. These programs convert financial information into the XBRL format. The file obtained after converting is used as an input to any program supporting XBRL files, without the need to manually convert all the information. After converting the file, specific tags are defined that ought to be used for certain data inputs. Kirk et al. (2016) argue that a problem arises when firms are creating firm-specific tags. Firms are creating new tags when they think that the pre-defined tags in for example the GAAP taxonomy do not match the items in their financial statements properly. The risk of creating firm-specific tags is that an abundance in tags is created Such an overuse in tags does not benefit the comparability between firms, which XBRL has as goal. Baldwin et al. (2006) describe another cause of the increase in firm specific tags. They argue that a lack of XBRL training leads to the possibility that mangers do not know which tag to use and therefore are creating a new tag. Lehavy et al. (2011) find in their study that less readable 10-K filings lead to a greater dispersion of analyst forecasts.

XBRL has as goal to improve the quality of information and thus improving the comparability among firms (SEC, 2009). Thereby, XBRL adoption should lead to the fact that it is easier for analyst to interpret financial information. For those reasons, I argue that the XBRL adoption leads to a less diverse interpretation of financial information. Hence, the third hypothesis is:

H_3: XBRL adoption lowers analysts' forecast dispersion.

2.3.4 Analyst forecast revision time

Lehavy et al. (2011) argue that the energy an analyst puts into following a firm can be measured by the average time from the firm's 10-K filings to the analyst's first report following the filing. Forecast revision time is seen as an important measure for analysts' behavior. Cooper et al. (2001) and Irvine (2003) indicate that both brokerage firms and analysts face advantages of timely forecast of analysts. When analysts make a timely forecast, brokerage firms benefit because there is a greater trading volume and this leads to a rise in commissions. This increase in trading volume and commission benefit analysts as well. However, when analysts have a short forecast revision time, it can be that the accuracy of their forecasts decreases. As they come up with their forecasts faster, they cannot observe the forecasts of other analysts and fail to include important information that became available after the forecast. Gleason and Lee (2003) indicate the importance of the analysts' revision time as well. They state the revision time is an important characteristic about the spreading of information about corporate earnings. Financial analysts revise their forecasts throughout the year and not just after quarterly or annual earnings announcement. Because of the fact that they revise their forecasts often and the timeliness of their forecasts, these revisions provide important information for financial statement users (Gleason and Lee, 2003).

In the last couple of years, a lot of technological developments in accounting has developed, which cause a decrease in time and costs of filing. XBRL is one of these technological developments, its data is available electronically and on the internet which cause that financial information is available in real time (SEC, 2009; Gleason and Lee, 2003). Thereby, XBRL demands their users to file the financial information in just one single format. So, once analysts are acknowledged with XBRL, they should need less time to process the filings because it is easier for them to understand and process the taxonomy used in the filings (Taylor and Dzuranin, 2010). Baldwin and Brand (2011) indicate that when firms do not make use of XBRL, analysts have to work themselves through paper-based financial information, while under XBRL information is available online. Bovee et al. (2002) examine in their paper that timeliness of financial information improves after XBRL adoption. Because of the expected consequences of XBRL as mentioned above I expect that the response of financial analysts after a firm's 10-K decreases. So, my fourth hypothesis is:

H_4: XBRL adoption lowers the average response time of financial analysts.

2.4.5 Information content

The fifth measure is the information content analysts perceive after XBRL filings. The information content indicates how informative the XBRL filings are according to analysts. As already mentioned as advantage of XBRL adoption, Ilias et al. (2015) state that implementation of XBRL decreases the cost of capital because of transparency. When information provided by firms is more transparent for users of financial information, information uncertainty among analysts decreases (Liu et al., 2014c). Pinsker and Li (2008) mention that survey respondents in their study indicate that by making financial information more transparent, the risk and uncertainty of providing capital decreases and thereby cost of capital decreases as well. Also, Efendi et al. (2016) examine whether 10-K filings in XBRL format provide increased information value. They scrutinize if the share price does vary on the day that the 10-K filing is filed. Since XBRL's objective is to improve transparency of information, I expect that after implementation of XBRL the information content of XBRL filings increases. So, the last hypothesis is:

H_4: XBRL adoption increases the information content for analysts.

2.4 Summary

This section started with providing more information about implementing XBRL. The SEC mentions several advantages of XBRL adoption, for example that because of standardized filings, less errors are made in the filings and that the information within the 10-K filings should be easier interpretable for analysts. Other possible advantages of XBRL adoption are increased transparency of filings and decreased asymmetric information among analysts. This because in XBRL format, information is standardized so analysts would have access to the same amount of information. Previous research has scrutinized the possible advantages. However, mixed results are found regarding XBRL adoption and analysts' behavior in Asian countries and the United States. In this thesis, I carry out different regressions in order to answer the hypotheses formulated in this section.

3. Methodology

In order to answer the research question, I perform a difference-in-difference test. Firms that report in XBRL format are the so-called treatment group and companies that did not format in XBRL are the control group. By carrying out a difference-in-difference test, I estimate what the effect of the treatment, adopting XBRL, is and prevent that changes in analysts' behavior are caused by alternative explanations, like changes in firm or stock market characteristics over time.

3.1 Variable description

3.1.1 Analyst following

Lehavy et al. (2011) measure, just as prior research from O'Brien and Bhushan (1990), analyst following as the number of analyst following the firm after the filing date of the 10-K report. O'Brien and Bhushan (1990) mention that the number of analyst following varies over the year, the number of analysts is namely increasing through the year. However, after the earnings announcement date the number of analyst following is steady. Blankespoor et al. (2014) also measure the number of analysts as 'the number of analysts covering the firm, taken most recent consensus analyst forecast measurement date prior to the earnings announcement date'. For that reason, I measure the number of analysts following as the number of analysts following the firm from the most recent consensus analyst forecast measurement date before the 10-K announcement date.

The number of analyst following is available in the I/B/E/S database and can be interpreted as the effort financial analysts put into the analysis of a firm. It is voluntary and not mandatory for firms to provide the number of analyst following. Therefore, it is possible that for some firms there is no data available but still this number in I/B/E/S is a reasonable approximation for analyst coverage. I/B/E/S excludes analysts that tend to be outliers or analyst whose estimates were stopped. The regression model is as following:

$$
\begin{aligned}
&Number\ of\ analyst \\
&\qquad = \beta_0 + \beta_1 XBRL + \beta_2 post + \beta_3 XBRL * post + \beta_4 Size + \beta_5 Leverage \\
&\qquad + \beta_6 EPS + \varepsilon
\end{aligned}
$$

3.1.2 Forecast Accuracy

The forecast accuracy of financial analysts says something about the quality of financial information. The lower the forecast error is, the higher the forecast accuracy of financial analysts is. Analyst forecast accuracy can be measured in different ways. Lehavy et al. (2011) measure forecast accuracy as the ''squared difference between I/B/E/S reported earnings and the analyst consensus forecast, scaled by the share price 90 days before the consensus forecast date". Another way to define the forecast accuracy is by defining forecast accuracy as the forecast error times -1 and normalized. The forecast error is measured by ''deflating the absolute difference between actual price earnings per share (EPS) and consensus forecast EPS by year-start share price'' (e.g. Lang and Lundholm, 1996; Liu and O'Farrell, 2013; Liu et al. 2017). I follow this previous research so that,

$$Forecast\ error = \frac{\left|Actual\ EPS_{i,t} - Consensus\ Forecast\ EPS_{i,t}\right|}{Share\ price_{i,t}}$$

Where actual EPS is the actual, realized earnings per share for firm *j* at time *t*. The consensus forecast EPS is equal to the median analyst forecast of earnings per share and where share price is equal to the price per share in year *t*.

To get the forecast accuracy, I derive from equation 1 the following equation:

$$Forecast\ accuracy = Forecast\ error * (-1)$$

The forecast accuracy is multiplied by minus 1, so that higher values stands for more accurate forecasts (Lang and Lundholm, 1996). In the regression model of analyst forecast, I incorporate the number of analyst covering a firm as control variable. Research have shown that analyst coverage has a moderating effect of the forecast accuracy. The regression model is as following:

$$\begin{aligned} Forecast\ accuracy \\ = \beta_0 + \beta_2 post + \beta_3 XBRL * post + \beta_4 Size + \beta_5 Leverage + \beta_6 EPS \\ + \beta_7 analyst + \varepsilon \end{aligned}$$

3.1.3 Dispersion

The forecasts of financial analysts can be different, dispersed. Dispersion arises because of two reasons. The first reason is that financial analysts use different forecasting models. The second one is that they have different information on which they base their forecasts (Lang and Lundholm, 1996). Analyst forecast dispersion is computed as “the standard deviation of analyst forecasts in the first analyst consensus annual earnings forecasts issued after the 10-K filing for the fiscal period following the 10-K filing divided by share price” (e.g. Lang and Lundholm, 1996; Lehavy et al., 2011). I compute forecast dispersion as the standard deviation of analyst EPS forecasts from the most recent consensus analyst forecast measurement date before the 10-K announcement date deflated by year-start share price.

$$Forecast\ dispersion = \frac{\sigma_{i,t}}{Stock\ price_{i,t}}$$

A negative forecast dispersion indicates that forecast dispersion decreases. A decrease in forecast dispersion entails that forecasts made by analysts are less dispersed. Because of the fact that the SEC mentions several advantages of XBRL implementation, I expect that the forecasts made by analysts are less dispersed after implementation of XBRL. For that reason, I expect a negative forecast dispersion. The regression model is as following:

$$\begin{aligned} &Forecast\ dispersion \\ &\qquad = \beta_0 + \beta_2 post + \beta_3 XBRL * post + \beta_4 Size + \beta_5 Leverage + \beta_6 EPS \\ &\qquad + \beta_7 analyst + \varepsilon \end{aligned}$$

3.1.4 Analyst forecast revision response time

Implementation of XBRL should fasten the response time of analysts after the 10-K filing because of the fact that information is available electronically and online. Analysts’ forecast revision response time is the time analysts need after the 10-K filing report date till the first annual or quarterly earnings forecast of each individual analyst following that firm (Lehavy et al., 2011). The analyst forecast revision response time is measured as the number of the days the analysts need to make a new forecast after the announcement date of the actual value of Earnings per Share (EPS). I make the assumption that the first forecast of the analysts after the announcement of the actual value of EPS is regarding the actual value announced and therefore defined as the analyst forecast revision response time. Since I want to use only data from 'active analysts', I exclude data from analyst who did not report any forecasts 90 days after the announcement date of the actual value of EPS. So, the duration of revision response time of analysts is measured by the length of time in working days between the announcement of the actual EPS value and the first forecast of the analyst following that announcement.

$$
\begin{aligned}
&Analyst\ forecast\ revision\ response\ time \\
&\qquad = \beta_0 + \beta_1 XBRL + \beta_2 post + \beta_3 XBRL * post + \beta_4 Size + \beta_5 Leverage \\
&\qquad + \beta_6 EPS + \beta_7 analyst + \varepsilon
\end{aligned}
$$

3.1.5 Information content

Frankel et al. (2006) examine the informativeness of analysts' reports. They argue that an analyst is better informed when brokerage profits are higher, which means a high trading volume and high volatility, and when analyst show bad news. With a rise in information processing costs, analysts tend to be less informative. A way to measure the informativeness of financial reports is to determine the average share price reactive to the release of analyst forecast revision. A study of Lehavy et al. (2011) is in line with the research conducted by Frankel et al. (2006). However, they measure the information content as ''the proportion of a firm's share returns related to analyst forecast revisions to the total firm's share return during the time period between the 10-K filing and the subsequent fiscal year''. In this thesis, I measure the information content of analysts with the use of the cumulative abnormal returns. The window used is the three-day window, meaning the change in the cumulative abnormal returns one day before the 10-K filing, the day of the 10-K filing and the day after the 10-K filing. Since XBRL has a goal to provide a higher quality of filings, I expect that the CAR for firms that adopted XBRL than for non-adopting XBRL firms. The regression for measuring the information content of XBRL filings is as following be

$$
\begin{aligned}
&Information\ content \\
&\qquad = \beta_0 + \beta_2 post + \beta_3 XBRL * post + \beta_4 Size + \beta_5 Leverage + \beta_6 EPS \\
&\qquad + \beta_7 analyst + \varepsilon
\end{aligned}
$$

3.1.6 Control variables

To make my model more reliable, I add control variables to my regressions. Bhushan (1989) examines which characteristics do affect analyst coverage. He concluded that industry, firm size, ownership structure, return variability, number of lines of business, correlation between firm return and market return have a relation with the analyst coverage of a firm. The first control variable is firm size. Wallace et al. (1994) indicate that the size of a firm and the amount of information available in annual reports are related. Lang and Lundholm (1996) state that there is a relation between the size of a firm and the number of analyst following and forecast accuracy of analysts. Bhushan (1989) agrees with this research by stating that larger firms have a greater analyst coverage. Also, other studies have shown that there is a negative relation between the size of a firm and information asymmetry (Chen et al., 2015; Cong et al., 2014). Gleason and Lee (2003) find in their research that price adjustment process is faster for firms that have a greater analyst coverage, controlling for other factors. The control variable for size is measured as the natural logarithm of total assets.

Leverage is added as next control variable. Leverage indicates the level of financial risk a firm is facing. A higher number of leverage has a negative effect on analysts' behavior. Following control variable is Earnings Per Share (EPS). According to Liu et al. (2014b), EPS is positively related with forecast accuracy. Barniv (2009) states in his paper that he uses EPS is used in his regression model to control for the magnitude for earnings and he agrees with Liu et al. (2014b) that actual earnings have positive relation with forecast accuracy of analysts.

Bhushan (1989) suggests that larger firms have a greater analyst coverage than smaller firms. For that reason, he argues that a greater number of analyst following causes greater private information acquisition. As a consequence, financial analysts can better forecast future earnings over time and earnings announcements tend to be less informative. So, the number of analyst following tends to have a moderating effect on the forecast accuracy of earnings. Except for the regression with the number of analyst following as dependent variable, the number of analyst is added as control variable as well. As control variable, this is measured as the log (1 + number of analyst following) (Blankespoor et al., 2014). An overview of the variables can be found in the appendices.

The predictive validity framework is visualized by the Libby Boxes and can be found in the appendix (Libby, 1981). The Libby Boxes visualize how the constructs of my model are operationalized. The first link of the Libby Boxes present the hypothesized causal relation between XBRL and analysts' behavior. Links two and three reflect the measurements of the dependent and in independent variables. The construct validity is the degree to how well the measurements I chose capture the underlying theoretical construct that it is supposed to measure. Link four is the link which I test in this thesis, the difference-in-difference effect of XBRL on the different measurements of analysts' behavior. Lastly, link five reflects the effect of the control variables used in the model.

The internal validity makes reference to how this thesis captures the causal relation mentioned in link four. Since it is very hard to incorporate all variables that affect analysts' behavior, I chose for the frequently used control variables above. More variables than included do affect the relation between XBRL and analysts' behavior, therefore the endogeneity problem arises. This means that there are omitted correlated variables. In section 4.3 an additional test is carried out to reduce the omitted variable bias. The external validity of model is subject to how well the results of this thesis can be generalized. Since I use only data from the United States, the model is limited in the extent to which this model can be applied to other settings or countries.

3.2 Sample

As mentioned in section 1, the SEC implemented XBRL in different waves in the United States. From these waves, it becomes clear that as of 2011 all public firms that use US GAAP are required to file their statements in XBRL. Prior research has shown that there is no significant relation between changes in the stock market and 10-Q filings (Blankespoor et al., 2014; Kim et al., 2012, Li and Ramesh, 2009). For that reason, I only make use of information relating to 10-K filings. First of all, I drop the observations of firms that reported in XBRL before 2011. As from 2011, all companies that are using US GAAP have to report their filings in XBRL. Because of that reason, and because of the self-adoption bias of voluntary adopters, I start using data from the treatment group as of 2011. Another benefit is that results are less vulnerable to the IT productivity paradox, the industry could become more familiar with the use of XBRL. For already a couple of years, firms were required to file in XBRL and therefore companies became more experienced with filing in XBRL and, possibly, less errors are made in the XBRL filings. So, to control for the fact that errors could affect the behavior of analysts. To carry out a difference-in-difference test, I have to remark a period as post period as well. As I only make use of XBRL filings after 2011, the post period is from 2011 onwards.

Table 1 Sample selection describes the sample selection process in more detail. Since, there are only filings date available for the XBRL filings, I merged those filings with the available fiscal years. Subsequently, I add all the fiscal years for the period (2000-2016). This makes a total of 67,778 available fiscal years for the treatment group (XBRL=1). To obtain data of the control group, I retrieve all fiscal years available on the entire database of North America with corresponding for the period between 2000 and 2016. Available fiscal years for the control group (XBRL=0) is equal to 123,243. For these total fiscal years available, I obtain data for analysts' behavior available on I/B/E/S, information regarding stock return data from CRSP and information regarding fundamentals from Compustat. After merging, I drop duplicates and data for missing data on either the independent variables or the control variables. A lot of data is missing, which causes a final sample of 60,102 observations.

Table 1 Sample selection

XBRL filings available on EDGAR	36,312
Fiscal years available for XBRL Filings	25,597
Fiscal years available treatment group	67,778
Fiscal years available control group	123,243
Total years available	191,021
Less: missing values for all variables	(130,861)
Less: missing values for control	(58)
Total observations left	60,102

3.3 Descriptive statistics

Since the variables used for the different regressions are not normally distributed, outliers are winsorized at a 1% level. Table 2 shows the descriptive statistics of the variables used in the research model. The mean of the number of analyst following a firm is positive with a number of

6.606 which means that on average for the total sample 6.606 analyst are following the firm. In table 3, a distinction is made between the treatment group (XBRL=1) and the control group (XBRL=0). This table also shows that there is a significant difference between the mean of the treatment and control group. On average, more analysts are following a firm that is filing in XBRL than a firm that does not file in XBRL. On average, almost eight analysts are following a firm that implemented XBRL and a bit more than five analysts follow a firm that did not implement XBRL. Since the treatment and control group are significantly different, in section 4 an extra robustness test is carried out without the control sample. The second variable's mean, forecast accuracy, is slightly negative for the total sample. Prior literature from Liu et al. (2014b) already showed that there is a negative relation between forecast accuracy and XBRL adoption. Their conclusion is that because of XBRL adoption the forecast accuracy of analysts decreases because of higher uncertainty. However, comparing the treatment control group, table 4 shows that the forecast accuracy of the treatment group is slightly less negative than the control group. The mean of the control group is -0.0199 whereas the mean of the treatment group is equal to -0.0112. This indicates that XBRL does somewhat help increasing the quality of the forecasts made by analysts. The next measurement of analysts' behavior is dispersion. The mean of dispersion is 0.0043 for the total sample. A positive number indicates that the forecast among analysts is dispersed, indicating that analysts make different forecasts. Table 3 distinguishes between the XBRL adopters and non XBRL adopters, this table displays that dispersion for the treatment group is lower, 0.0037, than for the control group, 0.00499. For both the forecast accuracy and dispersion, I would expect different numbers after the expectations of XBRL adoption the SEC made. The SEC argued that implementation of XBRL would generate benefits for firms, for example improved forecast accuracy and less dispersed forecasts. The descriptive statistics show that the forecast accuracy and the dispersion are not becoming better after implementation of XBRL. However, the forecast accuracy and dispersion of the control group is becoming even worse than the treatment group. So, XBRL seems to have a less negative effect on the forecast accuracy and dispersion of analysts. The fourth dependent variable is the analysts' forecast revision response time. The whole sample has an average response time of almost 18 days. The table 3 shows that the response time for the treatment group is five days faster than the control group. Since I deleted observations from analysts who did not report any revisions after 90 days, the maximum response time does not exceed 90 days. The treatment sample has on average a faster response, this in line with my expectations. The last variable used in the regression model is CAR. The mean of CAR for the whole sample is slightly below 0, -0.0002. This entails that after 10-K filing is filed, abnormal returns decrease slightly. Table 3 shows that the mean of the treatment sample is slightly above 0, and that the control sample has CAR below 0. Thus, the treatment group experiences positive abnormal returns after the filing date but that the control group experiences negative abnormal returns. These results are in line with my expectations.

Table 2 Descriptive statistics whole sample

Variable	N	Mean	Min	Max	Std. Dev	P25	Median	P75
numest	59,658	6,606	1	33	6.403	2	4	9
forecast_accuracy	49,994	-0.016	-1.02	0.00774	0.0590	-0.0079	0.0024	-0,0071
dispersion	43,714	0.0043	0	0.160	0.0115	0.0005	0.0013	0.0033
Response time	53,871	17.837	0	78.5	15.720	5.5	14.135	25.5
CAR	39,632	-0.0002	-0.2932	0.3128	0.0521	-0.2066	-0.0005	0.1973
XBRL	60,102	0.3658	0	1	0.4816	0	0	1
post	60,102	0.4940	0	1	0.5000	0	1	1
XBRL_post	60,102	0.2281	0	1	0.4196	0	0	0
EPS	54,147	1.190	-9.58	12.460	1.979	0.15	0.95	1.98
ln_totalassets	59,574	6.950	1.727	12.954	2.097	5.480	6.8602	8.279
leverage	59331	0.0406	0	0.6318	0.063	0	0.0098	0.0474
control_numest	59,658	1.719	0.693	3.5263	0.781	1.099	1.609	2.3026

Definitions of the variables above can be found in the appendices

Taking the control variables into account as well, table 2 shows that the mean of the EPS is equal to 1.190. For XBRL adopters the mean of EPS is higher than for non XBRL adopters, 1.238 versus 1.143. Interestingly to see is that the size of non XBRL adopters is larger than for the XBRL adopters. This is contrary with what I expect, since the different waves of adopting XBRL were based on the public float of a firm. The first wave of XBRL adoptions consisted of firms with a large public float had to adopt XBRL and only later on other firms. The next control variable is the leverage of a firm. The mean leverage is equal to 0.0448. Table 3 shows that the leverage of XBRL adopters is lower than for not XBRL adopters. Indicating that the total debt relatively to total assets is smaller for XBRL adopters compared with non-XBRL adopters. The last control variable is the number of analyst following a firm, measured as the log of 1 + the number of analysts. This is in accordance with research of Blankespoor et al. (2014). This number is higher for XBRL adopters which makes sense since the number of analyst following for a firm is higher for XBRL adopters as well.

Table 3 Descriptive statistics by group

XBRL=0								
Variable	**N**	**Mean**	**Min**	**Max**	**Std. Dev**	**P25**	**Median**	**P75**
numest*	30,178	5.358	1	33	5.646	1	3	7
forecast_accuracy*	24,930	-0.0199	-1.018	0.00774	0.7111	-0.1021	-0.003	-0.0008
dispersion*	19,961	0.00499	0	0.160	0.0125	0.0006	0.0015	0.0040
response time*	25,894	20.436	0	78.5	17.645	6.5	16.875	30
CAR*	15,972	-0.0010	-0.2932	0.3127	0.0471	-0.0185	-0.0064	0.0017
post*	34,676	0.2721	0	1	0.445	0	0	1
XBRL_post*	34,676	0	0	0	0	0	0	0
EPS*	27,492	1.143	-9.580	12.46	2.040	0.11	0.9	1.9
ln_totalassets*	34,310	7.200	1.781	13.01	2.234	5.474	6.945	8.445
leverage*	33,758	0.0560	0	11.96	0.140	0.0003	0.019	0.065
control_numest*	30,178	1.548	0.693	4.025	0.748	0.6931	1.386	2.079
XBRL=1								
Variable	**N**	**Mean**	**Min**	**Max**	**St. Dev**	**P25**	**Median**	**P75**
numest*	29,480	7.883	1	33	6.863	3	6	11
forecast_accuracy*	25,064	-0.0112	-1.018	0.0077	0.0433	-0.0062	-0.0020	-0.0006
dispersion*	23,753	0.0037	0	0.1600	0.0106	0.0004	0.0011	0.0029
response time*	27,977	15.335	0	78.5	13.554	4.8333	12.294	21.667
CAR*	23,660	0.0003	-0.2932	0.3127	0.0552	-0.0223	-0.0028	0.2176
post*	29,695	0.4616	0	1	0.4985	0	0	1
XBRL_post*	29,695	0.4616	0	1	0.4985	0	0	1
EPS*	26,655	1.238	-9.580	12.460	1.9127	0.18	1	2.06
ln_totalassets*	29,533	6.7000	1.726	12.954	1.9171	5.3847	6.6394	7.9550
leverage*	35,084	0.003	0	0.6318	0.0659	0	0.0494	0.0302
control_numest*	29,480	1.8943	0.693	3.5263	0.7769	1.3863	1.9459	2.4849

The asterisk (*) indicates that the means of the control group (XBRL=0) and treatment group (XBRL=1) are significantly different from each other at a 1% level. Definitions of the variables can be found in the appendices.

4. Empirical analysis

In this section, I first introduce the different OLS regression assumptions. Where necessary, adjustments are made in order to meet the assumptions. After the different assumptions, I introduce the correlation table of the variables. The correlation table shows that most signs are as expected. The second part of this chapter shows the results of the multivariate analysis and whether these results are in line with my expectations. To verify if the results found under 4.2 are robust, additional tests are carried out. This chapter finishes with a summary of the results.

4.1 Univariate analysis

4.1.1 OLS regression assumptions

A couple of assumptions have to be met before an Ordinary Least Square (OLS) regression can be carried out. One of these assumptions is homoscedasticity. Homoscedasticity implies that ''the variance of the error term is constant over various values of the independent variables'' (Pedace, 2013). When the null hypothesis of homoscedasticity is rejected, the model is subject to heteroscedasticity. Heteroscedasticity has as consequence that the assumptions for OLS are not met. So, in case of heteroscedasticity in the regression model, an adjustment has to be made. The Breusch – Pagan test is used to indicate if there is heteroscedasticity one of the five regression used. In the appendices is shown that the null hypothesis for all five regression models is rejected and therefore there is heteroscedasticity in the regression model. To adjust for heteroscedasticity, I make use of robust standard errors. This robustness check corrects for heteroscedasticity by correcting the standard errors of the regression model.

The next assumption of OLS is autocorrelation. Autocorrelation in the regression model indicates that the error time in one period is correlated with the error term in another period. To verify whether the model is exposed to autocorrelation, the Durbin – Watson test is carried out. A number close to 2 indicates that the regression models do not face autocorrelation. The numbers do deviate from 2 for the different tests carried. For this reason, an adjustment is made for these regressions. Autocorrelation is corrected by clustering the unique identifiers of the firms used in the sample.

The next assumption of OLS the normal distribution of the error terms. In order to verify if the errors are normally distributed, a skewness-kurtosis test is carried out. The appendices show that the null hypothesis for normal distribution is rejected, so the error time of the residuals is not normally distributed. As already mentioned in section 3.3, to correct for the fact that the variables are not normally distributed, outliers are winsorized. Ghasemi and Zahediasl (2012) mention that when using large samples, like the sample I use, rejection of the normality assumption does not cause major problems. For that reason, I did not make any additional corrections.

Multicollinearity exists when a linear relation between two more of independent variables in the regression model arises. It is important to check whether the regression is exposed to multicollinearity, because of the fact that multicollinearity can cause biased results. To check if the results are biased because of multicollinearity, a variance inflation factor (VIF) test is carried out for every independent variable. When the VIF value is higher than 10, there is multicollinearity in the regression model (Pedace, 2013). The appendices show the results of the VIF tests carried out for the different regression models. The means of all these tests lie far below 10 and therefore there is no multicollinearity in the model and no adjustment is necessary for multicollinearity.

4.1.2 Correlation matrix

Table 4 shows the correlation matrix of the variables used for the regression models. A correlation matrix describes how the independent variables and the dependent variables are correlated with each other. The variables with an asterisk are significant, whereas the variables without an asterisk are insignificant and therefore do not describe a consistent relation between the variables.

In table 4, the correlation between the number of analyst following a firm and the treatment group, XBRL, is significant and positive. This is corresponding with earlier research from Tan and Shon (2009) which indicates that the implementation of XBRL leads to increased analyst following. However, they only carried out a research of the effect of XBRL implement during the VFP whereas I carry out the effect of XBRL for a longer time frame. So, the positive relation between analyst following a firm and XBRL adoption still exists after the firm adopted XBRL a couple years ago. Other research from Liu et al. (2014a) also indicates that the number of analyst is increasing after XBRL adoption. The correlation between the interaction term (XBRL_post) and the number of analysts is positive and significant as well. This indicates that adopting XBRL causes more analysts following a firm than for firms that did not adopt XBRL, controlling for other economic developments. The control variables for the magnitude of earnings (EPS) and size (natural logarithm of total assets) are positively correlated with the number of analyst following a firm. Table 4 shows that the control variable of number of analyst following and analyst following is highly correlated. However, the control variable for the number of analysts covering a firm is not used in the regression with dependent variable analyst following and does therefore not cause any problems. The correlations mentioned above are in line with prior research and my expectations.

Secondly, the forecast accuracy and the interaction term are positively correlated. This correlation is significant as well. This indicates that XBRL adopters improve their forecast accuracy compared with the control group. This is contradictory with research from Liu et al. (2014b) who found a negative relation between forecast accuracy and XBRL adoption. Liu et al. (2014b) made use of data from early XBRL adoption and it could be that early adopters made errors in the 10-K filings in XBRL and for that reason the accurateness of forecasts decreased. Liu et al. (2014a) find a positive relation between forecast accuracy and XBRL adoption as well. Same as with the number of analysts following, the magnitude of earnings and the size of a firm a positively correlated with forecast accuracy and leverage negatively. For this regression, the number of analysts is added as control variable. This number is significant and positive, indicating that the number of analysts increases the forecast accuracy. The signs of the correlations between the variables and forecast accuracy are as I expected.

The correlation between dispersion and XBRL is negative and significant. However, the interaction term is positively correlated with dispersions. Indicating that forecasts are more dispersed for XBRL adopters comparing to non XBRL adopters. As expected, the control variables for the earnings magnitude, firm size and the number of analysts covering a firm decrease dispersion. Indicating that analysts' forecasts are less dispersed, meaning the forecasts are less different among analysts. The control variable leverage is positive correlated with dispersion. This means that a higher leverage induces a higher dispersion. Hence, the sign on the interaction term of post times XBRL is unexpected. Reason of a negative sign of the interaction term could be that analysts perceive difficulties with processing 10-K filings in XBRL format.

The fourth dependent variable used is forecast revision response time. The correlation matrix shows that the coefficient of the interaction term is negative and significant. This indicates that the number of days an analyst needs to revise his/her forecast decreases after XBRL adoption. The control variable leverage is positive. This effect is expected, since a higher leverage indicates a higher risk and for that reason analysts need more time to revise their forecasts. The control variables magnitude of earnings and firm size are positive correlated with the response time. This is not in line with what I expected. A possible explanation could be that a firm with higher EPS and larger size are more complex. As a result, analysts need more time to revise their forecasts. The number of analysts covering a firm is negative and significant. The more analysts are covering a firm, the less days are needed by analysts to revise their forecasts.

Lastly, table 4 shows that implementation of XBRL is negatively correlated with CAR. The interaction term on CAR is not significant and for that reason, no conclusion can be drawn. However, the control variables are significant. The numbers of EPS, firm size and number of analysts following a firm are positive. This entails that an increase in these numbers, increase the abnormal returns. The control variable leverage is negatively correlated with CAR. I expected all these signs of the variables.

Table 4 shows no high correlations between variables, it is unlikely that the regression models face multicollinearity. In the next section, additional tests are done to indicate whether these assumptions are true.

Table 4 Correlation matrix

	numest	forecast_accuracy	dispersion	Response time	CAR	XBRL	post	XBRL_post	EPS	ln_totalassets	leverage	control_ numest
numest	1.00											
forecast_accuracy	0.16	1.00										
	(0.00)*											
dispersion	-0.15	-0.55	1.00									
	(0.00)*	(0.00)*										
response time	-0.14	-0.03	0.01	1.00								
	(0.00)*	(0.00)*	(0.15)									
CAR	0.01	0.01	-0.01	-0.01	1.00							
	(0.18)	(0.22)	(0.03)**	(0.11)								
XBRL	0.20	0.07	-0.05	-0.17	0.00	1.00						
	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.38)							
post	0.11	0.05	0.06	-0.21	-0.00	0.18	1.00					
	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.41)	(0.00)*						
XBRL_post	0.15	0.04	0.02	-0.22	-0.01	0.52	0.72	1.00				
	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.23)	(0.00)*	(0.00)*					
EPS	0.32	0.26	-0.22	0.01	0.03	0.02	0.10	0.04	1.00			
	(0.00)*	(0.00)*	(0.00)*	(0.02)**	(0.00)*	(0.00)*	(0.00)*	(0.00)*				
ln_totalassets	0.47	0.12	-0.08	0.04	0.03	-0.13	0.12	-0.04	0.50	1.00		
	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.00)*			
leverage	-0.09	-0.14	0.09	0.07	-0.01	-0.14	-0.03	-0.07	-0.00	0.06	1.00	
	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.08)	(0.00)*	(0.00)*	(0.00)*	(0.48)	(0.00)*		
control_numest	0.93	0.21	-0.17	-0.17	0.01	0.22	0.11	0.15	0.31	0.46	-0.11	1.00
	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.06)	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.00)*	

*p<0.05 robust p-value in parentheses.
Definitions of the variables can be found in the appendix

4.2 Results

Table 5 provides the results of the different regressions carried out. As the table shows, almost all coefficients of the variables used are significant at the 1% level. Thereby, the F-statistics of the different regression all have a p-value of 0.000 indicating that all the models are significant at the 1% level. However, the R-squared of the models is different. As table 5 shows, the R-squared of the model of number of analyst following is much higher than model two till five. The higher R-squared of model I indicates that this model has a higher explanatory value than the other models. Especially model 5 has a very low R-squared and therefore a low explanatory value.

Table 5 Regressions effect of XBRL on analysts' behavior

The effect of XBRL on analysts' behavior

Variables	**I** **Number of analyst following**	**II** **Forecast Accuracy**	**III** **Dispersion**	**IV** **Response time**	**V** **CAR**
Intercept	-4.593***	-0.0313***	0.00624***	20.828***	-0.00549***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
post	-1.019***	0.00314***	0.00208***	-4.343***	-0.000275
	(0.000)	(0.000)	(0.000)	(0.000)	(0.735)
XBRL	2.307***	0.00256***	-0.000470***	-1.991***	0.00239***
	(0.000)	(0.000)	(0.002)	(0.000)	(0.001)
XBRL_post	1.838***	-0.00506***	-0.000623**	-1.224***	-0.00137
	(0.000)	(0.000)	(0.030)	(0.000)	(0.211)
EPS	0.270***	0.00717***	-0.00120***	0.122***	0.000472**
	(0.000)	(0.0000)	(0.000)	(0.001)	(0.015)
ln_totalassets	1.460***	-0.00166***	0.000468***	1.081***	0.000802***
	(0.0000)	(0.000)	(0.000)	(0.000)	(0.001)
leverage	-8.626***	-0.0925***	0.00900***	5.706***	-0.00380
	(0.000)	(0.0000)	(0.000)	(0.000)	(0.466)
control_numest	.	0.0117***	-0.00230***	-4.212***	-0.000961**
	.	(0.000)	(0.000)	(0.000)	(0.033)
F-statistics	0.000	0.000	0.000	0.000	0.000
Observations	53,453	49,344	40,308	49,062	37,831
R-squared	0.300	0.105	0.080	0.080	0.001

*** p<0.01, ** p<0.05, * p<0.1, robust p-value in parentheses.

The table shows coefficient estimates and p-value (in parentheses) from the regressions of the effect of XBRL on analyst following, forecast accuracy and dispersion with the corresponding control variables. Analyst following is determined as the number of analyst following a firm. Observations from analysts which tend to be outliers and analyst whose estimates were stopped are excluded. Second regression is the effect of XBRL on forecast accuracy. One extra control variable is added to this regression compared with the regression of the number of analyst following. This extra variable is the control variable for the number of analyst following a firm. The third regression is the effect of XBRL on dispersion. In this regression, the control variable number of analyst following is added as well. The fourth regression is the analyst revision response time. To include only active analysts, observations of analysts who did not make any forecasts within 90 days after the announcement of the actual were removed. The last regression is the effect of XBRL implementation on the information content of the filings, proxy for this variable is CAR. Definitions of all variables can be found in the appendices.

4.2.1 Analyst following

The first regression is the effect of implementation of XBRL on the number of analyst following. All coefficients in this model are as expected. The coefficient on XBRL is positive and significant which indicates that implementing XBRL possibly has a positive effect on the number of analyst following, which means that implementation of XBRL leads to a higher number of analyst following a firm. The question is whether the increased number of analyst following caused by the implementation of XBRL and not caused by other alternative explanations. For that reason, a difference-in-difference test has been carried out. The coefficient of interest in a difference-in-difference test is β_3, the interaction term XBRL*post. This coefficient estimates the effect of treatment group (XBRL=1) compared with the control group (XBRL=0). Table 6 gives a clearer view of the results of the difference-in-difference test. This table implies that the difference between the treatment group and the control group is equal to 1.838 and the number is significant. The coefficient of XBRL (β_2) is significant and is equal to 2.307, so the increase of the number of analysts following was not completely due to the implementation of XBRL. Other explanations could be economic factors. Since XBRL was adopted after the financial crisis, it could be that the recovering of the financial markets increased the numbers of analysts as well. However, it can still be concluded that the implementation of XBRL leads to higher number of analysts following the firm compared with firms that do not file in XBRL format.

Table 6 Difference-in-Difference test numest

Difference-in-difference test Number of Analyst Following

	Pre-2011 (post = 0)	Post-2011 (post = 1)	Difference
XBRL = 1 (Treatment group)	-5.612	-1.467	4.145
XBRL = 0 Control group	-4.593	-2.287	2.307
Difference	-1.019	0.82	1.838

The numbers can slightly differ from the numbers of the regression model because of rounding.

This result is in line with earlier research from Liu et al. (2014) and Tan and Shon (2009). These results are also in line with the coefficients in the correlation matrix. However, the effects of XBRL in a regression are stronger than in the correlation matrix. Interstingly to see is that the post coefficent (β_1) is negative. Indicating that after 2011, the number of analyst following is decreasing. As expected are the control variables for the magnitude of earnings (EPS) and firm size (ln_totalassets) positive correlated with the number of analyst following. This indicates that a higher magnitude of earnings and a larger firms size causes that more analysts are following a firm. The last control variable is leverage, this control variable is negatively correlated with number of analyst following a firm. Indicating that a higher leverage decreases the number of analyst following a firm. The coefficient of the control variables is following my expectations. All the control variables are significant at the 1% level.

4.2.2 Forecast Accuracy

The second regression carried out is the effect of XBRL on the forecast accuracy of analysts. I expect that implementation of XBRL has a positive effect of the forecast accuracy, indicating that the coefficient on the interaction term would be positive. Table 5 shows that the coefficient on XBRL is positive and significant, which could indicate that implementation of XBRL increases the forecast accuracy. To be sure that this positive coefficient is not due other economical explanations, a difference-in-difference test is carried out. Again, the coefficient of interest is (β_3). Surprisingly, the coefficient on the interaction term is negative. The higher the number on coefficient (β_3), the higher the accurateness of the forecasts would be. A negative number indicates that the forecasts of analysts are becoming worse after implementation of XBRL. So, in this case the accuracy of the forecasts is decreasing after implementation of XBRL.

Table 7 Difference-in-Difference test forecast_accuracy

Difference-in-difference test Forecast Accuracy

	Pre-2011 (post = 0)	Post-2011 (post = 1)	Difference
XBRL = 1 (Treatment group)	-0.028	-0.031	-0.003
XBRL = 0 Control group	-0.031	-0.029	0.002
Difference	0.003	-0.002	-0.005

The numbers can slightly differ from the numbers of the regression model because of rounding.

Previous research provides mixed results as well, where Liu et al. (2014a) found results that implementation of XBRL increaes the forecast accuracy of analysts, but Liu et al. (2014b) found results that forecast accuracy is deceasing after implementation of XBRL in PR China. Table 5 includes the effects of the control variables on the forecast accuracy. The coefficients of the control variables for the magnitude of earnings (EPS) and the number of analysts (control_numest) are positive and significant. This indicates that the a greater magnitude of earnings and a higher number of analyst covering a firm do increase the the accurateness of forecasts. I expected that the size of a firm is positively correlated with the forecast of accuracy, however, the coefficient on the size (ln_totalassets) is negatively and significant. This indicates that a larger firm size reduces the accuracy of forecasts. It could be that a larger firm size is more complex and therefore it is harder for analysts to make accurate forecasts. Reason for decreasing accurateness could be that analysts do not interpret the 10-K filings in format correctly or that errors are made in the XBRL filings, and for that reason make wrong forecast are made. As a consequence, the forecast accuracy decreases

4.2.3 Dispersion

The third regression is the effect of XBRL on the dispersion of forecasts made by analysts. Dispersion indicates how well the analysts understand the information given in the financial statements. A higher coefficient on β_3 indicates that the forecasts of analysts are more dispersed, meaning that analysts make diversed forecasts. So, a negative coefficient is preferred, indicating that the dispersion is reduced and better forecasts can be made. Table 5 shows that the coefficient of XBRL is as expected. The coefficient is negative, -0.00012, indicating that XBRL adoption could be a reason of decreased dispersion. Once again, I carry out a difference-in-difference test to estimate wheter these results are completely due to the implementation of XBRL or that these results are due other economic explantions. In table 8, a more clear view is presented of this test. The table shows the effect between the treatment and control group is -0.00113. This number is significant and negatively correlated with dispersion. So, XBRL implementation does decrease the dispersion of forecasts, however the real effect is somewhat lower than table 5 shows from the coefficient on XBRL. In this model the same control variables are used as in the model for the forecast accuracy. The coefficient of these variables are consistent with the findings for forecast accuracy. The control variables for the magntiude of earnings (EPS) and the number of analyst covering a firm (control_numest) decrease dispersion whereas the control variables for firm size (ln_totalassets) and levearge (levearge) are postively correlated with dispersion. This entails that an increase of these control variables cause more dispersed forecasts. All these signs are as expected.

Table 8 Difference-in-Difference test dispersion

Difference-in-difference test Dispersion

	Pre-2011 (post = 0)	Post-2011 (post = 1)	Difference
XBRL = 1 (Treatment group)	0.008	0.007	-0.001
XBRL = 0 Control group	0.006	0.006	-0.000
Difference	-0.002	-0.001	-0.001

The numbers can slightly differ from the numbers of the regression model because of rounding.

4.2.4 Analyst forecast revision time

The fourth dependent variable is the analyst forecast revision time. The reaction time of analysts is measured by the number of days an analyst need to make the revised forecast. After XBRL implementation, it is expected that the number of days analysts need decreases. For that reason, I expect that coefficient on the interaction term XBRL times post is negative. Table 4 shows that the coefficient on the interaction term is indeed negative. After implementation of XBRL, analysts need less time to make revised forecasts. The difference-in-difference effect is better displayed in table 9. The table shows that for both the treatment and control group, the revision response time decreases after 2011. The analysts following a firm of the control group needed it before 2011 on average 20.828 to revise their forecasts where that was 18.837 after 2011. Analysts following a firm in the treatment group needed less time than the control group before 2011, on average 16.497 days. After 2011, they only needed 13.263 to revise their forecasts. So, XBRL adopters need less time to revise their forecasts. On average, the XBRL adopters revised their forecast revision 1.243 days faster than non-XBRL adopter. The control variable for the number of analysts following is as expected. The coefficient is negative, indicating that a higher number of analyst covering a firm leads to a faster forecast revision response time. The coefficient of leverage is also as expected. The coefficient on leverage is positive, meaning that a higher leverage ratio leads to a delayed forecast revision response of analysts. The number of the coefficients on the magnitude of earnings (EPS) and firm size (ln_totalassets) are not as expected. I presumed that a greater magnitude of earnings would lead to a shorter forecasts revision response time. I also presumed that a larger firm size would lead to a shorter forecasts revision response time, but also the coefficient on firm size is positive. An explanation could be that larger firms file more complex information in their 10-K filings which is harder for the analysts to interpret and for that reason analysts need more time to make a revised forecast.

Table 9 Difference-in-Difference test reaction

Difference-in-difference test Reaction			
	Pre-2011 (post = 0)	Post-2011 (post = 1)	Difference
XBRL = 1 (Treatment group)	16.497	13.263	-3.234
XBRL = 0 Control group	20.828	18.837	-1.991
Difference	4.331	3.088	-1.243

The numbers can slightly differ from the numbers of the regression model because of rounding.

4.2.5 Information content

The last dependent variable used to measure the effect of XBRL adoption on analysts' behavior is CAR. CAR is used to measure the information content of the 10-K filings. A higher coefficient on the interaction term indicates that more information is released after the 10-K filing. As XBRL has as goal to improve the quality of information, I expect that the coefficient on the interaction term is positive. Table 5 displays the opposite of this. It presents that the coefficient for the XBRL adopters is positive, indicating that for the firms that implemented XBRL the abnormal returns do increase after the 10-K filing. However, as table 5 and 10 present, the interaction term of post*XBRL is negative. Important to note is that this number is not significant and for that reason no conclusion can be drawn from this coefficient. The coefficients on the magnitude of earnings (EPS) and firm size (ln_totalassets) are as presumed. The numbers of both coefficients are both positive and significant. This means that higher earnings and a larger firm size would lead to greater abnormal returns. The number on the control variable of the number of analysts covering a firm is negative and significant. This could be logical, because when more analysts are following a firm, more and more information is becoming public. The coefficient on leverage is negative but insignificant. Hence, no conclusion can be drawn from the effect of XBRL adoption on the information content.

Table 10 Difference-in-Difference test CAR

Difference-in-difference test Information Content			
	Pre-2011 (post = 0)	Post-2011 (post = 1)	Difference
XBRL = 1 (Treatment group)	-0.005	-0.004	-0.001
XBRL = 0 Control group	-0.005	-0.003	-0.002
Difference	-0.000	-0.001	-0.001

The numbers can slightly differ from the numbers of the regression model because of rounding.

4.3 Robustness check

In this section, I carry out some additional analysis to verify if the results found under 4.2 are robust. A robustness check gives extra information how certain outcomes are and if the results are behaving differently under other circumstances. This section discusses three additional analyses, which includes mean centering, the fixed effects model and a regression based on only the treatment group.

4.3.1 Mean centering

The first additional analysis carried out is mean centering, indicating that I estimate the center of each variable and subtract it from each observation. Subsequently, a robust regression is carried out by using the centered variables in order to verify whether the coefficients in 4.2 are robust. According to Bramati and Croux (2007) mean centering has a consequence that mean centering makes a reduction in the number of parameters because it cancels out the fixed effects parameters. Kreft and de Leeuw (1998) mention another advantage op mean centering. They indicate that mean centering leads to a lower correlation between intercepts and slopes of coefficients. The results of mean centering can be found in table 9. This table shows that, after mean centering, all coefficients of the variables are still the same as in the results provided in section 4.2. The only coefficient that does change, is the coefficient of the intercept. Results of the robustness check can be found in table 11.

Table 11 Robustness check

Robustness check effect of XBRL on analysts' behavior

Variables	I Centered Number of analyst following	II Centered Forecast accuracy	III Centered Dispersion	IV Centered Reaction	V Centered CAR
Constant	-1.082***	-0.00242***	0.000163*	3.645***	-0.000939*
	(0.0000)	(0.0000)	(0.0548)	(0.0000)	(0.0582)
post	-1.019***	0.00314***	0.00208***	-4.331***	-0.000275
	(0.0000)	(0.0003)	(0.0000)	(0.0000)	(0.7808)
XBRL	2.307***	0.00256***	-0.000470***	-1.991***	0.00239***
	(0.0000)	(0.0001)	(0.0001)	(0.0000)	(0.0006)
XBRL_post	1.838***	-0.00506***	-0.000623***	-1.243***	-0.00137
	(0.0000)	(0.0000)	(0.0037)	(0.0000)	(0.2538)
centered_EPS	0.270***	0.00717***	-0.00120***	0.122***	0.000472***
	(0.0000)	(0.0000)	(0.0000)	(0.0023)	(0.0030)
centered_leverage	-8.626***	-0.0925***	0.00900***	5.706***	-0.00380
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.3610)
centered_ln_totalassets	1.460***	-0.00166***	0.000468***	1.081***	0.000802***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
centered_control_numest	.	0.0117***	-0.00230***	-4.212***	-0.000961**
	.	(0.0000)	(0.0000)	(0.0000)	(0.0316)
F-statistics	0.000	0.000	0.000	0.000	0.000
Observations	53,453	49,344	40,308	49,070	37,831
R-squared	0.300	0.105	0.080	0.080	0.001

*** p<0.01, ** p<0.05, * p<0.1, robust p-value in parentheses.

Since the VIF score of the regressions were already low, it makes sense that mean centering does not change the coefficients.

4.3.2 Fixed effects model

The second additional test is the fixed effects model. Since it is very hard to include all variables in the regression model that affect the effect of XBRL adoption on analysts' behavior, there is an omitted variable bias. This means not all variables are included in the regression, and for that reason the results found in section 4.2 could be biased. Greene (2011) mentions that the fixed effects model can reduce this omitted variable bias by controlling for example for industry and firm factors. Controlling by using the fixed effects model can help to prevent for a spurious association between variables and time-constant unobserved characteristics. The fixed effects model makes it possible that the variables and the unobserved characteristics are correlated with each other in the regression model (Greene, 2011). I include the fixed effects for industry and fiscal year in the fixed effects model.

The results of the fixed effects model can be found in table 12. As the table displays, the numbers on the regression of the effect of XBRL implementation on the number of analyst following slightly change. The interaction term in the regression model in section 4.2 was equal to 1.838 whereas the coefficient of the interaction term in the fixed effects model is equal to 1.354. So, after controlling for the fixed effects of industry and years, the number of analyst is less than without controlling for the fixed effects. The same happens with the coefficient of the interaction term in the regression of forecast accuracy. Indicating that under the fixed effects model, forecasts made by analysts are less worse. However, still after implementation of XBRL, the forecast of accuracy decreases. Table 12 shows that coefficients of the interaction term of dispersion and analyst revision response time increase. This means that forecasts of analysts were less dispersed and that analysts on average need less time to revise their forecasts. The interaction term in the regression model of CAR is still not significant and for that reason I can still not derive a conclusion on the effect of XBRL implementation on the information content, measured by CAR.

The fixed effects model has compared with the model in section 4.2 higher R-squared for the different regressions. The higher the number of the R-squared the more closely the data of the variable are fitted to the regression line. Indicating that a higher number of R-squared means that the model is better explaining the variability of the data. It is logical that in the fixed effects model the R-squared is higher, because of the fact that a correction is made for possible omitted variables.

Table 12 Fixed Effects model

The effect of XBRL adoption on analysts' behavior using the Fixed Effects model

Variables	I Number of analyst following	II Forecast Accuracy	III Dispersion	IV Reaction	V CAR
Constant	-6.122***	-0.0341***	0.00687***	29.71***	0.0305***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
post	-1.843***	0.000849	0.00443**	-13.86***	-0.0431***
	(0.0048)	(0.7819)	(0.0242)	(0.0000)	(0.0000)
XBRL	1.702***	0.00600***	-0.00140***	-0.760*	0.00276***
	(0.0000)	(0.0000)	(0.0000)	(0.0509)	(0.0091)
XBRL_post	1.354***	-0.00407***	-0.000792	-2.262***	-0.000997
	(0.0001)	(0.0031)	(0.1021)	(0.0000)	(0.2185)
EPS	0.272***	0.00712***	-0.00116***	0.152	0.000477*
	(0.0001)	(0.0000)	(0.0000)	(0.1020)	(0.0897)
ln_totalassets	1.790***	-0.00164**	0.000376	0.999***	0.000938***
	(0.0000)	(0.0147)	(0.1078)	(0.0000)	(0.0019)
leverage	-5.409***	-0.104***	0.0125***	4.898***	-0.00396
	(0.0004)	(0.0000)	(0.0001)	(0.0042)	(0.5063)
control_numest	.	0.0128***	-0.00248***	-4.439***	-0.00107*
	.	(0.0000)	(0.0000)	(0.0000)	(0.0664)
Fyear					
2001	-0.0834	-0.0108**	0.00141**	-5.723***	-0.0434***
	(0.5303)	(0.0399)	(0.0416)	(0.0000)	(0.0000)
2002	-0.326**	0.00560***	-0.000872**	-2.415***	-0.0400***
	(0.0257)	(0.0092)	(0.0246)	(0.0001)	(0.0000)
2003	-0.164	0.000706	0.000508	-7.346***	-0.0353***
	(0.3334)	(0.7600)	(0.2770)	(0.0000)	(0.0000)
2004	-0.316	0.00793***	-0.000901***	-7.358***	-0.0353***
	(0.1538)	(0.0000)	(0.0084)	(0.0000)	(0.0000)
2005	-0.658***	0.00738***	-0.000547*	-9.277***	-0.0387***
	(0.0040)	(0.0000)	(0.0864)	(0.0000)	(0.0000)
2006	-0.700***	0.00561***	-0.000162	-10.96***	-0.0361***
	(0.0018)	(0.0010)	(0.6203)	(0.0000)	(0.0000)
2007	-1.141***	0.00430**	0.000163	-11.21***	-0.0363***
	(0.0000)	(0.0118)	(0.6535)	(0.0000)	(0.0000)
2008	-1.401***	-0.00780**	0.00117***	-12.39***	-0.0380***
	(0.0000)	(0.0150)	(0.0042)	(0.0000)	(0.0000)
2009	-1.096***	-0.0145***	0.00450***	-11.70***	-0.0359***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
2010	-0.572**	-0.000435	0.00200***	-11.97***	-0.0375***
	(0.0127)	(0.8628)	(0.0000)	(0.0000)	(0.0000)
2011	0.572	0.00225	-0.00157	3.687***	0.00549
	(0.2817)	(0.4376)	(0.3965)	(0.0028)	(0.2049)
2012	0.520	-0.00123	-3.20e-05	2.233**	0.00670*
	(0.3251)	(0.6808)	(0.9864)	(0.0498)	(0.0858)
2013	0.639	-0.000715	-0.00164	2.021**	0.00525
	(0.1911)	(0.7969)	(0.3597)	(0.0427)	(0.1765)
2014	0.449	0.00258	-0.00244	-0.0607	0.00570
	(0.3459)	(0.3153)	(0.1810)	(0.9538)	(0.1674)
2015	0.454	-0.000441	-0.00117	0.139	0.00470
	(0.3613)	(0.8523)	(0.5199)	(0.8997)	(0.2373)
F-statistics	0.000	0.000	0.000	0.000	0.000
Observations	53,453	49,344	40,308	49,062	37,831
R-squared	0.348	0.123	0.103	0.108	0.002
Number of sic2	70	69	68	70	69

Robust pval in parentheses. *** p<0.01, ** p<0.05, * p<0.1

4.3.3 Treatment sample only

The last additional test carried out is a regression model without the control sample. In section 3.3, I carried out a t-test of the variables between the treatment and control group. As concluded in that section, the two groups are significantly different. This entails that the means of the variables are significantly different comparing the treatment group with the control group. For that reason, the control group used in section 4.2 is not sufficient enough to carry out a difference-in-difference test. For that reason, in this section I carry out a regression using only the treatment group. The following general regression model is used:

$$Analysts'\ behavior = \beta_0 + \beta_1 XBRL + \beta_2 control\ variables + \varepsilon$$

In this regression, XBRL is a dummy variable, it is equal to 1 when a firm implemented XBRL and 0 otherwise. To avoid the self-selection bias, observations from voluntary adopters are dropped. Voluntary adopters are adopters which float was smaller than $5 billion and $700 million in respectively fiscal year 2009 and 2010. The results of the effect of XBRL on analysts' behavior without the control sample can be found in table 13. Some of the results showed in table 13 are in line with the results found in section 4.2. The number of analysts covering a firm is increasing after XBRL adoption in this model as well. In the regression model of XBRL implementation on forecast accuracy can be seen that that the coefficient on XBRL (β_1) is negative. This means that after implementation of XBRL, the forecasts made by analysts are less accurate. Also, these findings are in line with the findings found in section 4.2. However, the coefficient found on the regression model of dispersion is different. Table 13 shows a positive coefficient on XBRL. A higher number on dispersion means that the forecasts made by analysts are more dispersed. This indicates that the difference in forecasts made by analysts increases after XBRL implementation. This finding is not in line with the findings found in section 4.2.3. In section 4, table 5 showed that dispersion decreased after implementation of XBRL. Because of the fact that the control group is not sufficient, the findings in section 4.2 could be biased and for that reason it is not possible to draw a conclusion on the effect of XBRL adoption on dispersion. The effect of XBRL adoption on the forecast revision response time is line with results in 4.2.4. Also in the table below, it is shown that the number of days analysts need to revise their forecasts decrease. The last regression is the effect of XBRL on the information content. Whereas the coefficient of the interaction term post and XBRL in table 5 was not significant, in table 13 the coefficient on XBRL is significant and negative. Indicating that after implementation of XBRL the information content of the 10-K filings is decreasing. Comparing the R-squared of the model without control sample and the difference-in-difference test performed in section 4.2, the R-squared of the model in table 13 is higher. So, the model used in table 13 explains better the changes in variables.

Table 13 Additional analysis treatment sample only

The effect of XBRL adoption on analysts' behavior treatment sample only

Variables	I Number of analyst following	II Forecast Accuracy	III Dispersion	IV Reaction	V CAR
Constant	-6.232***	-0.0195***	0.00511***	18.28***	-0.00263*
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0903)
XBRL	0.878***	-0.00106***	0.00108***	-5.061***	-0.00161**
	(0.0000)	(0.0064)	(0.0000)	(0.0000)	(0.0129)
EPS	0.293***	0.00474***	-0.00117***	0.156***	0.000322
	(0.0000)	(0.0000)	(0.0000)	(0.0015)	(0.1517)
ln_totalassets	1.925***	-0.00167***	0.000362***	0.911***	0.000587**
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0269)
leverage	-9.894***	-0.0301***	0.00488***	6.774***	-0.00150
	(0.0000)	(0.0000)	(0.0009)	(0.0000)	(0.8160)
control_numest	.	0.00837***	-0.00164***	-3.341***	-0.000452
	.	(0.0000)	(0.0000)	(0.0000)	(0.3647)
F-statistics	0.000	0.000	0.000	0.000	0.000
Observations	32,167	30,421	26,069	30,795	27,821
R-squared	0.361	0.110	0.090	0.059	0.001

Robust pval in parentheses. *** p<0.01, ** p<0.05, * p<0.1

4.4 Summary results

This section starts with elaborating on the assumptions that have to be met before an OLS regression can be carried out. Adjustments are made when the regression models I use in this thesis do not met one of the assumptions. Once these adjustments are made, the regressions are carried out. Results of the effect of XBRL on analysts' behavior are not consistent in the difference-in-difference test. As expected, the number of analyst following a firm increases, forecasts made by analysts are less dispersed and analyst revision response time decreases. Unexpected, the forecast accuracy and CAR decrease after XBRL implementation.

To verify whether the results are robust, some additional tests are carried out. The first robustness test is mean centering. These tests conclude that the coefficients found in the difference-in-difference test are robust since the coefficients do not change in the mean centering model. The second additional analysis, the fixed effects model, shows that the effect of XBRL adoption is slightly less for the dependent variables number of analyst following and forecast accuracy. For the dependent variables dispersion and forecast revision response time the effect is opposite, the coefficients of the interaction term are strengthened. Indicating that in the fixed effects model, forecasts are less dispersed and analysts need less time to revise their forecasts. The coefficient for the interaction term for CAR is still not significant and for that reason no conclusion can be made. The last robustness test is a model where I carried out regressions of only the treatment sample. Whereas in the other two robustness tests, the signs of the coefficients were the same as in the difference-in-difference tests, this does not hold for the last robustness test. Where the forecasts of analysts were less dispersed in the difference-in-difference tests, in the model without the control sample, dispersions of forecasts arise when XBRL is adopted. In this model, the effect of XBRL adoption on the information content is significant but negative. Indicating that the information content of 10-K filings in XBRL format are less informative. The effect of XBRL on the other dependent variables is the same as in the difference-in-difference model. Namely, the number of analyst covering is increasing, the forecast accuracy is decreasing and the forecast revision time is decreasing. Table 14 provides a summary of the results of the different tests. The plus and minus signs indicate the direction of the coefficients of interest in the different models used. Whereas a dot indicates that there is no significant results found.

The firsts hypothesis of this thesis states that XBRL adoption has no association with the number of analysts covering a firm. This hypothesis can be rejected. The different models indicate that XBRL implementation has a positive effect of the number analysts covering a firm. So, XBRL adoption increases the number of analysts covering a firm. The second hypothesis states the forecast accuracy of analysts increases after XBRL adoption. This hypothesis is also rejected. The coefficients of interest in all models used is negative, indicating that the forecast accuracy of analysts decrease after XBRL implementation.

Table 14 Results of the different tests

Summary results

	Number of analyst following	Forecast accuracy	Dispersion	Response time	CAR
Diff-in-diff	+	-	-	-	.
Mean centering	+	-	-	-	.
Fixed effects model	+	-	-	-	.
Treatment sample only	+	-	+	-	-

The plus and minus signs indicate the direction of the coefficients of interest in the different models used. Whereas a dot indicates that there is no significant results found.

As mentioned before, an explanation could be that analysts still perceive difficulties with interpreting 10-K filings in XBRL format and for that reason make less accurate forecasts. The third hypothesis states that implementation of XBRL decreases dispersion among analysts. Table 14 displays that in three of the four models carried out, dispersion indeed decreases. However, in the model with only the treatment sample, XBRL adoption has a positive association with dispersion, indicating that dispersion increases after implementation of XBRL. So, no consistent evidence is found. In hypothesis four, I mentioned that XBRL adoption reduces the forecast revision response time. Table 14 shows evidence to accept this hypothesis. In all different models, the sign of the coefficient of interest is negative, indicating the number of days the analysts need decreases. The last hypothesis of this thesis states that the information content is higher after XBRL adoption. Table 14 shows that three of the four tests do not give significant results. The last test, carried out with only the treatment sample, indicates that the information content decreases after XBRL adoption. For these reasons, hypothesis five is rejected.

5. Conclusion

In 2008, the SEC issued a new mandate regarding how firms have to format their 10-K filings. Starting in 2009, depending on the public float of a firm, firms have to file their 10-K filings in XBRL format in the United States. This new format is an electronically way of filing information, called XBRL. The SEC argues that the implementation of this new technology has several advantages. They mention that transparency of financial data increases because a common digital language is used. Since the way of formatting information is standardized, less human errors are made and managers should not be able to manipulate data. Another advantage of implementation of XBRL increases the comparability among firms. Because of the fact that 10-K filings in XBRL format are available online, analysts should be able to process the information in financial statements faster. If analysts indeed perceive these benefits is examined in this thesis.

The two most important theories used in this thesis are asymmetric information and the IT productivity paradox. Asymmetric information entails that not all analysts have access to the same kind of information. In the case of asymmetric information, one party has more information available to make forecasts than other analysts. XBRL has as goal to increase the quality of information by standardizing financial statements. So, all analysts have to have access to the same information. The second theory described is the IT productivity paradox. This indicates that in the beginning of a newly introduced technology, errors are made. For that reason, this thesis goes further than previous research, by examining the long-term effect of XBRL adoption. In this thesis, I dive deeper into the possible effects of XBRL implementation on analysts' behavior. Therefore, the research question answered in this thesis is

What is the effect of XBRL adoption on overall analysts' behavior?

To answer this research question, several measures of analysts' behavior are used. These are the number of analysts covering a firm, the analyst forecast revision time and the accurateness, dispersion and information content of forecasts made by analysts. From the sample, voluntary adopters of XBRL are removed to avoid the self-selection bias. Regarding the number of analysts covering a firm, in the different analysis used, the number of analysts is increasing after XBRL adoption. So, the hypothesis formulated in section 2.3.1 is accepted. The results found on the effect of XBRL adoption on forecast accuracy are contradictory to my expectations. After implementation of XBRL, the forecast accuracy decreases. These findings are consistent throughout the different analyses carried out. The hypothesis formulated in section 2.3.2 is therefore rejected. Possible reason for this could be the increased use of firm-specific tags or the difficultness of interpreting XBRL files. On the third proxy used, dispersion, is no clear evidence found. In three of the four analysis carried out, dispersion is decreasing. This indicates that XBRL tends to provide analysts with better information. However, in the analysis without the control sample, dispersion increases. Since no consistent evidence is found regarding the effect of XBRL on dispersion, no conclusion can be drawn. The fourth hypothesis formulated in this thesis is that XBRL adoption reduces the average time needed for analysts to revise their forecasts decreases. Throughout the analysis carried out, I conclude that XBRL adoption indeed reduces the time analysts need to revise their forecasts. CAR is in three of the four analysis insignificant. Since CAR is used as proxy for the information content of analysts, no conclusion can be drawn on the effect of XBRL implementation on the information content of analysts

A few limitations could cause that the evidence found is not consistent. The first limitation of this thesis is regarding the control group of the difference-in-difference test. As mentioned in section 3.3, the treatment and control group are significantly different. Carrying out a difference-in-difference test with a control sample that does not difference significantly from the treatment group could provide less biased results. Another limitation is the external validity. In this thesis, I only used data from the United States. Previous research already showed that there are mixed results on the implementation of XBRL in China and United States. For that reason, it is not possible to assume that results will be the same in other countries. Secondly, not all variables that affect analysts' behavior are included in the model. This causes that the results found could be biased. Further research could reduce this problem by implanting more control variables. Lastly, because of the fact that the SEC does not hold firms liable for errors made in the 10-K filings and that no assurance is needed on 10-K filing in XBRL format, errors in the filings could remain. If firms make errors in their filings, it is unavoidable that forecast accuracy and dispersion are becoming worse. Therefore, the SEC should think about the role of CPAs. CPAs could help to create more accurate financial information and this could subsequently help analysts by making forecasts. However, it is doubtful whether the benefits of assurance outweigh the costs.

Because of the fact that this thesis does not include the number of tags used by firms in creating 10-K filings. Future research could incorporate the number of firm-specific tags used by firms and if this number does affect the behavior of analysts in a different way. Another possibility for future research could be to scrutinize the effect of XBRL assurance on the behavior of XBRL. If such a research is made, it could be discussed whether mandatory assurance of XBRL is effective or not.

Overall, no consistent effect of XBRL adoption on analysts' behavior is found. It is clear from the results above, that analysts' behavior does change after XBRL adoption. However, the direction is not consistent. For that reason, I can answer my research question by stating that XBRL adoption does change the effect of analysts but that the effect is ambiguously. This thesis contributes to previous research in the way that the benefits of XBRL promised to analysts, firms and other stakeholders are still doubtful. For that reason, the SEC should think of next steps to make XBRL an unambiguously success.

Bibliography

Ahrendt, B. (2009, December 16). What are the costs and benefits of XBRL in the financial service industry? *Economie & Informatica*.

AICPA. (2017). *American Institute of CPAs*. Retrieved from http://www.aicpa.org/InterestAreas/FRC/AccountingFinancialReporting/XBRL/Pages/BenefitsandPotentialUsesofXBRL.aspx

Akerlof, G. (1970). The Market for "Lemons": Quality Uncertainty and the Market Mechanis. *The Quarterly Journal of Economics, 84*(3), 488-500.

Baldwin, A., & Brand, T. (2011). The impact of XBRL: A Delphi investigation.

Baldwin, A., Brown, C., & Trinkle, B. (2006). XBRL: An Impacts Framework and Research Challenge. *Journal of Emerging Technologies in Accounting, 3*, 97-116.

Barniv, R. (2009). Does foreign investor demand for information affect forecast accuracy? Evidence from the Chinese stock markets. *Journal of International Accounting, Auditing and Taxation, 18*(2), 101-118.

Barniv, R. (2009). oes foreign investor demand for information affect forecast accuracy? Evidence from the Chinese stock markets. *Journal of International Accounting, Auditing and Taxation, 18*(2), 101-118.

Bergeron, B. (2004). *Essentials of XBRL: Financial reporting in the 21st century* (Vol. 30). John Wiley & Sons.

Bhushan, R. (1989). Firm characteristics and analyst following. *Journal of Accounting and Economics, 11*(2), 255-274.

Bizarro, P., & Garcia, A. (2010). XBRL-Beyond the basics. *The CPA Journal, 80*(5), 62-71.

Blankespoor, E. (2012). *The Impact of Investor Information Processing Costs on Firm Disclosure Choice: Evidence from the XBRL Mandate.* Doctoral dissertation, The University of Michigan.

Blankespoor, E., Miller, B., & White, H. (2014). Initial evidence on the market impact of the XBRL mandate. *Review of Accounting Studies, 19*(4), 1468–1503.

Bloomfield, R. (2002). The ''incomplete revelation hypothesis'' and financial reporting. *Accounting Horizons, 16*(3), 233-243.

Boritz, J., & No, W. (2009). Assurance on XBRL-Related Documents: The Case of United Technologies Corporation. *Journal of Information Systems, 23*(2), 49-78.

Bovee, M., Ettredge, M., Srivastava, R., & Vasarhelyi, M. (2002). Does the year 2000 XBRL taxonomy accomodate current business financial-reporting practice? *Journal of Information Systems, 16*(2), 165-182.

Bramati, C., & Croux, C. (2007). Robust estimators for the fixed effects panel data model. *The econometrics journal, 10*(3), 521-540.

Burnett, R., Friedman, M., & Murthy, U. (2006). Financial reports: Why you need XBRL. *The Journal of Corporate Accounting and Finance, 17*(5), 33-40.

Chen, S., Harris, L., Li, W., & Wu, D. (2015). How Does XBRL Affect the Cost of Equity Capital? Evidence from an Emerging Market. *Journal of International Accounting Research, 14*(2), 123-145.

Cong, Y., Hao, J., & Zou, L. (2014). The Impact of XBRL Reporting on Market Efficiency. *Journal of Information Systems, 28*(2), 181-207.

Cooper, R., Day, T., & Lewis, C. (2001). Following the leader: a study of individual analysts' earnings forecasts. *Journal of Financial Economics, 61*(3), 383-416.

Debreceny, R., Chandra, A., Cheh, J., Guithues-Amrhein, D., Hannon, N., Hutchison, P., . . . Vasarhelyi, M. (2005). Financial Reporting in XBRL on the SEC's EDGAR System: A critique and Evaluation. *Journal of Information Systems, 19*(2), 191-210.

Efendi, J., Park, J., & Smith, L. (2014). Do XBRL filings enhance informational efficiency? Early evidence from post-earnings announcement drift. *Journal of Business Research, 67*(6), 1099-1105.

Efendi, J., Park, J., & Subramaniam, C. (2016). Does the XBRL Reporting Format Provide Incremental Information Value? A Study Using XBRL Disclosures During the Voluntary Filing Program. *A Journal of Accounting, Finance and Business Studies, 52*(2), 259-285.

European Securities and Markets Authority. (2017). *ESMA proposes new digital format for issuers' financial reporting*. Retrieved from https://www.esma.europa.eu/press-news/esma-news/esma-proposes-new-digital-format-issuers%E2%80%99-financial-reporting

Frankel, R., Kothari, S., & Weber, J. (2006). Determinants of the informativeness of analyst research ☆ *Journal of Accounting and Economics, 41*(1-2), 29-54.

Ghasemi, A., & Zahediasl, S. (2012). Normality tests for statistical analysis: a guide for non-statisticians. *International journal of endocrinology and metabolism, 10*(2), 486-489.

Gleason, C., & Lee, C. (2003). Analyst Forecast Revision and Market Price Discovery. *The Accounting Review, 78*(1), 193-225.

Greene, W. (2002). *Econometric Analysis* (Fifth Edition ed.). New Jerset, United States of America: Prentice Hall.

Groenewegen, J., Spithoven, A., & van den Berg, A. (2010). *Institutional economics: an introduction.* London: Palgrave Macmillan.

Grosu, M., & Chersan, I. (2013). Information Risk and Insurance Theory in Financial Audit.

Healy, P., & Palepu, K. (1995). The challenges of investor communication The case of CUC International, Inc. *Journal of Financial Economics, 38*(2), 111-140.

Healy, P., & Palepu, K. (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.

Hong, H., & Kubik, J. (2003). Analyzing the Analysts: Career Concerns and Biased Earnings Forecast. *The Journal of Finance, 58*(1), 313-351.

Ib, C., Jide, I., & Zik-Rullahi, A. (2015). The Impact of XBRL on Financial Reporting: A Conceptual Analysis. *International Journal of Empirical Finance, 4*(2), 78-85.

Ilias, A., Razak, M., & Rahman, R. (2015). The Expectation of Perceived Benefit of Extensible Business Reporting Language (XBRL): A Case in Malaysia. *The Journal of Developing Areas, 49*(5), 263-271.

Irvine, P. (2003). The incremental impact of analyst initiation of coverage. *Journal of Corporate Finance, 9*(4), 431-451.

Jensen, M., & Meckling, W. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics, 3*(4), 305-360.

Kernan, K. (2008). *XBRL around the world*. Retrieved from Journal of Accountancy : http://www.journalofaccountancy.com/Issues/2008/Oct/XBRLAroundTheWorld.htm

Kim, J., Lim, J.-H., & No, W. (2012). The Effect of First Wave Mandatory XBRL Reporting across the Financial Information Environment. *Journal of Information Systems, 26*(1), 127-153.

Kirk, M., Vincent, J., & Williams, D. (n.d.). From Print to Practice: XBRL Extension Use and Analyst Forecast Properties.

Kreft, I., & de Leeuw, J. (1998). Introducing multilevel modeling. *Sage*, 106-114.

Lang, M., & Lundholm, R. (1996). Corporate Disclosure Policy and Analyst Behavior. *The Accounting Review, 71*(4), 467-492.

Lehavy, R., Li, F., & Merkley, K. (2011). The Effect of Annual Report Readability on Analyst Following and the Properties of Their Earnings Forecasts. *American Accounting Association, 86*(3), 1087-1115.

Leuz, C., & Verrecchia, R. (2000). The Economic Consequences of Increased Disclosure. *Journal of Accounting Research, 38*(3), 91-124.

Li, E., & Ramesh, K. (2009). Market reaction surrounding the filing of periodic SEC reports. *The Accounting Review, 84*(4), 1171-1208.

Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics, 45*(s2-3), 221-247.

Li, O., Lin, Y., & Ni, C. (2012). Does XBRL Adoption Reduce the Cost of Equity Capital?

Libby, R. (1981). *Accounting and human information processing: Theory and applications.* Prentice Hall.

Liu, C., & O'Farrell, G. (2013). The role of accounting values in the relation between XBRL and forecast accuracy. *International Journal of Accounting & Information Management, 21*(4), 297-313.

Liu, C., Luo, X., & Wang, F. (2017). An empirical investigation on the impact of XBRL adoption on information asymmetry: Evidence from Europe. *Decision Support Systems, 93*, 42-50.

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

Liu, C., Wang, T., & Yao, L. J. (2014a). XBRL's impact on analyst forecast behavior: An empirical study. *Jorunal of Accounting and Public Policy, 33*(1), 69-82.

Liu, C., Yao, L., Sia, C., & Wei, K. (2014b). The impact of early XBRL adoption on analysts' forecast accuracy - empirical evidence from China. *Electronic Markets, 24*(1), 47-55.

Madden, P. (2009). Standard Business Reporting: An Idea Whose Time Starts Now. *Economic Round-up*(3), 1-18.

Marshall, B., Mortenson, K., Bourne, A., & Price, K. (2010). Visualizing Basic Accounting Flows: Does XBRL + Model + Animation = Understanding? *The International Journal of Digital Accounting Research, 10*, 27-54.

Mikhail, M. B., Walther, B. R., & Willis, R. H. (1999, April 2). Does Forecast Accruacy Matter To Security Analysts? *Thea Accounting Review, 74*, 185-200.

O'Brien, P., & Bushan, R. (1990). Analyst following and institutional ownership. *Journal of Accounting Research*, 55-76.

Pedace, R. (2013). *Econometrics for dummies.* Hoboken: N.J.: John Wiley & Sons.

Peng, J., & Chang, J. (2010). Applying XBRL in an Accounting Information System Design Using the REA Approach: An Instructional Case. *Accounting Perspectives, 9*(1), 55-78.

Pinsker, R., & Li, S. (47-50). Costs and benefits of XBRL adoption: Early evidence. *Communications of the ACM, 51*(3), 2008.

Plumee, R., & Plumlee, M. (2008). Assurance on XBRL for Financial Reporting. *Accounting Horizons, 22*(3), 353-368.

Premuroso, R., & Bhattacharya , S. (2008). Do early and voluntary filers of financial information in XBRL format signal superior corporate governance and operating performance? *International Journal of Accounting Information Systems, 9*(1), 1-20.

Rai, A., Patnayakuni, R., & Patnayakuni, N. (1997). Technology investment and business performance. *Magazine Communications of the ACM, 40*(7), 89-97.

Ravi, R., & Hong, Y. (2014). Firm opacity and financial market information asymmetry. *Juornal of Empirical Finance, 25*, 83-94.

Rock, S., Sedo, S., & Willenborg, M. (2001). Analyst following and count-data econometrics. *Journal of Accounting and Economics, 30*(3), 351-373.

Securities and Exchange Commission . (2009). *SEC Rules for Reporting Financial Statements in XBRL Format*. Retrieved from https://www.sec.gov/rules/final/2009/33-9002.pdf

Shaw, K. (2003). Corporate disclosure quality, earnings smoothing, and earnings' timeliness. *Journal of Business Research, 56*(12), 1043-1050.

Srivastava, R., & Kogan, A. (2010). Assurance on XBRL instance document: A conceptual framework of assertions. *International Journal of Accounting Information Systems, 11*(3), 261-273.

Tan, C., & Shon, J. (2009). XBRL and its financial reporting benefits: Capital Market Evidence. *In th International Conference at The University of Kansas, Lawrence, KS*, 24-25.

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

The Dutch Taxonomy Project. (2008, December). *From NTP to SBR: perspectives on a special project.* Retrieved from www.xbrl-ntp.nl: https://web.archive.org/web/20090327123748/http://www.xbrl-ntp.nl/english/ntptosbr

Trites, G. (2006). Interactive Data: The Impact on Assurance-New Challenges for the Auditing Profession. *XBRL International*.

Vasarhelyi, M., Chan, D., & Krahel, J. (2012). Consequences of XBRL standardization on financial statement data. *Journal of Information Systems, 26*(1), 155-167.

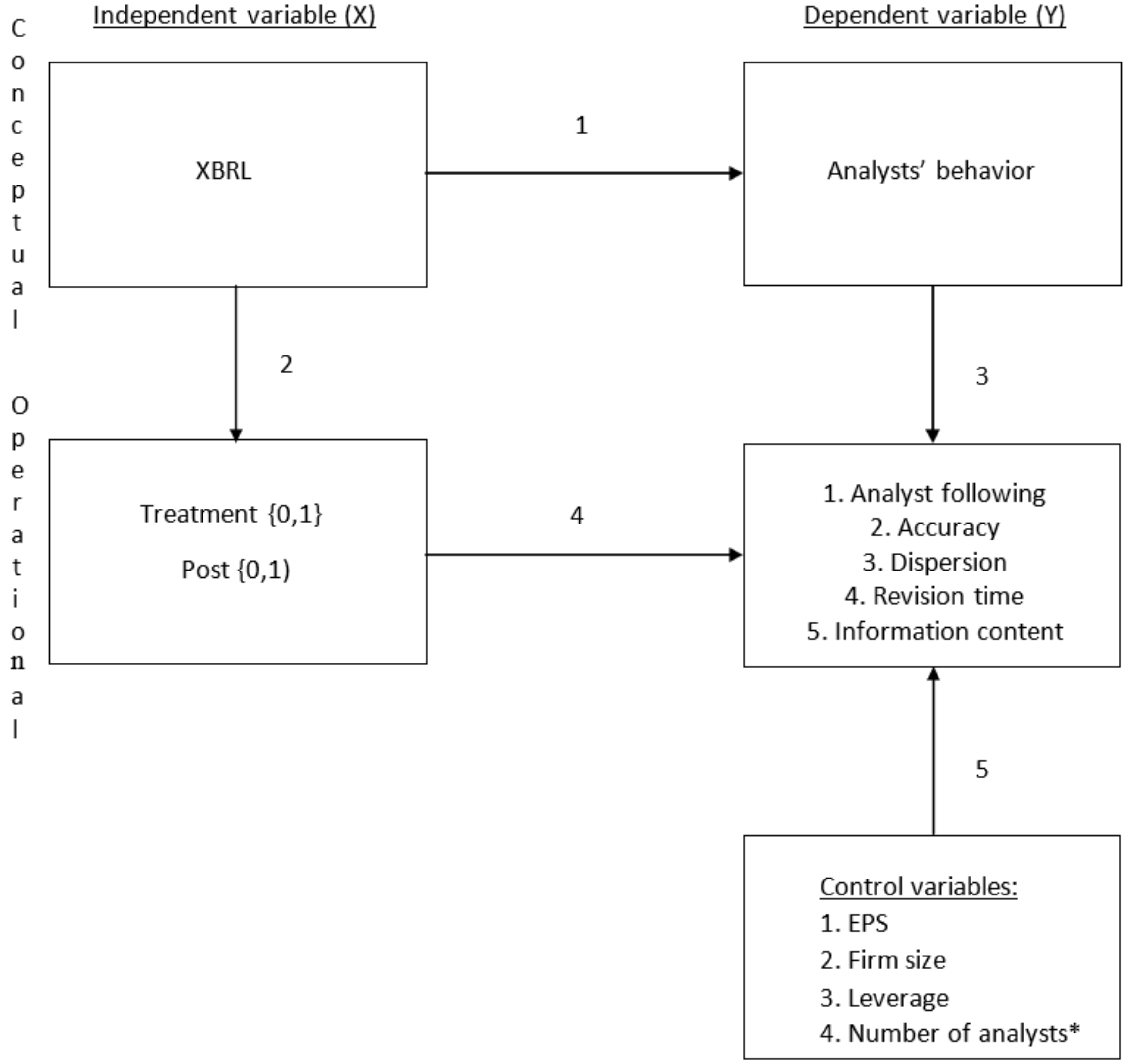
Yoon, H., Zo, H., & Ciganek, A. (2011). Does XBRL adoption reduce information asymmetry. *Journal of Business Research, 64*(2), 157-163.

Yu, M. (2010). Analyst forecast properties, analyst following and governance disclosures: A global perspective. *Journal of International Accounting, Auditing and Taxation, 19*(1), 1-15.

Appendices

Appendix A: Libby boxes

Figure 1 Libby Boxes



*This control variable is only used for regression models two until 5.

Appendix B: Variable definitions

Table 15 Variable definitions

Variable name	Variable description
numest	The number of analyst covering a firm. Computed as the number of analyst following from the most recent consensus analyst forecast measurement date before the 10-K announcement date.
forecast_accuracy	The level of accurateness of forecasts made by analysts. Computed as the forecast error * (-1). Whereas the forecast error is calculated as the absolute difference between actual price of EPS minus the consensus forecasts EPS divided by the year-start share price.
dispersion	The level of understandability of forecasts made by analysts. Computed as the standard deviation of analyst EPS forecast from the most recent consensus analyst forecast measurement date before the 10-K announcement date divided by the year-start share price.
response time	The average time in working days that analysts need to revise their forecasts. The time is computed as the time analysts need after the announcement of the actual EPS until their first revision.
CAR	The three-day window of the cumulative abnormal returns based on the market model.
XBRL	A dummy variable which is 1 for the treatment group (XBRL adopters) and is equal to 0 for the control group (non-XBRL adopters).
post	A dummy variable which is equal to 1 when the fiscal year is 2011 or later and 0 otherwise.
EPS	Variable that controls for the magnitude of earnings. The actual Earnings per Share (EPS) of a firm.
ln_totalassets	Variable that controls for the size of the firm. Computed as the natural logarithm of total assets.
leverage	Variable that controls for the financial risk a firm is facing. Computed as total debt divided by total assets.
control_numest	Variable that controls for the number of analyst covering a firm. Computed as log (1 + number of analysts).

Appendix C: OLS assumptions tests number of analyst following

Table 16 Breusch - Pagan test

Breusch – Pagan / Cook – Weisberg test for heteroscedasticity
H_0: Constant variance

Variables: fitted values of numest
chi2(1) = 10920.70
Prob > chi2= 0.000

Table 17 Durbin - Watson test

Durbin – Watson tests for Autocorrelation	
Number of gaps in sample:	5,937
Durbin – Watson d-statistics (7, 53453) =	1.751938

Table 18 Skewness - Kurtosis test

Skewness / Kurtosis tests for Normality

				-----joint------	
Variable	Obs	Pr (Skewness)	Pr (Kurtosis)	Adj chi2(2)	Prob>chi2
Residuals	53,453	0.0000	0.0000	.	.

Table 19 VIF test

VIF test for multicollinearity

Variable	VIF	1/VIF
XBRL_post	3.20	0.312071
post	2.50	0.400293
XBRL	1.54	0.650148
ln_totalassets	1.41	0.706140
EPS	1.35	0.569726
leverage	1.04	0.986681
Mean VIF	1.84	

Appendix D: OLS assumptions tests forecast accuracy

Table 20 Breusch - Pagan test

Breusch – Pagan / Cook – Weisberg test for heteroscedasticity
H_0: Constant variance
Variables: fitted values of dispersion

chi2(1) = 72932.86
Prob > chi2 =0.000

Table 21 Durbin - Watson test

Durbin – Watson tests for Autocorrelation	
Number of gaps in sample:	5,757
Durbin – Watson d-statistics (8, 49344) =	1.248942

Table 22 Skewness - Kurtosis test

Skewness / Kurtosis tests for Normality

				-----joint------	
Variable	Obs	Pr (Skewness)	Pr (Kurtosis)	Adj chi2(2)	Prob>chi2
Residuals	49,344	0.0000	0.0000	.	.

Table 23 VIF Test

VIF test for multicollinearity

Variable	VIF	1/VIF
XBRL_post	3.25	0.306792
post	2.58	0.387751
ln_totalassets	1.75	0.570547
EPS	1.35	0.574326
XBRL	1.59	0.628731
control_numest	1.48	0.679177
leverage	1.06	0.956789
Mean VIF	1.87	

Appendix E: OLS assumptions tests dispersion

Table 24 Breusch - Pagan test

Breusch – Pagan / Cook – Weisberg test for heteroscedasticity
H_0: Constant variance
Variables: fitted values of dispersion

chi2(1) = 23574,42
Prob > chi2 =0.000

Table 25 Durbin - Watson test

Durbin – Watson tests for Autocorrelation

Number of gaps in sample:	6,901
Durbin – Watson d-statistics (8, 53,311) =	0.9171389

Table 26 Skewness - Kurtosis test

Skewness / Kurtosis tests for Normality

				-----joint------	
Variable	Obs	Pr (Skewness)	Pr (Kurtosis)	Adj chi2(2)	Prob>chi2
Residuals	40,308	0.0000	0.0000	.	.

Table 27 VIF test

VIF test for multicollinearity

Variable	VIF	1/VIF
XBRL_post	3.39	0.294437
post	2.68	0.373012
ln_totalassets	1.75	0.571367
EPS	1.35	0.581184
XBRL	1.58	0.630201
control_numest	1.40	0.716618
leverage	1.07	0.944833
Mean VIF	1.89	

Appendix F OLS assumptions Reaction

Table 28 Breusch - Pagan test

Breusch – Pagan / Cook – Weisberg test for heteroscedasticity
H_0: Constant variance
Variables: fitted values of dispersion

chi2(1) = 6328,22
Prob > chi2 =0.000

Table 29 Durbin - Watson test

Durbin – Watson tests for Autocorrelation	
Number of gaps in sample:	6,618
Durbin – Watson d-statistics (8; 49,062) =	1.378828

Table 30 Skewness - Kurtosis test

Skewness / Kurtosis tests for Normality

				-----joint------	
Variable	Obs	Pr (Skewness)	Pr (Kurtosis)	Adj chi2(2)	Prob>chi2
Residuals	49,062	0.0000	0.0000	.	.

Table 31 VIF Test

VIF test for multicollinearity

Variable	VIF	1/VIF
XBRL_post	3.27	0.306068
post	2.55	0.392234
ln_totalassets	1.75	0.572927
EPS	1.35	0.740789
XBRL	1.60	0.740789
control_numest	1.41	0.740789
leverage	1.06	0.944852
Mean VIF	1.85	

Appendix G OLS assumptions CAR

Table 32 Breusch Pagan - Test

Breusch – Pagan / Cook – Weisberg test for heteroscedasticity
H_0: Constant variance
Variables: fitted values of dispersion

chi2(1) = 2272.83
Prob > chi2 =0.000

Table 33 Durbin - Watson test

Durbin – Watson tests for Autocorrelation

Number of gaps in sample:	5,855
Durbin – Watson d-statistics (8, 37831) =	1.58309

Table 34 Skewness - Kurtosis test

Skewness / Kurtosis tests for Normality

				-----joint------	
Variable	Obs	Pr (Skewness)	Pr (Kurtosis)	Adj chi2(2)	Prob>chi2
Residuals	37,831	0.0000	0.0000	.	.

Table 35 VIF test

VIF test for multicollinearity

Variable	VIF	1/VIF
XBRL_post	3.98	0.250957
post	3.19	0.313944
ln_totalassets	1.95	0.511513
EPS	1.36	0.736119
XBRL	1.60	0.623661
control_numest	1.66	0.602664
leverage	1.05	0.952999
Mean VIF	2.11	