ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics

Master Thesis [programme Accounting and Finance]

Title thesis: Determinants and value relevance of customer base disclosures.

Name student: Joost Huijsdens

Student ID number: 456928jh

Supervisor: Michael Erkens

Second assessor: Jeroen Suijs

Abstract

The goal of this thesis is to examine whether there is a link between competition and customer base disclosure. In addition, I also examine whether customer base disclosures are value relevant to investors. I find that firms operating in more competitive markets are more likely to disclose the size of their customer base. I also find that these disclosures are value relevant. Additional tests suggest that active user disclosures are more value relevant than traditional customer base disclosures. The implication is that despite often being represented as generic information about a company, the number of customers a firm has is value relevant in addition to traditional accounting figures. Managers should be conscious of this value relevance in their disclosure decisions if they were not already before. Not being aware of this value relevance could lead managers to disclose information that affects the perceptions of shareholders without them realizing.

Key words: Voluntary disclosure, non-GAAP disclosure, customer base disclosure, competition, value relevance.

1. Introduction

Whether they are the tenants of a landlord, the 265 million Walmart shoppers (Walmart, 2020) or Facebook's 2,6 billion active users (Facebook, 2019), all successful companies have customers in one way or another. These customers form the basis of most revenue a company generates. The way firms generate revenue through their customers can be different per company. Some companies generate revenue through a limited number of key customers; others generate their revenues through a large and varied base of customers. In addition, most companies generate their revenues directly through monetary transactions with their customers. Other firms attempt to generate a large customer (or user) base first, and then generate revenues through their large customer (or user) base.

Some companies disclose the number of customers that they have in their annual 10-k report. I examine which companies disclose information about the size over their customer base and whether these disclosures have any value relevance. In addition to this, I also examine the difference in value relevance for the firms that report a customer base number and the firms that report the number of active users. The main research question that I answer in this thesis is formulated as follows:

Does competition influence a firm's decision to disclose the size of their customer base and are these customer base disclosures value relevant?

Consistent with several previous papers (Li, 2010; Burks, Cuny, Gerakos, & Granja, 2018), I hypothesize that the amount of competition that a firm faces in their main operating market affects their disclosure decision. There are reasons why competition can have a positive or a negative effect on the likelihood of a firm disclosing the size of their customer base. Competition could have a positive effect because customers are more valuable in more competitive markets. It could also have a negative effect because firms that face less competition might disclose the size of their customer base to signal the amount of market power that they have. A previous study has found that measures of market penetration are value relevant (Amir & Lev, 1996). In order to test the this first hypothesis a probit regression using ten years of data from the S&P 500 index is carried out. This regression will result in an overview of the effect of competition on the probability of disclosure. In this regression I will account for various control variables like firm size, analyst following and capital market dependency. The data on which firms disclose will be hand collected from the SEC Edgar database. The data on control variables will be collected through various online databases. The findings suggest that firms operating in more competitive markets are more likely to disclose the size of their customer base. This leads me to accept the first hypothesis. Competition does indeed affect a firm's decision to disclose the size of their customer base. This finding contributes to the relatively small amount of existing literature that examines a link between competition and disclosure. The findings are in line with (Li, 2010) who also found a link between competition and disclosure.

Besides what affects the decision to disclose I will also examine the value relevance of these disclosures. A large customer base is often described as a very important intangible asset (Gupta, Lehmann, & Stuart, 2004). However, this does not have to mean that the reported customer base is value relevant. It is also possible that all the value that this customer base provides is already represented in traditional accounting figures. I hypothesize that the reported customer base is value relevant. In addition to that, I also hypothesize that the reported customer base is more value relevant for companies that report active users instead of just customers. The value relevance is tested using the Ohlson model. The Ohlson model attempts to model the share price of a company using several accounting values. In order to test the value relevance, I expand the Ohlson model with a variable that documents the growth in reported customer base. The data on the reported customer base figures will be hand collected through the SEC Edgar database. The data on earnings, book values market values will be collected through the Compustat and CRSP databases.

Based on ten years of data from the S&P 500 index, the growth in reported customer base does appear to be value relevant. This confirms the second hypothesis. Reported active users have a higher incremental value relevance compared to reported customer bases. These results do fall in line with other studies that examined the value relevance of similar customer measures (Gupta et al., 2004; Choi, Lee, Yoo, & Yoo, 2019).

Answering the research question can be of interest to various stakeholders. First, the question can be of interest to academia that study the link between competition and disclosure. Secondly, it can interest managers of companies. The number of customers is sometimes listed as a key performance measure. However, it is usually listed within the general description of the company. This could imply that some managers might disclose this figure without viewing it as a meaningful performance measure. If the size of the customer base of a firm turns out to be value relevant it could indicate that investors do consider it to evaluate performance. Finally, the answer to the research question might be of interest to regulators. For most industries, there is little or no regulation regarding these kinds of disclosures. There is relatively low transparency as to how the number of customers is calculated. If this number influences the investors' perceptions of a company it might be important to regulators to ensure that the figure is accurate and not susceptible to manipulation. The most important implications of these findings are as follows. The reported number of customers is value relevant. This is especially true for firms that report an active users number. For managers this means that they should be conscious about their decision disclose this information, since it is more than just general information about the company. The second implication relates to regulations. There are relatively few regulations regarding these disclosures; an opportunistic manager could for instance attempt to inflate the number of customers by counting customers that have not bought anything from the firm in years. Regulators could look into this to ensure that disclosures about a firm's customer base remain accurate.

2. Theoretical background

Companies disclose information on both mandatory and voluntary bases. Examples of voluntary disclosures are conference calls, social media-based disclosures and information inside mandatory filings that goes beyond what is mandated by law. These voluntary disclosures are a popular topic of research for accounting researchers. Researchers try to answer questions related to why managers chose to disclose voluntarily, the importance of these disclosures to investors and the consequences voluntary disclosure.

There are several hypothesised motivations for voluntary disclosures. Healy and Palepu (2001) review prior literature and summarize multiple motivations. The first motivation is that managers disclose more to reduce information asymmetry prior to capital market transactions. In this scenario managers disclose information prior to public offerings in order to attempt to reduce the risk that investors assign to their stock. When an investor or analyst has relatively little information about a company, the investor faces a higher risk investing in that company. By increasing the information that an investor has, managers hope to reduce the risk that investors assign to their stock. The second motivation is that managers disclose more to avoid undervaluation of the stock. Whenever a manager believes that his or her firm is undervalued, the managers can decide to disclose more information about his company to inform investors on the value of the company. A firm could be undervalued when the value of the firm is not properly captured in the disclosure that the firm gives out. Undervaluation can be a serious problem for managers; it can cause issues with debt covenants resulting in the manager to lose control of (part of) the company. Another problem for the manager that can be caused by undervaluation is tied into the third motivation: managers disclose strategically in order to maximize their personal compensation. When a manager has company shares in his compensation package, undervaluation effectively reduces his compensation. Another way that a manager can use disclosure to maximize his own personal compensation is when his compensation package includes stock options: it is beneficial to disclose bad news (that can cause the share price to drop) when he is granted stock options. This could result in options to be granted with a lower strike price. Subsequently, when the manager is close to exercising his options, he will disclose good information. This is done to increase the share price (further) beyond the option strike price. This increases the value of the managers options and thus increases his compensation. The fourth motivation can be a legal threat that can lead to an increase or a decrease in disclosure. Managers are required to disclose material events when they become aware of them. Failing to disclose important information can lead to litigation; disclosing wrong and or intentionally misleading information can also lead to litigation. The fifth motivation is that managers simply want to signal how talented they are. This is not just because managers are vain and would like to let everyone know how talented they are. Being known as a talented manager increases their own future job prospects. The sixth and final motivation that Healy and Palepu list is the proprietary cost of disclosure. Managers want to avoid disclosing information that can be of use to their competitors. This is the only one of the six motivations where there is no conflict of interest between the shareholders, debtholders, and manager of a company. An extension of the proprietary costs motivation is the hypothesis that

competition drives disclosure decisions. The hypothesis is that the potential proprietary costs are higher for companies with higher competition. The treat of giving away company secrets is higher when you have more competitors that monitor a firm's disclosures. Li (2010) finds that competition affects disclosure decisions in several ways. Another recent study finds that within the banking industry increased competition does lead to an increase in disclosure (Burks et al., 2018).

Value relevance tests are common for accounting figures. The goal of standard setters when deciding on different accounting treatments is to produce information that is both relevant and reliable (Barth, Beaver, & Landsman, 2001). A figure is deemed relevant once it influences the decision making of a financial statement user. A figure is reliable if it accurately represents what it aims to represent. Value relevance tests are testing both the reliability and relevance of information. After all, an investor will not base his decision making on non-reliable or irrelevant data. In order to see whether investors react to voluntary disclosure researchers conduct studies regarding the value relevance of disclosures. This is typically done by checking for a contemporaneous association between the disclosed variable and the market value of equity. When there is a relation after controlling for earnings and other factors one would conclude that the disclosed variable is value relevant. An example of this is a study that found that environmental performance can be value relevant using this method (Hassel, Nilsson, & Nyquist, 2005). Alternatively, one can research abnormal stock returns around the date and time of voluntary disclosure. However, this approach is not always practical when a company releases a lot of different data at the same time during a press release or SEC filing. Because there are so many different forms of voluntary disclosure it is not possible to say that either none or all voluntary disclosure is value relevant.

The economic consequences of voluntary disclosure are also summarized by Healy and Palepu (2001). Research has found several economic consequences of voluntary disclosure. The first consequence is increased stock liquidity (Leuz & Verrecchia, 2000). Increased disclosure leads to more liquidity of the companies' stocks. This could potentially relate to the attention effect, both institutional and individual investors tend to trade stocks that are in the news more often. More extensive and more frequent disclosure could lead to stocks being in the news more and thus getting traded more. This increase in trades generally reduces the bid-ask spread of stocks. The second consequence is reduced cost of capital (Lambert, Leuz, & Verrecchia, 2007). Research has found that voluntary disclosure reduces the cost of capital, which is especially interesting for firms that seek to issue equity or debt. Managers are aware of this; research has

found that companies disclose more prior to public debt and equity offerings (Frankel, McNichols, & Wilson, 1995). The third consequence is increased information intermediation. Although there is not a lot of research done on this consequence the theory behind it is sound. Increased disclosure gives analysts more resources to work with, which allows them to deliver more value when analysing companies; because they have more information to base their analysis on and because these analysts can now also fill a roll in selecting and prioritizing information when there is potentially an overflow of information for investor to deal with.

A popular form of voluntary disclosure are accounting based figures that are derived from, but not in line with, the Generally Accepted Accounting Principles (GAAP). These disclosures are often referred to as non-GAAP disclosures. Although non-GAAP disclosures are often seen as a subset of voluntary disclosure it should be noted that not all non-GAAP disclosure is voluntary. For example, since 2002 S&P requires that all companies in their equity indexes report core earnings. Core earnings is a non-GAAP earnings measure which includes several deviations from GAAP earnings. The number of firms that disclose non-GAAP figures has increased over the past decade (Black, Christensen, Ciesielski, & Whipple, 2018). Examples of financial non-GAAP disclosures are non-GAAP earnings per share (EPS), non-GAAP revenues and EBITDA. There has been a large amount of research on non-GAAP disclosures (Black et al., 2018). As the amount of companies reporting non-GAAP figures has increased so has the interest of academics in these disclosures. Most of this research is focussed on non-GAAP earnings or non-GAAP EPS but there is also an increasing interest in other financial non-GAAP disclosures such as EBITDA. For example, several researchers found that firms reporting EBITDA have a larger analyst following and lower information asymmetry (Cormier, Demaria, & Magnan, 2017). Just like the broader research on voluntary disclosure non-GAAP research focuses on the motivations for non-GAAP disclosure, the value relevance of non-GAAP disclosure and the consequences of non-GAAP disclosure.

Research examining the motivations behind non-GAAP disclosures mostly boils down to one question. Do managers disclose non-GAAP figures to inform certain stakeholders more effectively? Or do these managers opportunistically use non-GAAP figures to hide bad performance? A lot of research has been carried out on this subject. Isidro & Marques (2015) found that, within Europe, managers are more likely to issue non-GAAP earnings when GAAP earnings fail to meet important earnings benchmarks. This indicates that managers use these non-GAAP figures to beat the benchmarks sets by analysts. Additionally, this behaviour of just meeting benchmarks using non-GAAP measures is more common in more developed countries

with stricter regulations and more efficient law enforcement (Isidro & Marques, 2015). This suggests that managers resort to non-GAAP measures when manipulating GAAP figures is not an option. However, there is no shortage of research which argues the opposite that non-GAAP measures are not only disclosed in an opportunistic attempt to mislead investors, but rather help to inform them more effectively. For example, Curtis, McVay, & Whipple (2013) find that most firms disclose non-GAAP earnings to better inform investors. Only a minority of firms seem to use non-GAAP disclosure in an opportunistic way. Another motivation related area of research is not whether managers use non-GAAP disclosure opportunistically, but rather whether the usage of non-GAAP figures is manager or analyst driven. Or in other words, is it the managers that push these non-GAAP metrics or are managers simply responding to the demand form analyst. The evidence on this subject is mixed. Some managers claim that they report non-GAAP metrics because they more accurately represent their own internal metrics (Black et al., 2018). While other managers claim that they report non-GAAP metrics because analysts prefer to use these metrics as opposed to or in addition to their GAAP counterparts. The evidence is also mixed on the analysts part. Some analysts claim that they use non-GAAP data because managers put emphasis on it, while other analysts claim that the managers reporting choices have very little impact on the figures that they decide to use (Black et al., 2018).

The value relevance of financial non-GAAP disclosure is often compared to their respective GAAP counterparts. This is especially true for EPS. When there is no relevant GAAP counterpart, the research often focusses on value relevance in addition to traditional accounting figures. The arguments in favour of non-GAAP disclosure being more value relevant are generally that by deviating from inflexible accounting standards managers can communicate more effectively to stakeholders. The argument for GAAP disclosure being more value relevant is that managers deviate from GAAP in an opportunistic manner. Managers could try to do this in order to present their performance as more positive than their real performance. An example of research into the value relevance of non-GAAP figures are Wieland, Dawkins, & Dugan (2013) who researched whether core earnings (a proprietary non-GAAP earnings measure) are value relevant to investors. They did this by checking for an association between core earnings and stock price returns. The result was that this non-GAAP earnings measure is more value relevant than GAAP earnings. Another example of this kind of research is a paper by Venter, Emanuel, & Cahan (2014). This paper examines the value relevance of headline earnings compared to GAAP earnings. Headline earnings is yet another non-GAAP earnings measure that is used in South Africa. By regressing the book value of equity and the GAAP-earnings or

headline earnings on the share price and subsequently comparing the R2 of these two regressions they found that these headline earnings (the non-GAAP measure) are more value relevant than GAAP-earnings. Both papers are in line with most other papers on this subject that overwhelmingly find these non-GAAP earnings to be more value relevant than their GAAP counterparts. This evidence also circles back to the motives section. When non-GAAP earnings are more value relevant than their GAAP it indicates that these disclosures are in fact informative to investors.

The consequences of financial non-GAAP disclosure that are being researched are different from the consequences of voluntary disclosure in general. Consequences of non-GAAP disclosure are generally more focused at whether different types of investors are influenced by opportunistic non-GAAP reporting. The expectation is that institutional investors are capable to see through opportunistic non-GAAP reporting, while less sophisticated individual investors might be fooled by opportunistic non-GAAP reporting. The evidence on this subject is mixed. Frederickson and Miller (2004) found that within an experiment setting less sophisticated investors are more heavily influenced by non-GAAP earnings. However, Christensen, Drake, and Thornock (2014) found that short sellers, who are generally considered sophisticated investors, trade more surrounding earnings releases that include non-GAAP figures.

Note that all the papers mentioned about non-GAAP figures have examined financial non-GAAP metrics so far, but non-GAAP figures do not have to be financial in nature. Some examples of non-financial non-GAAP disclosures are CO_2 emissions, number of new stores opened or the amount of web-traffic that a webstore receives. GAAP mainly applies to financial information only. As a result of this, most non-financial disclosures can be considered as non-GAAP disclosures. Therefore, most of these non-GAAP non-financial disclosures are often simply referred to as non-financial disclosures.

Most papers on non-GAAP disclosures are based on financial non-GAAP figures, but there is some research on non-financial disclosures. The vast majority of this non-financial research is focussed on a few topics. These topics are Corporate Social Responsibility (CSR), environment and intellectual capital. Just like the general disclosure research, these studies typically examine the motivations behind these disclosures, the value relevance of these disclosures and the consequences of these disclosures.

The potential motivations for non-financial disclosure depend on whether the disclosure is done voluntarily or whether it was required by law or other regulations that a firm must adhere to. An example of mandated non-financial disclosure is CSR disclosure in Denmark where large

firms are required to disclose their CSR policy since 2008. If they do not have a CSR policy, firms are required to disclose that they do not have a CSR policy. Whenever companies are required by law to disclose information it is not necessary to research a manager's motivations behind disclosing information. In this case it is more common to research the usefulness or effect of the disclosure. The motivations for voluntary non-financial disclosure are similar to the motivations for voluntary disclosure in general. Commonly researched motivations are the capital market motive and the proprietary costs motive. Research has found that managers will disclose more non-financial information in order to benefit their company's position on capital markets (Gao, Dong, Ni, & Fu, 2015). The theory behind this is that these non-financial disclosures can include information that is not reflected in a firm's financial statements. This additional information can make investors perceive this firm's stock as less risky. The reduced risk will gain the firm several capital market benefits in debt and equity markets. At the same time managers are worried that disclosing to much proprietary information could help competitors (Rezaee & Tuo, 2017). Both the capital market motive and the proprietary costs motive are supported by survey data (Graham, Harvey, & Rajgopal, 2005). This contradiction leads managers to attempt to balance their disclosures optimally to maximize capital market benefits while minimizing proprietary costs.

However, for this motivation to hold any ground these non-financial measures should be value relevant. Investors will not assign lower risk to a stock based on information that they deem to be irrelevant. Research has shown that non-financial disclosures can be value relevant by explaining market value fluctuations that cannot be explained by earnings (Amir & Lev, 1996; Ittner & Larcker, 1998). However, the wide variety of non-financial performance metrics makes it hard to make generalized statements about value relevance of non-financial disclosure. One example of a study that has found non-financial information to be value relevant is a study by Ittner and Larcker (1998) who found that customers satisfaction can be used to explain variations in stock prices that are not explained by book values. There is also no shortage of research that examines the value relevance of intellectual capital disclosure. Vafaei, Taylor, and Ahmed (2011) find that intellectual capital disclosure is indeed value relevant based on data from several different countries.

The consequences of non-financial disclosure that are researched are not always capital market related. For example, research on CSR can also focus on the effect of disclosure on environmental performance. One example of this is Luo & Tang (2014) who examined whether firms that have a higher level of carbon disclosure had better underlying environmental

performance; they found that there was indeed a positive relation between disclosure and performance. This result is consistent with signalling theory. According to signalling theory good performers will disclose their performance in a way that bad performers cannot. This is a way for good performers to distinguish themselves from bad performers for investors. The research on consequences that are related to capital markets investigates three different potential consequences, which are the same as those for disclosure research in general. The first two consequences are a lower cost of capital and a larger analyst following (Dhaliwal, Li, Tsang, & Yang, 2011). The third consequence that is often examined is higher stock liquidity (Gao et al., 2015).

One specific form of non-financial non-GAAP disclosure is customer base disclosure. Many firms have an active base of customers that keep buying their products or keep using their services. This customer base can be seen as an important intangible asset of a firm (Gupta et al., 2004). While some of the value of this customer base is potentially captured via revenues, the full value of having a customer base is not always captured by a firm's financial statements. This is evidenced by previous studies that have found disclosures regarding customer relationships to be value relevant for several industries. For example, research has found that customer equity (a calculated measure of the value of a firm's customer base) does have incremental value relevance beyond earnings within the mobile telecom industry (Choi et al., 2019). The extra value that having a larger customer base appears to provide can be caused by several factors. For example, network effects create value that is not captured by financial statements (Rajgopal, Venkatachalam, & Kotha, 2003), while Amir & Lev (1996) find that measures of market penetration can be value relevant. A relatively new form of customer base disclosure is *active users* disclosure. This kind of disclosure is most prevalent in the technology sector. Technology firms who provide software services will report the number of users that actively use their services. This is usually dome through either a Daily Active Users (DAU) figure or a Monthly Active Users (MAU) figure. These active user disclosures are unique, because with these kinds of disclosures there does not have to be a monetary transaction for someone to be counted as an active user. Typically, the business model of these companies relies on one or multiple of three revenue sources. The first revenue source is selling advertising space to companies that wish to advertise to their users. The second revenue source is selling user data to different parties. The third revenue source is upselling premium features to their active user base. The usage of this metric appears to be increasing in recent years. There is no research on this subject yet, based on the fact thatthe number of companies that disclose the

DAU or MAU figure is still quite limited and it is a relatively recent trend. However, there is some anecdotal evidence which suggests that the number of active users is value relevant to equity investors. For example, in April 2020 the share price of videoconferencing software provider ZOOM increased substantially after it reported a large increase in *active users* (CNBC, 2020).

Given this value relevance one might expect managers to disclose information about this in order to reduce information asymmetry and reduce cost of capital. It is true that some companies disclose the size of their customer base in their annual reports. There has also been some research on customer disclosure. Research has found that web-based disclosures related to customer value do reduce implied cost of capital (Orens, Aerts, & Lybaert, 2013). Another study has found that companies are more likely to voluntarily disclose names of large customers prior to SEOs (Ellis, Fee, & Thomes, 2012). This is consistent with the hypothesis that managers disclose information in order to improve access to capital markets. Nevertheless, similar to other forms of disclosure, proprietary costs also appear to influence a manager's disclosure decisions. For example, Ellis et al. (2012) found that firms with higher potential proprietary costs are less likely to disclose information about large customers. They did this by assuming that firms with high R&D cost, advertising costs and intangibles face higher proprietary costs.

Overall, the research on this type of disclosure is still quite limited. There is not a lot of information on the motivations and consequences of customer base disclosure. The studies on the value relevance of customers often rely on measures that require additional information that is typically not found in a company's annual report. This thesis aims to contribute to disclosure literature by investigating what companies disclose the size of their customer base in their annual report. Consistent with several studies (Li, 2010; Burks et al., 2018), I hypothesize that there is a link between competition and disclosure. In a more competitive market, a larger customer base is more valuable. The cost of maintaining a large customer base is higher because customers have more options to switch to another supplier. The opposite can also be true. In a market with low competition a large customer base can signal that a firm has market power. The firm could theoretically use that power to drive up prices and increase profit margins. As discussed in section 3, Amir & Lev (1996) found market penetration to be value relevant for the telecommunications industry. Being aware of the amount of market power that a firm has can also be valuable. Because of these two contradicting rationales behind the effect of competition on customer base disclosure the hypothesis being tested is two sided.

H1: Firms customer base disclosure decisions is related to the competitiveness of the market in which they operate.

Based on the literature discussed in section 3 it would be reasonable to expect the number of customers to be value relevant information but most of the literature mentioned in section 3 does not directly measure the value relevance of just the number of customers. For example, Choi et al. (2019) investigates the value relevance of 'customer equity'. This is a term that also includes the average revenue per customer, average contribution margin, costs of retaining customers, customer retention rate and new customer acquisition costs. Customer equity thus requires much more data and assumptions in order to calculate. Not all firms disclose this data in their SEC filings. Therefore, there is a difference between customer equity and just the number of customers a firm has. However, the literature mentioned in section 3 does indicate that there might be a positive relation between the number of customers and firm performance. Because of this the second hypothesis is formulated as a positive relation. To examine whether the number of customers is value relevant the following hypothesis will be tested:

H2: The (change in) number of customers which a firm has provides value relevant information beyond financial statements to explain contemporaneous market value of equity.

There is also the possibility that the value of having a large customer base is already mostly reflected in financial statements. In this case the null hypothesis would hold.

It should be noted that customer base disclosure can be grouped in two different forms. The first form is a simple disclosure of the number of customers. An example of this would be: "we serve approximately 30 million customers". The second form is disclosure about the number of *active users* a company has. This is a more recent form of disclosure that is mostly used by technology companies. The key difference between *active users* and the previous form of disclosure is that there does not have to be a monetary transaction between the user and company for a user to be counted towards the number of *active users*. Because there is a considerable difference between these two types of customer disclosure and the type of firms that disclose it, the value relevance for these two types of disclosure will be tested separately. The hypothesis is that *active users* are more value relevant than customers. *Active users* are hypothesized to be more value relevant because it can be used to value intangible assets which are typically considered harder to value (McKinsey & Company, 2016). New software can be difficult to value and one piece of information to assess the value of a of new software can be the number of *active users*. In addition to that, having a large number of *active users* is usually the basis of the key revenue streams for the firms that report it. The main revenue streams of

many companies that report this metric are usually either selling user data, selling advertising space or upselling free users with more premium features. A third reason why *active users* are hypothesized to be more value relevant is because the term more reliably represent what it claims to represent. MAU counts users that are active at least once a month. However, someone who bought products five years ago could theoretically still be considered a customer. This degree of freedom might make the customer measure less reliable. It is this reliability that makes up a key element of value relevance (Barth et al., 2001). The final reason why active users disclosure can be more value relevant is because active users disclosure is more specific to certain business models. Almost any company can release customer base disclosure, but only companies with specific business models can release meaningful active users disclosure. The fact that active users disclosure is generally more specific to a firm than customer base disclosure could be another potential cause of higher value relevance. *Active users* can both be relevant to investors in equity and debt. However, the equity investors are usually the ones who will see most of the profit if the company becomes profitable. The future profitability that *active users* might show is therefore most relevant to equity investors. Investors in debt are generally more interested in the financial stability of the company. This is because they are mainly interested in whether the company will default on loans and bonds or not. The number of *active users* appears to provide less information about the financial stability of a company as opposed to potential future profitability.

3. Research Design

To test whether competition has an influence on a firm's disclosure decisions a regression analysis will be used. A probit regression will be used to find out what variables contribute to a firm's disclosure decision. The regression used is the following:

$$DISCLOSE = \beta 0 + \beta 1 * HHI + \beta 2 * SIZE + \beta 3 * ISSUE + \beta 4 * INST + \beta 5 * MTB + \beta 6 * ANALYST \quad (1)$$

Where DISCLOSE is a dummy variable with the value of 1 when a firm does disclose the size of their customer base. The measure used for competition and the variable of interest is the Herfindahl-Hirschman Index (HHI). The HHI is a measure for market concentration which is used to measure market competitiveness. This measure is commonly used by researchers who study the effect of competition (Ali, Klasa, & Yeung, 2014; Li, 2010). The HHI is calculated as follows: $HHI = \sum S_i^2$. Where S_i is market share of a company expressed as a whole number. This means that a monopoly industry would have an HHI of 10.000, while an industry made up

of 100 equally sized competitors would have an HHI of only 100. A lower HHI indicates a higher degree of competitiveness. The HHI will be calculated per GICS sub-industry per year using Compustat data. Because the distribution of HHI is very skewed, the natural log of a firms HHI will be used.

To make sure that the results measured are not a result of other omitted variables, several different control variables will be included. It has been documented that firm size does have an influence on a firm's voluntary disclosures (Baginski & Hassell, 1997). Therefore, the first control variable that is included is firm size (SIZE). Another factor that influences a firm's disclosure decisions is their reliance on capital markets. Firms disclose more information in the years prior to public equity or debt offerings (Frankel et al., 1995). They do this to lower the risk that investors assign to their company. This helps reduce the cost of capital that the company incurs during public offerings. I control for this in the same way that Li (2010) does, by including a dummy variable (ISSUE) that is equal to 1 if a firm issues public equity or debt in a subsequent two-year period. I also control for corporate governance by using the percentage of institutional investors as a control variable (INST). Institutional investors are generally seen as more active investors (Li, 2010) and they do a better job at monitoring a company, which improves corporate governance. Another influence on a firm's disclosure policies is the number of analysts that follow the firm (Baginski & Hassell, 1997). To control for this influence, I include ANALYSTS as a control variable, which counts the number of analysts that follow a firm. Finally, I control for a firm's information asymmetry by including market-to-book ratio (MTB) in the regression. There is a higher demand for disclosure from firms with a higher degree of information asymmetry, which leads these firms to disclose more (Verrecchia, 1990). Besides these control variables I also control for firm year fixed effects.

There are multiple different methods of examining the value relevance of accounting information. The most popular methods are the Return Regression Model (RRM) and Price Regression Model (PRM), the PRM is also often referred to as the Ohlson Model.

The RRM regresses the difference in EPS on stock returns. The differences in EPS can be either the year-on-year change in EPS or the difference between the real EPS and the analyst expected EPS. Using the difference between the real EPS and analyst expectation of EPS is the most common and assumes that the analyst expectation of EPS is already priced into the market. Therefore, the new information would be the difference between the real EPS and the analyst expectation of EPS.

A method that is commonly used in many value relevance studies and is often referred to as the Ohlson model. This methodology is based on a paper by Ohlson (1995) in which he suggests this equity driven model as a way of valuing companies. Based on three assumptions Ohlson develops a way to model the share price or the market value of equity of a company. These three assumptions are as follows. First of all, the market value is the discounted price of all future dividends. Secondly, all changes in book values are incorporated in either earnings or dividends. The third assumption is that abnormal earnings are never persistent. To examine the value relevance, a regression explaining the share price of a affirm using the book value and earnings of that firm is performed. Then, the variable of interest is added to the basic Ohlson model and researchers examine whether the adjusted R2 increases and whether the variable of interest is significant. This method has been used to research the value relevance of several different customer related variables (Choi et al., 2019; Amir & Lev, 1996). The usage of this model is not just limited to customer related variables. There are many papers which use this method to test the value reliance of a wide range of variables. One example of this are Franzen & Radhakrishnan (2009) who use the model to examine the value relevance of R&D. Another example is research which investigates the effect of adopting international accounting standards on value relevance (Barth, Landsman, & Lang, 2008).

All but one of the previously mentioned value relevance papers so far uses (a variant of) the Ohlson model to come to their conclusions (Choi et al., 2019; Hassel et al., 2005; Rajgopal et al., 2003; Venter et al., 2014; Wieland et al., 2013). This popularity is not just exclusive to the papers that I have cited, it appears that the Ohlson model is the most used model to test value relevance of different types of disclosures. I will apply the Ohlson model to test the value relevance of the size a company's customer base. First, the regression model based on equitation (2) is performed. Secondly, I will add the variable CUSTGROW to equation (2) and which gives equation (3). CUSTGROW measures the increase in size of a companies reported customer base. Then a regression will be performed based on equation (3). Finally, I will look at the difference in adjusted R2 between the two regressions and look at the whether the (CUSTGROW) variable is significant or not.

$$MVPS = \beta 0 + \beta 1 * BVPS + \beta 2 * EPS \quad (2)$$

$$MVPS = \beta 0 + \beta 1 * BVPS + \beta 2 * EPS + \beta 3 * CUSTGROW \quad (3)$$

Where MVPS is the market value of equity per share, BVPS is the accounting book value per share, EPS is the accounting earnings per share and CUSTGROW is the percentage growth of

a firms disclosed customer base. The reason for using per share values is to avoid larger companies from weighing to heavily in the model estimation. The coefficient of interest here is $\beta 3$. If it is significant and adding it causes an increase in adjusted R2 than the size of a firm's customer base can be considered value relevant.

The research will be carried out using all S&P 500 firms as the sample. This index was chosen to get a sample that is representative of a wide variety of industries. By maintaining an American sample potential issues with different regulation and different legislators between countries are largely avoided. Utility firms will be excluded from the sample. These firms are often required by law to share information regarding the size of their customer base. This makes the disclosure by these firms non-voluntary. Data regarding their customer base disclosures for the years 2010-2019 will be hand collected from their 10-k forms through the SEC EDGAR database. This collecting is done by reading through every firms 10-k filings from 2010 through 2019. Due to time limitations hand collecting the data has to be sped up using a couple of techniques. First of all, the 10-k filings are searched for certain key words like: customers, clients, users, active accounts and sometimes others depending on the industry that the firm operates in. This is done because the disclosures that are sought often include these words. Secondly, I still looked through every report fully to see if I could find any disclosure that did not include the key words mentioned previously. Additional attention was paid to certain sections of the 10-k that were more likely to include the disclosure. These sections are Item 1- Business, Item 6 – Consolidated financial data and Item 7 – Management's Discussion and Analysis of Financial Condition and Results of Operations. If a firm only discloses the number of customers or users per business unit, the amounts are added up to get a total figure. This adding up is only done when a firm discloses the number of customers for all of their business segments. This is done this way, because the goal is to examine a firms customer base disclosure and not the disclosure of individual business segments. The data collection is done on a firm by firm basis.

Data for the book value of equity and earnings will be collected through Compustat. The market value per share will be retrieved through CRSP. As for the control variables, firm size and market to book ratios will also be collected through Compustat. The amount of analyst that are following each firm will be retrieved from the IBES database. The percentage of institutional ownership and the share price will be retrieved through the Thomson Reuters database. Equity and debt offerings are retrieved through compustat data. Although this data is not available directly it is inferred from different datapoints. A significant increase in outstanding shares or

long-term liabilities indicates a public equity or debt offering. A significant increase is defined of an increase larger than 5 percent.

4. Results

The original sample consists of all S&P 500 firms and all firm years within the period 2010-2019. Only 480 out of 500 firms had all of the Compustat and CRSP data that is required for the first hypothesis. The availability of analyst data from IBES reduced the sample size from 480 to 478 firms. Furthermore, the availability of data regarding the ownership statistics has was only available for 467 firms. In addition to that, this data was only available for the years 2010-2018 reducing the amount of firm-years. After excluding utility firms and winsorizing all continuous variables at the 1 and 99 percent levels there are 421 firms and 2,968 firm years left. The full sample selection process is tabulated in table 1.

Stage	*Firms*	*Firm-Years*
SP500 from 2010-2019	500	4.577
Availability of Compustat and CRSP data	480	4.479
Availability of analyst data from IBES	478	4.364
Availability of ownership data from Thomson Reuters (2019 data not available)	467	3.504
Data left after excluding utility firms	439	3.284
Data left after winsorizing all continuous variables	421	2.968
Final Sample	**421**	**2.968**

Table 1: Sample selection proces Hypothesis 1

Some of the control variables were heavily skewed or had large outliers. Particularly HHI SIZE and MTB were heavily skewed and INST has large outliers. In order to address these issues all continuous variables were winsorized at the 1 and 99 percent levels. However, SIZE and HHI remained heavily skewed. Therefore, the natural log of these variables was used. The firm years range from 2010-2018 and are relatively equally distributed across the individual years. The amount of observations per firm year can be found in table 2. The number of firms is not equal each year; this is a result of missing datapoints of control variables for certain firms in certain years.

YEAR	DISCLOSE=0	DISCLOSE=1	TOTAL
2010	279	51	330
2011	266	53	319
2012	270	53	323
2013	274	51	325
2014	274	50	324
2015	265	58	323
2016	262	56	318
2017	265	54	319
2018	314	73	387

Table 2: Datapoints per year

Histograms and normal quantile plots of all continuous variables are included within the appendix. The descriptive statistics of the data used to test the first hypothesis is summarized in table 3.

VARIABLE	OBS	MEAN	STD.DEV.	MIN	MAX
DISCLOSE	2,968	0.168	0.374	0	1
HHI	2,968	6.87	0.699	4.908	8.662
SIZE	2,968	9.695	1.358	6.724	13.736
ISSUE	2,968	0.565	0.496	0	1
INST	2,968	0.802	0.146	0.024	1.095
MTB	2,968	4.097	6.554	-69.278	52.781
ANALYSTS	2,968	16.293	7.348	2.565	38.083

Table 3: Descriptive statistics Hypothesis 1

To test the first hypothesis, the probit regression in equation 1 is carried out. Before this is done it is important to examine potential issues with the data before running the regression. Common causes of model misspecification are high or near perfect correlation between explanatory variables, non-normal distribution of variables and heteroscedasticity. In table 4 the correlations between all variables are shown. Below the correlations the significance levels are presented, these significance levels represent the likelihood that the coefficient is significant from zero. There appear to be no variables that are strongly correlated enough to potentially present issues when estimating the model. Other observations from the correlation matrix are that HHI is negatively correlated with disclosure within the sample. It should be noted that the correlation is less significant than some of the other variables in the model. The strongest correlations amongst the control variables are that larger firms appear to be less owned by institutional investors and have a larger analyst following.

	DISCLOSE	HHI	SIZE	ISSUE	INST	MTB	ANALYSTS
DISCLOSE	1.000						
HHI	-0.030 (0.101)	1.000					
SIZE	-0.074 (0.000)	-0.103 (0.000)	1.000				
ISSUE	0.033 (0.077)	0.035 (0.056)	-0.045 (0.014)	1.000			
INST	0.033 (0.068)	-0.072 (0.000)	-0.338 (0.000)	0.056 (0.002)	1.000		
MTB	0.072 (0.001)	0.074 (0.000)	-0.130 (0.000)	0.006 (0.763)	-0.019 (0.309)	1.000	
ANALYSTS	0.013 (0.464)	0.128 (0.000)	0.177 (0.000)	0.019 (0.309)	-0.130 (0.000)	0.114 (0.000)	1.000

Table 4: Correlation matrix hypothesis 1

The probit model is estimated in table 5. The year effects have an insignificant impact and are found in table 11 within the appendix. Estimating the model with robust standard errors does not affect the results either. The model with robust standard errors can also be found in table 12 within the appendix. The most important result is that firms with a larger HHI are less likely to disclose the size of their customer base in their 10-k filings. Thus, firms that operate in a more concentrated industry are less likely to disclose the size of their customer base. This result

indicates that competition does affect a firm's disclosure decisions and thus confirms the first hypothesis. If a firm operates in a more competitive industry, the firm is more likely to disclose the size of their customer base. The -0.102 can be interpreted as follows. As the log of a firms HHI increases by 1 the z-score of the probability of disclosure decreases by 0.102. In a more practical sense if a firms HHI index increases from 1100 to 3000 and all other variables remain at their means the probability of disclosure drops from approximately 0.161 to 0.138. A presentation on the estimated probability of disclosure and marginal effects at different HHI levels is given in table 6. HHI does have a significant impact on the disclosure decision. However, the impact is quite modest. As can be seen in table 6, a change in HHI from 148 to 8103 only decreases the probability of disclosure by 8.3 percentage points. For reference, 148 is a very low market concentration while 8103 is quite close to being a monopoly (with 10,000 being a pure monopoly).

Disclose	**Coef.**	**St. Err.**	**t-value**	**p-value**	**95% Confidence Interval**	
Constant	0.218	0.448	0.49	0.626	-0.660	1.096
HHI	-0.102**	0.041	-2.48	0.013	-0.182	-0.021
SIZE	-0.084***	0.023	-3.58	0.000	-0.129	-0.038
ISSUE	0.087	0.057	1.53	0.127	-0.025	0.198
INST	0.094	0.208	0.46	0.649	-0.312	0.501
MTB	0.014***	0.004	3.08	0.002	0.005	0.022
ANALYST	0.006	0.004	1.43	0.152	-0.002	0.013

*** p<0,01, ** p<0,05, * p<0,1

Pseudo r-squared	0.015	**Number of observations**	2.968
Chi-square	40.983	**Prob > chi2**	0.000

Table 5: Hypothesis 1 probit regression model

HHI	LN(HHI)	PROBABILITY	DY/DX
148	5	0.216	-0.030
403	6	0.187	-0.027
1096	7	0.161	-0.025
2981	8	0.138	-0.022
8103	9	0.116	-0.020

Table 6: Effect of HHI on probability of disclosure

This relation between disclosure and competition is consistent with (Li, 2010) who also found that the level of competition affects disclosure. Of the control variables SIZE and MTB are significant. Starting with the latter, the positive relation between MTB and customer base disclosure is in line with prior disclosure literature. For example, (Verrecchia, 1990) found that firms with a higher level of information asymmetry and a higher MTB ratio disclose more. The sing of SIZE is negative. This is in contrast with most disclosure literature that finds a company's size to be positively correlated with the amount of disclosure. Larger companies supposedly disclose more, while in this model this appears not to be the case. A potential explanation for this could be that it is harder for larger firms to estimate an accurate figure that represents the size of their customer base. Larger firms are more likely to operate in many different countries and sell their products through many different resellers. These could be factors that make it harder for firms to give an accurate figure that represents the size of their customer base. Some of these larger companies disclose the number of stores that they operate or the number of stores in which their products are sold. These kinds of disclosures are not included in the model because they do not directly inform about the size of a firms customer base.

In order to test the value relevance of the disclosures only the firm-years that do disclose are of use. After excluding utility firms there are 778 firm-year observations in which firms disclose the size of their customer base. Limited availability of MVPS, EPS and BVPS further reduces the sample to 759 firm-years. In order to calculate the CUSTGROW variable two consecutive years of disclosure are needed. This requirement further reduces the amount of firm years to 652 The sample before winsorizing any variables consists of 97 firms and 652 firm-years

ranging from 2011 to 2019. The sample selection process for the second hypothesis is tabulated in table 7.

Stage	*Firms*	*Firm-Years*
SP500 from 2010-2019 that do disclose customer base figures	104	778
Availability of Compustat and CRSP data	103	759
Calculation of CUSTGROW (two consecutive years required)	97	652
After winsorizing all continuous variables	97	610
Final sample	**97**	**610**

Table 7: Sample selection process Hypothesis 2

The descriptive statistics of the data used to examine the value relevance of the reported customer figures are found in table 8. This data is after winsorizing all the variables at the 1 and 99 percent levels. Winsorizing these variables is done to limit the effect of outliers on the model estimation. The final sample after winsorizing consists of 97 firms and 610 firm years. The average reported growth of customer base was seven percent year-on-year within the sample. This is reasonable when factoring in the growing population of the United States and the fact that the sample period of 2011-2019 was generally a period of economic growth. Thus, this average increase does not necessarily indicate opportunistic behaviour of managers with the freedom they have in disclosing customer base figures.

VARIABLE	**OBS**	**MEAN**	**STD.DEV.**	**MIN**	**MAX**
MVPS	610	86.311	61.18	10.24	357.41
BVPS	610	21.624	18.093	-14.436	103.493
EPS	610	3.517	2.793	-2.58	16.94
CUSTGROW	610	.076	.126	-.193	1.235

Table 8: Descriptive statistics hypothesis 2

Table 9 includes all of the correlations between all of the variables used. Below the correlations the significance levels are presented, these significance levels represent the likelihood that the coefficient is significant from zero. The accounting figures (BVPS and EPS) have the strongest correlation with market value. However, customer growth is also positively correlated with market value. All of these correlations are significant at the five percent level.

	MVPS	BVPS	EPS	CUSTGROW
MVPS	1.000			
BVPS	0.293 (0.000)	1.000		
EPS	0.568 (0.000)	0.460 (0.000)	1.000	
CUSTGROW	0.097 (0.017)	-0.092 (0.024)	-0.077 (0.058)	1.000

Table 9: Correlation matrix hypothesis 2

As discussed in the research design section, the model is effectively run four times. Once without the CUSTGROW variable in order to generate a baseline r-squared to compare the other models with. The second model is the main model that includes the CUSTGROW variable. The third model only includes data from firms that reported active users. Finally, model four only includes data from firms that disclose a number of traditional customers. Model three and model four are created to examine whether active user disclosure is more value relevant compared to traditional customer base disclosure. As discussed in the literature review and hypothesis development section, the hypothesis is that active users are more value relevant compared to traditional customer base disclosure. Active users are hypothesized to be more value relevant because it can be used to value intangible assets which are typically considered harder to value (McKinsey & Company, 2016). New software services can be difficult to value and one piece of information to assess the value of a of new software services can be the number of active users. Secondly, having a large number of active users is usually the basis of the key revenue streams for the firms that report it. The third reason why active users are hypothesized to be more value relevant is because the term more reliably represent what it claims to represent. MAU counts users that are active at least once a month. However, someone who bought products five years ago could theoretically still be considered a customer. The results of the four models can be found in table 10.

MVPS	Model 1 (Basic)	Model 2 (All)	Model 3 (Only Users)	Model 4 (No Users)
Constant	41.020 *** (10.44)	34.441*** (8.99)	10.932 (1.12)	36.802*** (8.79)
BVPS	0.137 (1.08)	0.172 (1.37)	0.252 (0.39)	0.160 (1.22)
EPS	12.034*** (14.62)	12.173*** (14.99)	16.699*** (5.95)	11.812*** (13.65)
CUSTGROW		69.948*** (4.36)	124.225** (2.41)	68.041*** (4.01)
Adjusted R2	0.3218	0.3414	0.5426	0.3230
Observations	610	610	70	540
**** p<0.01, ** p<0.05, * p<0.1*				

Table 10: OLS models hypothesis 2

Table 10 includes several interesting results. First, the adjusted r-squared does increase from model 1 to model 2. It increases slightly from 0.3218 in model 1 to 0.3414 in model 2. This indicates that the growth in a firm's reported customer base does have incremental value relevance. This result does confirm the second hypothesis. The change in number of customers reported by a firm provides value relevant information beyond financial statements to explain market value of equity. This is consistent with several other studies that have found customer related statistics to be value relevant (Gupta et al., 2004; Choi et al., 2019). According to model 2 a 10 percent growth in reported customer base would be associated with an increase of 6.95 dollars of the predicted share price. The average share price in the sample is 86.31 dollars. Thus, on average this 6.95 dollar increase would represent an increase of 8.05 percent.

Model 3 yields the highest adjusted r-square of 0.5426. This does confirm the hypothesis that active users have a higher value relevance compared to traditional customers. In addition to this the coefficient of the CUSTGROW variable is also larger compared to both models 2 and 4. According to model 3 a 10 percent growth in active users would be associated with an increase of 12.42 dollars. The average share price of firms within the sample that reported active users is 72.63 dollars. Thus, on average this 12.42 dollar increase would represent an increase in share price of 17.10 percent. However, one should not read into the specific coefficients too much as it is not the intend to show a causal link between the customer growth and share price. Instead, the main objective of the model is to examine the value relevance of the customer growth

disclosure. This is done by comparing the r-squared of the different models. It also should be noted that the amount of observations on which these results are based on is relatively low. The number of firms that disclosed active users within the sample was quite small and this should be kept in mind when examining these results. It is also this smaller sample size that leads to the lower significance of all of the variables. Further research could be carried out with a larger sample that includes more firms which disclose this data.

It should be noted that this value relevance is only based on firms that do already voluntarily disclose the number of customers or user that they have in their 10-k filings. Firms for which this information is less value relevant might be less inclined to disclose this information. For example, firms who generate almost all of their revenue through one customer often do not disclose the number of customers that they serve. Instead, these firms might disclose information about the relationship with this one client.

Finally, one interesting observation is the insignificance of the BVPS coefficient in all four models. Even though this is not of any particular interest to the hypothesis being tested it is an interesting observation. The insignificant coefficient is quite an unexpected result. It is expected that firms with a higher book value of equity also have a higher market value of equity. This is also expected because of the positive correlation found between these two variables. One potential explanation of this result is that due to a low variance in BVPS the constant factor in the model captures a large part of the MVPS that would otherwise be explained by BVPS.

5. Conclusion

In this thesis I examine the link between competition and customer base disclosure. In addition to that, I also examine whether these customer base disclosures are value relevant. In order to examine potential motivations and the value relevance of customer base disclosure the following research question has been formulated.

Does competition influence a firm's decision to disclose the size of their customer base and are these customer base disclosures value relevant?

Based on the findings in the previous section I can conclude that competition does indeed influence a firm's decision to disclose the size of their customer base. The sign for the HHI in the probit regression is negative and significant at the 10 percent level. Thus, firms that operate in more concentrated markets (less competition) are less likely to disclose the number of customers that they have. These results are consistent with prior papers that have also found a relation between disclosure and competition (Li, 2010). This thesis expands on the relatively

small body of work examining competition and disclosure that already exist by examining a link between non-financial disclosure and competition.

Also based on the findings from within the previous section I can conclude that these customer base disclosures do have incremental value relevance beyond what is explained by traditional earnings. Adding a variable for customer growth does increase the explanatory power of the model. This increase is a lot higher when only firms that disclose active users are considered. This indicates that active users disclosures have a higher incremental value relevance than normal customer base disclosures. Customer base disclosures being value relevant is consistent with several other papers which examine different but similar customer related metrics (Gupta et al., 2004; Choi et al., 2019).

These results have some implications. For managers, the implication is that they should be conscious about their decision disclose this information because it is more than just general information about the company. The second implication relates to regulations. There are relatively few regulations regarding these disclosures. This does mean that an opportunistic manager could attempt to inflate the number of customers. For example, by counting customers that have not bought anything from the firm in years. Regulators could potentially look into this in order to ensure that disclosures about a firm's customer base remain accurate.

There are a number of limitations to the research. First, the sample size that is used to examine the value relevance of active users is relatively small. This small sample size is a result of the small number of firms that disclose this figure and time limitations related to the hand collection of data. This small sample size can make it difficult to generalize the findings to a broader population. Potential future research could examine the value relevance of active users specifically. This could allow the research to focus on this issue and test the value relevance of active users with a larger sample.

The second limitation is that only 10-k filings were used to gather data regarding the disclosures. It is possible that a firm might disclose the number of customers that they have outside of these filings. For data collection purposes I only got data from the firm's 10-k filings. However, if firms do disclose this information through other channels it could affect the results within this thesis. As they would be marked as a firm that does not disclose the size of their customer base, while in fact they do. Therefore, the findings only apply to disclosures within a firms 10-k form. Potential future research could examine the effects of customer base disclosure outside the official 10-k filings.

The third and final noteworthy limitation is that the sample consisted of only S&P 500 firms. Due to time limitations it was not possible to hand collect data on a larger sample. However, the S&P 500 generally includes larger companies. This means that smaller companies and companies outside of The United States are not included within the sample. Further research could address this issue by taking a larger sample of potentially the S&P 1500 index. The S&P 1500 index also includes smaller and midsized firms.

References

Ali, A., Klasa, S., & Yeung, E. (2014). Industry concentration and corporate disclosure policy. *Journal of Accounting and Economics*, *58*(2–3), 240–264.

Amir, E., & Lev, B. (1996). Value-relevance of nonfinancial information: The wireless communications industry. *Journal of Accounting and Economics*, *22*(1–3), 3–30.

Baginski, Stephen P, & Hassell, John M. . (1997). Determinants of Management Forecast Precision. *The Accounting Review*, *72*(2), 303–312. Retrieved from https://www.jstor.org/stable/248558

Barth, M. E., Beaver, W. H., & Landsman, W. R. (2001). The relevance of the value relevance literature for financial accounting standard setting: another view. *Journal of Accounting and Economics*, *31*(1–3), 77–104.

Barth, M. E., Landsman, W. R., & Lang, M. H. (2008). International Accounting Standards and Accounting Quality. *Journal of Accounting Research*, *46*(3), 467–498.

Black, D. E., Christensen, T. E., Ciesielski, J. T., & Whipple, B. C. (2018). Non-GAAP reporting: Evidence from academia and current practice. *Journal of Business Finance & Accounting*, *45*(3–4), 259–294.

Burks, J. J., Cuny, C., Gerakos, J., & Granja, J. (2018). Competition and voluntary disclosure: evidence from deregulation in the banking industry. *Review of Accounting Studies*, *23*(4), 1471–1511.

Choi, Y. B., Lee, J., Yoo, S., & Yoo, Y. K. (2019). Value relevance of customer equity beyond financial statements: evidence from mobile telecom industry. *Asia-Pacific Journal of Accounting & Economics*, *26*(3), 281–300.

Christensen, T. E., Drake, M. S., & Thornock, J. R. (2014). Optimistic reporting and pessimistic investing: Do pro forma earnings disclosures attract short sellers?. *Contemporary Accounting Research*, *31*(1), 67-102.

CNBC. (2020, April 23). Zoom shares pop after users grow from 200 million to 300 million in a matter of days. Retrieved from https://www.cnbc.com/2020/04/23/zoom-shares-pop-after-users-grow-from-to-300-million.html

Cormier, D., Demaria, S., & Magnan, M. (2017). Beyond earnings: do EBITDA reporting and governance matter for market participants? *Managerial Finance*, *43*(2), 193–211.

Curtis, A. B., McVay, S. E., & Whipple, B. C. (2013). The Disclosure of Non-GAAP Earnings Information in the Presence of Transitory Gains. *The Accounting Review*, *89*(3), 933–958.

Dhaliwal, D. S., Li, O. Z., Tsang, A., & Yang, Y. G. (2011). Voluntary Nonfinancial Disclosure and the Cost of Equity Capital: The Initiation of Corporate Social Responsibility Reporting. *The Accounting Review*, *86*(1), 59–100.

Ellis, J. A., Fee, E. C., & Thomes, S. E. (2012). Proprietary Costs and the Disclosure of Information About Customers. *Journal of Accounting Research*, *50*(3), 685–727.

Facebook. (2019). *Form 10-K 2019*. Retrieved from SEC EDGAR website http://www.sec.gov/edgar.shtml

Frankel, R., McNichols, M., & Wilson, G. P. (1995). Discretionary Disclosure and External Financing. *The Accounting Review*, *70*(1), 135–150.

Franzen, L., & Radhakrishnan, S. (2009). The value relevance of R&D across profit and loss firms. *Journal of Accounting and Public Policy*, *28*(1), 16–32.

Frederickson, J. R., & Miller, J. S. (2004). The Effects of Pro Forma Earnings Disclosures on Analysts' and Nonprofessional Investors' Equity Valuation Judgments. *The Accounting Review*, *79*(3), 667–686.

Gao, F., Dong, Y., Ni, C., & Fu, R. (2015). Determinants and Economic Consequences of Non-financial Disclosure Quality. *European Accounting Review*, *25*(2), 287–317.

Graham, J. R., Harvey, C. R., & Rajgopal, S. (2005). The economic implications of corporate financial reporting. *Journal of Accounting and Economics*, *40*(1–3), 3–73.

Gupta, S., Lehmann, D. R., & Stuart, J. A. (2004). Valuing Customers. *Journal of Marketing Research*, *41*(1), 7–18.

Hassel, L., Nilsson, H., & Nyquist, S. (2005). The value relevance of environmental performance. *European Accounting Review*, *14*(1), 41–61.

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

Isidro, H., & Marques, A. (2015). The Role of Institutional and Economic Factors in the Strategic Use of Non-GAAP Disclosures to Beat Earnings Benchmarks. *European Accounting Review*, *24*(1), 95–128.

Ittner, C. D., & Larcker, D. F. (1998). Are Nonfinancial Measures Leading Indicators of Financial Performance? An Analysis of Customer Satisfaction. *Journal of Accounting Research*, *36*, 1–35.

Lambert, R., Leuz, C., & Verreccia, R. E. (2007). Accounting Information, Disclosure, and the Cost of Capital. *Journal of Accounting Research*, *45*(2), 385–420.

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

Li, X. (2010). The impacts of product market competition on the quantity and quality of voluntary disclosures. *Review of Accounting Studies*, *15*(3), 663–711.

Luo, L., & Tang, Q. (2014). Does voluntary carbon disclosure reflect underlying carbon performance? *Journal of Contemporary Accounting & Economics*, *10*(3), 191–205.

McKinsey & Company. (2016). *Valuing high-tech companies*. Retrieved from https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/valuing-high-tech-companies

Orens, R., Aerts, W., & Lybaert, N. (2013). Customer value disclosure and cost of equity capital. *Review of Accounting and Finance*, *12*(2), 130–147.

Rajgopal, S., Venkatachalam, M., & Kotha, S. (2003). The Value Relevance of Network Advantages: The Case of E-Commerce Firms. *Journal of Accounting Research*, *41*(1), 135–162.

Rezaee, Z., & Tuo, L. (2017). Voluntary disclosure of non-financial information and its association with sustainability performance. *Advances in Accounting*, *39*, 47–59.

Vafaei, A., Taylor, D., & Ahmed, K. (2011). The value relevance of intellectual capital disclosures. *Journal of Intellectual Capital*, *12*(3), 407–429.

Venter, E. R., Emanuel, D., & Cahan, S. F. (2014). The Value Relevance of Mandatory Non-GAAP Earnings. *Abacus*, *50*(1), 1–24.

Verrecchia, R. E. (1990). Information quality and discretionary disclosure. *Journal of Accounting and Economics*, *12*(4), 365–380.

Walmart. (2020). About Us. Retrieved 24 July 2020, from https://corporate.walmart.com

Wieland, M. M., Dawkins, M. C., & Dugan, M. T. (2013). The Differential Value Relevance of S&P's Core Earnings Versus GAAP Earnings: The Role of Stock Option Expense. *Journal of Business Finance & Accounting*, *40*(1–2), 55–81.

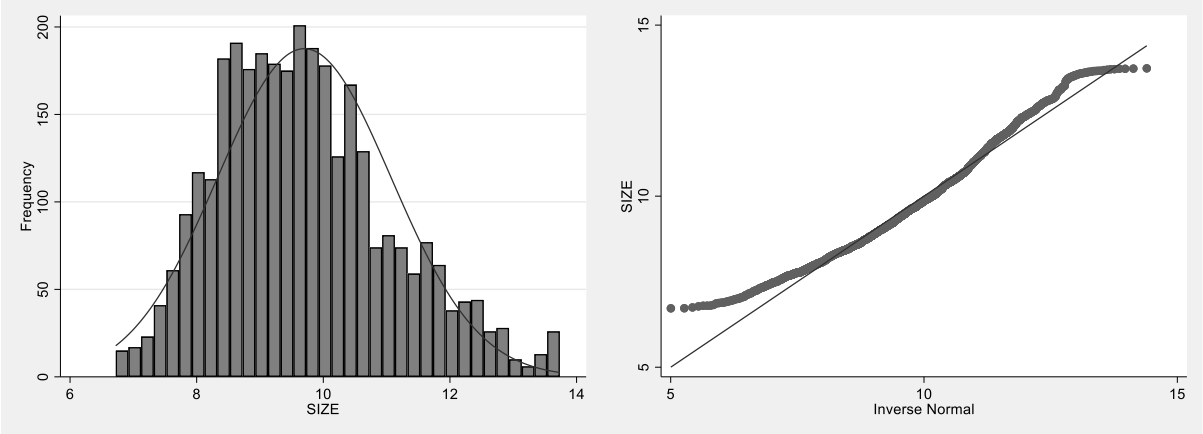
Appendix

Variable definitions

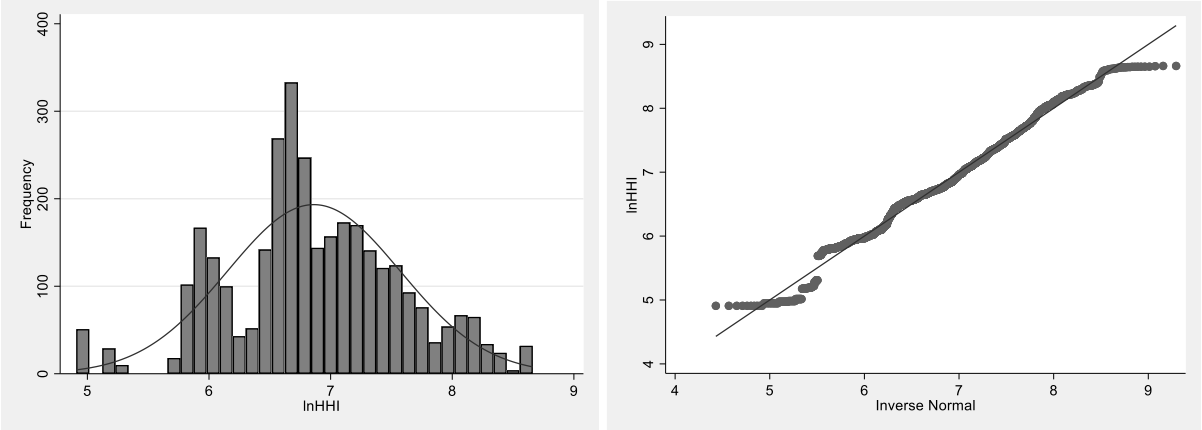
VARIABLE	DESCRIPTION	SOURCE
ANALYSTS	The number of analysts that are following a firm.	IBES
BVPS	A firm's book value of equity per share	Compustat
CUSTGROW	The year on year percentage growth in the reported customer base, this is calculated based on the hand collected data.	Hand-collected SEC-EDGAR
DISCLOSE	A dummy variable which is 1 when a firm does disclose the size of their customer base in their 10-k filings.	Hand-collected SEC-EDGAR
EPS	A firm's earnings per share	Compustat
HHI	The Herfindahl–Hirschman Index, measured as the sum of squared market shares of all firms in an industry. Calculated using all available US Compustat data.	Compustat
INST	The percentage of a company's equity that is owned by institutional investors.	Thomson Reuters
ISSUE	A dummy variable which is 1 if a company issues a significant amount equity or debt in the current year or in the next year. Calculated based on share and long term debt data from Compustat.	Compustat
MTB	A firm's market-to-book ratio. Calculated by dividing MVPS by BVPS.	CRSP/Compustat
MVPS	A firm's share price	CRSP
SIZE	The size of a company, measured as the natural log of their total assets.	Compustat

Histogram and normal quantile plots

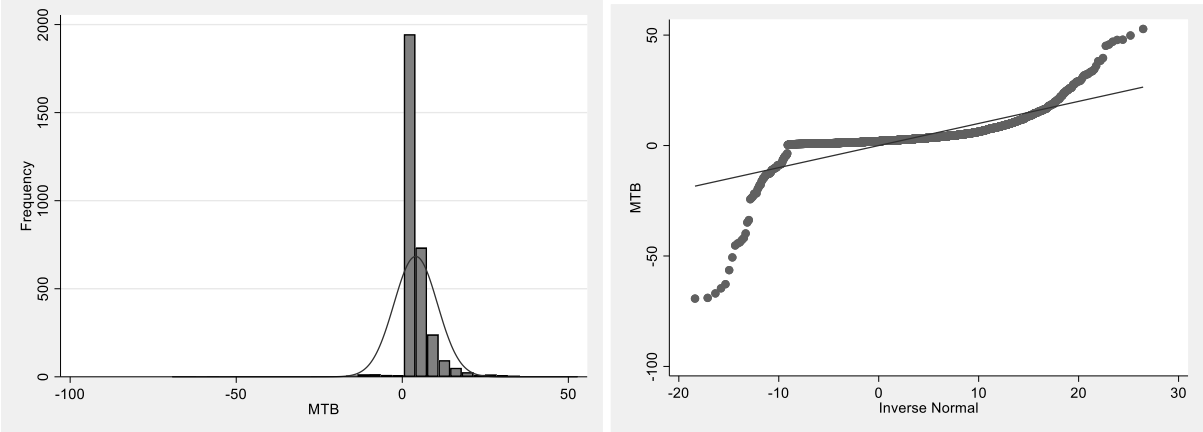
Size – the log of total assets



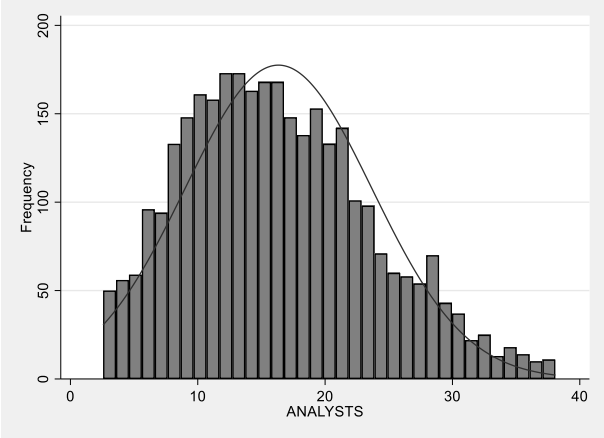
HHI – the log of the HHI index

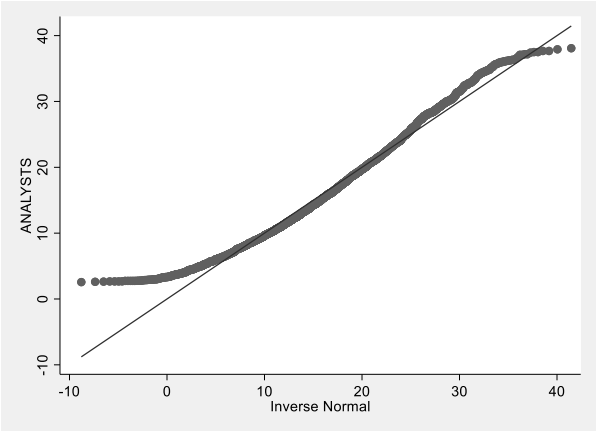


MTB – the market-to-book ratio trimmed at 1% and 99%

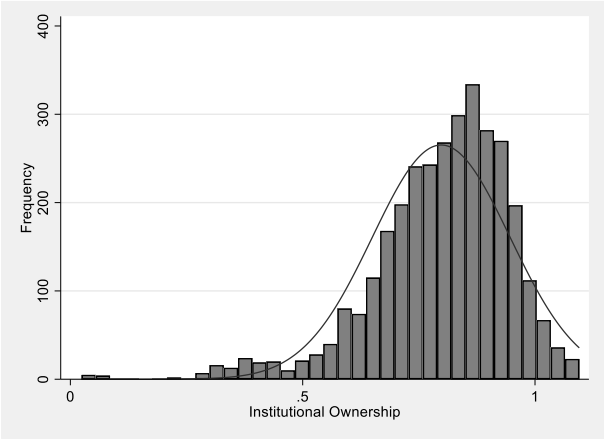


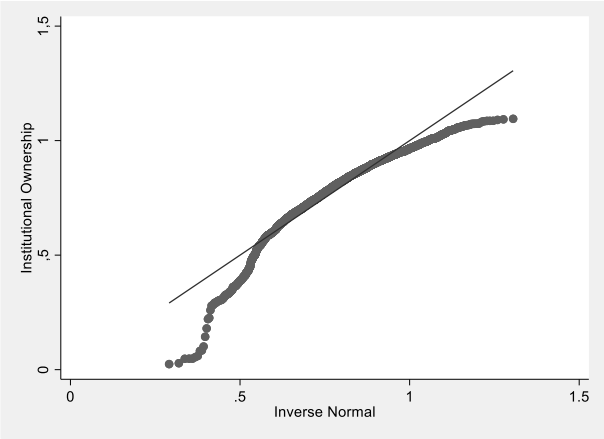
ANALYST – The number of analysts that follow a firm trimmed at 1% and 99%



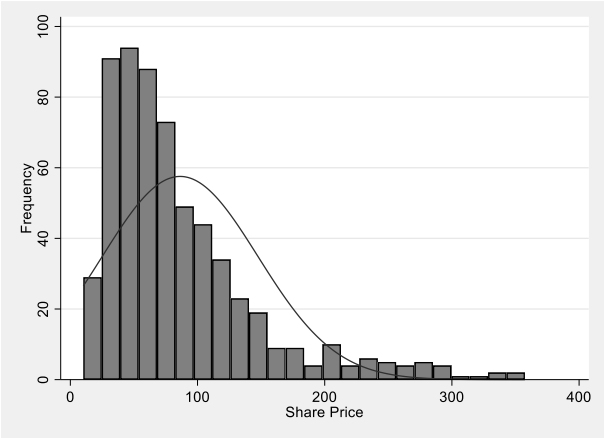


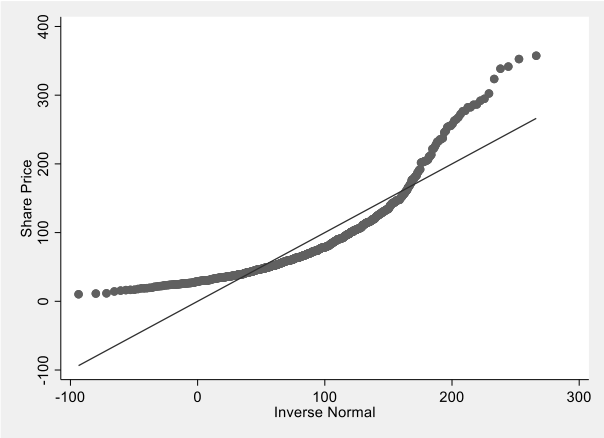
INST – The percentage of institutional ownership trimmed at 1% and 99%



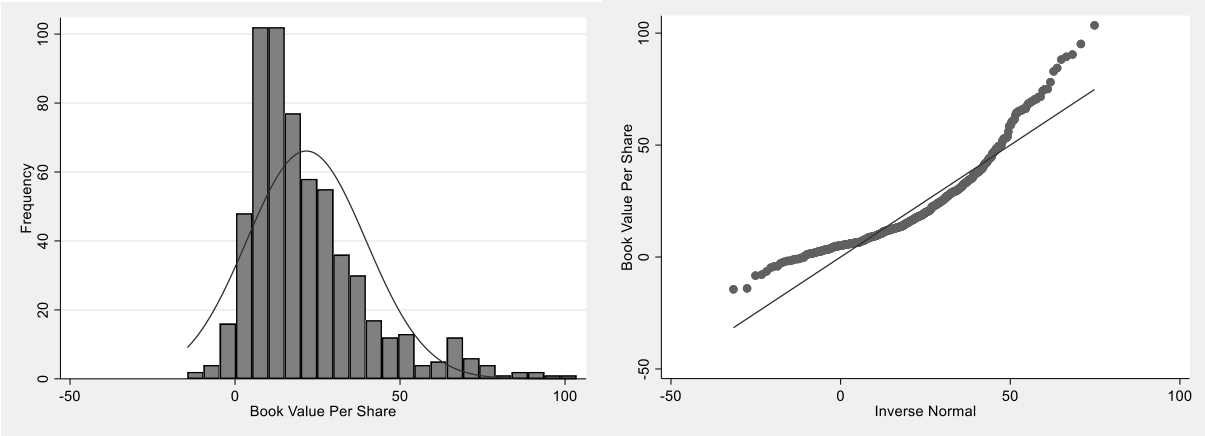


MVPS – Market value per share

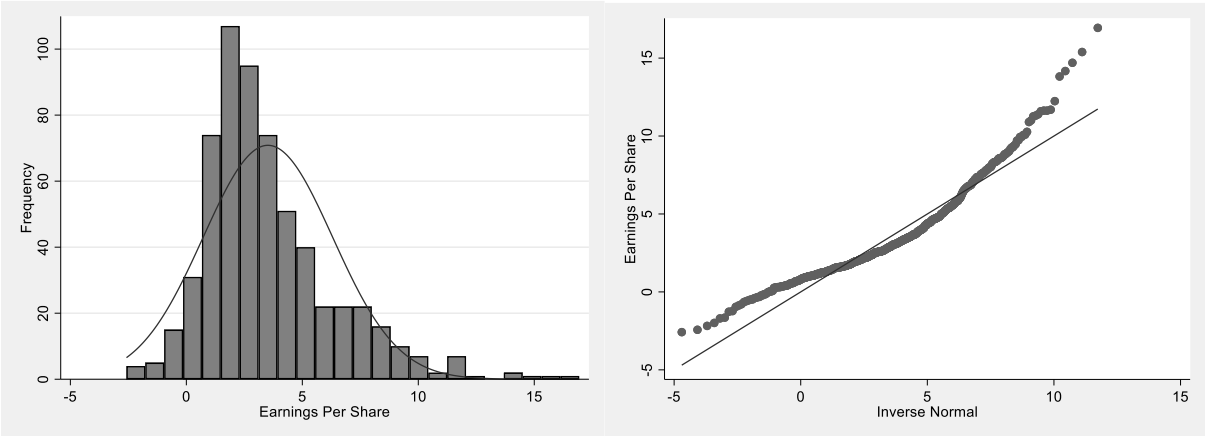




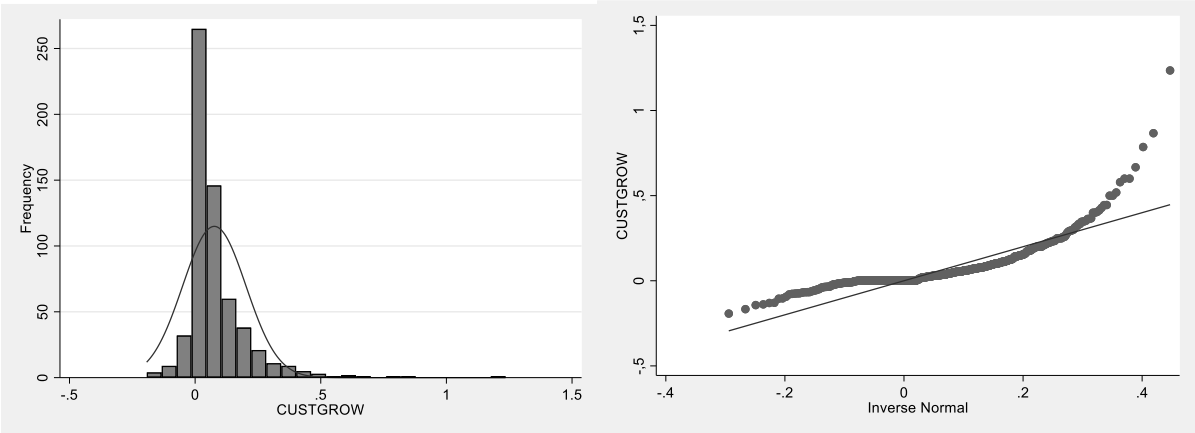
BVPS – Book value of equity per share trimmed at 1% and 99%



EPS – Earnings per share trimmed at 1% and 99%



CUSTGROW – Year-on-year growth in reported customer base trimmed at 1% and 99%



Year fixed effects and model with robust standard errors

Year fixed effects from the probit model in table 5

DISCLOSE	COEF.	ROBUST ST.ERR.	T-VALUE	P-VALUE	95% CONFIDENCE INTERVAL	
2010.YEAR	0.000	.	.	.	.	.
2011.YEAR	0.033	0.120	0.27	0.785	-0.203	0.268
2012.YEAR	0.010	0.120	0.08	0.935	-0.226	0.246
2013.YEAR	-0.022	0.121	-0.18	0.857	-0.259	0.215
2014.YEAR	-0.050	0.122	-0.41	0.681	-0.290	0.190
2015.YEAR	0.068	0.120	0.56	0.572	-0.168	0.304
2016.YEAR	0.053	0.121	0.44	0.661	-0.184	0.290
2017.YEAR	0.036	0.121	0.29	0.768	-0.201	0.273
2018.YEAR	0.126	0.114	1.10	0.272	-0.099	0.350

Table 11: Fixed effects of main probit model

DISCLOSE	COEF.	ROBUST ST.ERR.	T-VALUE	P-VALUE	95% CONFIDENCE INTERVAL	
HHI	-0.102**	0.042	-2.45	0.014	-0.183	-0.020
SIZE	-0.084***	0.023	-3.64	0.000	-0.129	-0.039
ISSUE	0.087	0.056	1.55	0.121	-0.023	0.196
INST	0.094	0.209	0.45	0.651	-0.315	0.504
MTB	0.014***	0.005	2.63	0.008	0.003	0.024
ANALYST	0.006	0.004	1.42	0.156	-0.002	0.013
2010.YEAR	0.000	.	.	.	.	.
2011.YEAR	0.033	0.120	0.27	0.785	-0.202	0.268
2012.YEAR	0.010	0.120	0.08	0.935	-0.226	0.246
2013.YEAR	-0.022	0.120	-0.18	0.856	-0.258	0.214
2014.YEAR	-0.050	0.121	-0.41	0.678	-0.287	0.187
2015.YEAR	0.068	0.119	0.57	0.569	-0.166	0.302
2016.YEAR	0.053	0.120	0.44	0.658	-0.182	0.288
2017.YEAR	0.036	0.120	0.30	0.765	-0.199	0.270
2018.YEAR	0.126	0.113	1.11	0.266	-0.096	0.347
PSEUDO R-SQUARED		0.015	**Number of observations**		2.968	
CHI-SQUARE		38.614	**Prob > chi2**		0,000	

Table 12: Main probit model with robust standard errors