ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics

Master Thesis Accounting & Auditing

The effect of SOX 404(b) on operational efficiency and financial performance

Name student: Matthias Pach

Student ID number: 508565

Supervisor: Jaeyoon Yu

Second assessor: Ying Gan

Date final version: 8th of July 2022

Abstract

This research investigates the consequences of section 404 of the Sarbanes-Oxley Act. It extends prior research by looking at whether compliance with SOX 404(b) increases operational efficiency and financial performance. This through improvements in the internal control environment as a result of the audit and the fact that the company knows its control environment is under audit.

The results indicate that the audit under SOX 404(b) does not provide these benefits. In contrast, they suggest that SOX 404(b) negatively impacts financial performance. Combined with previous literature, this indicates that the implementation of SOX 404(b) was not economically justifiable as the costs seem to heavily outweigh the benefits.

Keywords: SOX 404(b), Sarbanes-Oxley Act, Internal Controls, Operational Improvements

Special thanks to Professor Jaeyoon Yu for his guidance during the writing process and to Professor Roy Schmardebeck (University of Tennessee) for his help in obtaining the public float data

The views stated in this thesis are those of the author and not necessarily those of the supervisor, second assessor, Erasmus School of Economics or Erasmus University Rotterdam.

1. Introduction

This research examines whether the introduction of the Sarbanes-Oxley Act Section 404(b) (SOX 404(b)) has led to the benefit of operational improvements for compliers. SOX was implemented to improve the quality of financial reporting, increase the investor confidence and to reduce the possibilities of corporate fraud (Lui, 2020). To comply with Section 404(a) companies must write a report evaluating their Internal Controls (IC), incentivizing the companies to maintain them. All companies filing under the SEC requirements must comply with Section 404(a), but only accelerated filers need to comply with part (b) as well. When the Sarbanes-Oxley Act was first implemented, accelerated filter were filers with a public float of more than 75 million USD. The public float being the aggregate value of the portion of shares that is held by non-affiliates of the company (SEC), i.e., the value of the shares that are publicly tradable by investors, excluding restricted shares that are for example bought back by the company. Non-accelerated filter, i.e., companies with a public float lower than 75 million, where exempted from complying with Section 404(b). Effective from 2020, the exemption is extended to all companies with a public float of less than 100 million, instead of the previous 75 million. Of course, companies below the threshold are allowed to comply voluntarily if they want to do so.

Compliers with SOX 404(b) need to have their report written under Section 404(a) audited by an external auditor. This entails that the auditor must test and report on the effectiveness of the internal controls over the financial reporting of a company. SOX 404(b) brings extra costs for the compliers, in the form of e.g., higher audit fees and higher CFO compensation (McCallen et al., 2020). To justify this there should be benefits.

According to supporters of the regulation, SOX 404(b) could incentivize improvements within the control environment of a complying company. This would be beneficial for the stakeholder as it would lead to better internal controls, more informative internal control reports and a higher quality of financial reporting (Ashbaugh-Skaife (2009); McCallen, 2020). Not all these benefits seem to concur in practice however, as current academic research fails to find that SOX 404(b) positively affects the financial reporting quality (McCallen et al., 2020), removing that benefit from the list. But not all possible benefits of SOX 404(b) are investigated upon. Meaning that even though the Sarbanes-Oxley Act was implemented almost two decades ago, there is an ongoing debate about the benefits of the Act against its downsides. This research contributes to that debate by investigating upon some of the other potential benefits of SOX 404(b), trying to fill the gap in the academic literature that currently exists in this area and helping regulators evaluate on their decision making and helping them with their future rule proposals. This research investigates if SOX 404(b) has the benefit of leading to better internal controls. These improved internal controls would benefit the stakeholder as the quality of internal controls has an economically significant positive effect on firm operations (Feng, McVay, Skaife, 2015) and internal controls with less deficiencies are associated with improved operational efficiency (Cheng, Goh, Kim, 2018). Following this, the operational performance will lead to a better financial performance in the form of greater revenue and sales (Bendickson & Chandler, 2019).

Thus, if complying with the additional audit requirements of SOX 404(b) improves the complier's internal controls, this will be visible in the form of operational efficiency improvements and subsequently in the form of better financial performance. This would show that complying with SOX 404(b) has some direct benefits for the company, in contrast to the indirect benefits (e.g., lower cost of equity through more investor confidence because of higher quality financial reporting) that don't even seem to be there. If these advantages of complying are in place, this will help to justify the increased costs associated with a company's compliance with SOX 404(b). However, if there are no positive effects on operational and financial performance, the results of this research would, combined with the findings of McCallen et al (2020), suggest that the costs of the implementation of SOX 404(b) outweigh the benefits.

This paper investigates the question if the benefits of SOX 404(b) are greater than the costs by investigating if there exists an association between compliance with the regulation and operational improvements as well as financial improvements. Based on logic extending on prior research, the hypothesis is that compliance is positively associated with increased operational efficiency and financial improvements. According to Coates and Srinivasan (2014) research into the effects of SOX is hard because of the lack of a control group. This makes it hard rule out any possible endogeneity concerns. To fix this problem, in this research a Regression Discontinuity Design (RDD) is used that uses the regulatory cut-off point for accelerated and non-accelerated filers as an instrument variable to predict SOX 404(b) compliance. Important to note here is that this works as a valid test given that there are no other firm fundamentals related to this cut-off point, meaning that the effect of SOX 404(b) will be well isolated. Following McCallen (2020), a bandwidth around the cut-off point of 100 million is used to ensure that the companies on either side of the cut-off point have mostly similar underlying characteristics, except for their SOX 404(b) compliance. The RDD is used in combination with different dependent variables to proxy for the operational efficiency and the financial performance. For the operational efficiency an efficiency frontier is made using Data Envelopment Analysis (DEA) and for the financial performance different financial performance factors are used separately. The sample consists of US companies that filed with the SEC between 2009 and 2019 that reported a public float between 25 and 125 million USD in their annual reports.

Evidence indicates that SOX 404(b) did not increase either the operational efficiency nor the financial performance. In fact, some of the results even suggest that it has significant negative effects on a firm's financial performance.

The paper is relevant as it possibly helps regulators reflect on their decision making and can guide them in their future decisions as well as helping end the on-going discussion about SOX 404(b) between compliers and regulators.

It is also relevant to the academic literature as it extends prior research on relations between internal controls and firm performance (Chen et al., 2018; Stoel and Muhanna, 2011) by specifically looking at how these internal controls are affected by SOX 404(b) and by incorporating the effect of smaller improvements instead of only the remediation of severe internal control weaknesses. Furthermore, where the paper of McCallen (2020) looks into some potential costs and benefits of SOX 404(b) this paper extends on that by investigating upon a potential benefit (the effect of SOX404 on operational performance) that is excluded in that paper and that is specifically stated as a shortcoming that requires further research there.

A shortcoming of this paper is that it does not take companies of all sizes into account. The found results might have different implications for larger companies that fall outside of the scope of the sample. Furthermore, due to data availability, the timeframe starts multiple years after the implementation and ends in 2019, three years before the writing of this paper. Furthermore, as shown in section 4.1., the data is skewed towards certain years in terms of number of observations. Some years have a very limited number of observations when compared to others, meaning that the results might be biased to certain years, and could be different if more data for the other years was available. A last shortcoming is that due to the complexity of obtaining company public float data, around 15% of the initial sample is dropped. Although there is not clear reason suggesting this, it could potentially be the case that this 15% has different characteristics from the rest of the sample, biasing the results.

The further structure of the paper is as follows. First, the theoretical framework will follow, together with the development of the hypotheses. After that the used sample methodology is explained following the results, which are strengthened with robustness checks. Lastly, the conclusion and final remarks are given.

2. Theoretical background and Hypothesis Development

2.1. Background

When the Sarbanes-Oxley Act was announced in 2002 to improve the financial reporting quality Section 404 might have been the section that sparked the most debate (Honinsberg and Rajgopal, 2019; Barth et al. 2019). This because of the high amount of additional work required both for the filer as well as the external auditor and the consequential costs that this entails. SOX 404 requires companies to establish adequate internal controls and procedures over financial reporting and to ensure SOX compliance these companies must maintain these controls, document about them and test them to ensure their proper performance. They must include a section in their annual report containing a report about these internal controls. Under part b of the section, SOX 404(b), the complying company needs to hire an independent external auditor that must assess the report in order to give certification that all the above actions required under SOX 404 were carried out accordingly.

From the moment SOX was announced, adversaries of the rule have advocated that SOX 404 brings too much costs (Powel, 2005; Romano, 2005), especially for smaller companies who according to Iliev (2010) basically had to create control systems from scratch, causing significant costs to be made in the form of creating and staffing completely new departments in order to create and maintain these new internal controls. However, proponents of the rule argue that this was necessary to increase the quality of financial reporting to a sufficiently high standard. These arguments suggested that especially after a few years after the initial costs were made, the benefits would outweigh the costs (Alexander, 2013). Whether the net benefit was positive or negative remains unclear (Chhaochharia and Grinstein, 2007).

In 2018, the SEC came up with amendments that brought changes to the definitions of (large) accelerated and non-accelerated filers. For this research, the distinction between large and normal accelerated filers is not relevant as the difference between them only has implications for the filing deadline of their annual report. Therefore, in this research, large accelerated filers and accelerated filers are both referred to with further mentions of accelerated filers. The change that the SEC proposed was to increase the number of companies that are non-accelerated filter by increasing the threshold of public float for being an accelerated filer from 75 million USD to 250 million USD. The increasement of the threshold, according to the SEC itself, was made in order to decrease the financial burden that SOX places upon compliers in order because the additional fees required for compliance with SOX404(b) often where significant for the relatively small companies that are affected by the amendments (i.e., companies with a public float between 75 and 250 million USD). By implementing these changes effective from 2020 they recognize the initial concerns that opposers of the rule brought up when it was first implemented. To further illustrate these costs, the additional costs for a company to comply with section 404(a) where a bit over 90 thousand USD on average as estimated by the SEC, but according to Coates and Srinivisan (2014), these costs were significantly higher in reality. The additional audit costs for the attestation of the report (i.e., part b of Section 404) even lead to audit costs more than doubling for compliers in the first year that SOX was effective (Iliev, 2010). Adversaries of the rule therefore argue that only using Section a of SOX 404 is a viable alternative (DeFond and Lennox, 2017). This recent change has brought new fuel to the fire in the debate between the costs and benefits for compliers and shows that, even though it was originally announced two decades ago, SOX Section 404(b) remains relevant to this day.

While after the implementation of SOX, proponents argued that these additional costs would decline in the years following the initial year of compliance, McCallen et al. (2020) shows that the costs remain significantly high. In the research, it is shown that the additional audit cost associated with compliance with SOX 404(b) remain significant and have not materially declined since the implementation of the regulation. This is despite the SEC actively trying to bring down the costs of compliance by for example by implementing Auditing Standard 5 in 2007, trying to tailor the requirements for the internal controls report more to the complexity of the company (Doogar at al. 2010).

Furthermore, the research looks at potential benefits associated with SOX 404 compliance. As current academic research undividedly seems to agree that there are significant costs associated with SOX 404(b) (Clay and Kim, 2017), in order for the regulation to be economically justifiable the regulation also should entail significant economic benefits. These benefits could manifest themselves in multiple ways, one obvious benefit being an improved financial reporting quality, seeing that SOX 404 clearly focusses on increasing the internal controls over this financial reporting. Surprisingly, this does not seem to be the case. According to the research of McCallen, no significant financial reporting quality improvements are found for companies complying with SOX 404(b). This conclusion is not in favor of SOX 404(b), but of course there are more aspects that need to be taken into account to form a well thought opinion on the regulation.

Continuing, McCallen also investigates if SOX 404(b) improves the internal controls over financial reporting. The theory behind this being that if the audit of the ICFR report would find additional internal control weaknesses, this could help the company improve their ICFR for the future. Although they find evidence suggesting that the internal controls became more effective in the years in their sample period, they are not able to prove that it was audits under SOX 404(b) that managed to materially improve these internal controls over financial reporting. This because they saw a similar trend for companies not complying with SOX 404(b). Furthermore, it is important for this research to note that McCallen only tests this by looking at if the ICFR audit helps to remediate internal control deficiencies (ICDs). Internal control deficiencies are major weaknesses and/or problems in the internal controls. The paper does not take into account possible smaller improvements to internal controls, or improvements to internal controls that are already sufficient, but could be made even better.

Adding on this, ICDs are only about problems that could lead to mistakes in the financial statements, not about operational processes that might be inefficient, but that do not have the potential to directly lead to material misstatements. Potential improvements to internal controls

that improve operational processes but that form no harm for misstatements are thus not captured.
These are ways that the SOX 404(b) audit could improve the internal controls that do not fall under the scope of ICD and that thus are not investigated in the paper of McCallen.

The final subject that the paper of McCallen investigates that is discussed in this research is if the audit under SOX 404(b) could help to improve the informativeness of the management report on ICFR. McCallen tests this by looking if ICDs for compliers are more predictive of financial reporting failures than non-complying companies. They are again unable to find that this effect is in place. Also, there only is looked at ICDs, smaller improvements and improvements to areas that already where sufficient are not taken into account. This research extends on that of McCallen by capturing, among other things, the potential effect of these other potential improvements that fall outside of the scope of ICDs, but that potentially could help the company improve. This plays to a specific shortcoming mentioned in the paper, namely that the paper only investigates possible financial reporting benefits and not potential operational improvements.

For this paper to investigate if the compliance with SOX 404(b) could have other potential benefits besides the ones investigated in the paper of McCallen something that is investigated is if minor improvements to internal controls also could have benefits for the respective company.

A paper looking at the effects of internal control quality beyond financial reporting is that of Cheng, Goh and Kim (2018). This paper investigates whether better internal controls have the potential to increase the internal efficiency. They find evidence suggesting that this is indeed the case. The paper however only captures the improvements of internal controls as a dummy variable looking if the company had a material weakness in its internal controls. Therefore, it also does not capture potential smaller improvements that could be made. It however does suggest that also these smaller improvements could positively impact the operational efficiency. The paper of Cheng uses a sample of only accelerated filers, so it does not investigate upon the differences that might be there between accelerated filers and filers that are exempted from complying with SOX 404(b). This research therefor extends on this paper as well as that of McCallen by investigating whether there are other improvements made to either a company's internal controls or its operational processes as a consequence of the audit under SOX 404(b), capturing these improvements by looking at their consequential effect on operational efficiency.

A paper looking if internal controls have a direct positive effect on operational performance is that of Feng, Li, McVay and Skaife (2015). Their paper hypothesizes that in the case of lower quality internal controls, the decision making of the management will be based on less reliable information. Consequentially, this would negatively affect their operational performance. They find evidence that maintaining an effective internal controls system can indeed positively impact the operational performance of a company. Feng also finds that these operational improvements are associated with significant economical improvements for the firm. The paper does however does not take into account either SOX or smaller improvements to internal controls.

2.2. Hypotheses development

This paper extends on the papers mentioned above by looking into the effect of all potential improvements made as a result of the audit under SOX 404(b) and directly investigating improvements made to the operational efficiency and financial performance of the firm. Compliance with SOX 404(b) made companies put more time and resources into creating an effective control environment (Iliev, 2010). In order for the report made under SOX 404(a) to be audited, extra discussion between clients and auditors are necessary. Auditors are experienced professionals on internal controls and business processes, so they have the ability to help improve these for the client, especially after performing their tests on the current internal controls (as necessary under SOX 404(b). This paper hypothesizes that these extra efforts of the auditor can lead to improvements made to areas of the control environment also outside of the material weaknesses, i.e., improvements that are either with regards to controls that where already sufficient or improvement of a smaller size, but that together could aggregate to significant improvements. These less severe improvements can also impact the operational efficiency, as suggested by Cheng et al. (2018). This paper also captures potential improvement to controls that would not necessarily lead to material misstatements if they were to contains weaknesses, but that could negatively affect the operational performance as a consequence of management making decisions based on less reliable information (Feng et al., 2009; Feng et al., 2015) that could for example weaken their tax planning (Bauer, 2016; Gallemore and Labro, 2015) or decrease their forecasting accuracy (Cassar and Gibson, 2008). Furthermore, the mere fact that the compliant company knows that their report on their internal controls will need to get audited can incentivize them to put more time and effort into improving their internal processes, in order to have less change of the auditor finding something in their tests of the internal controls. This would form an indirect benefit of the audit under SOX 404(b).

These possible benefits combined lead to the first hypothesis:

H 1: *Compliance with SOX 404(b) is positively associated with operational efficiency.*

Furthermore, following Feng et al. (2015), if these positive effects are in place, they would consequentially affect the company's financial performance. This logically leads to the forming of the second hypothesis:

H 2: *Compliance with SOX 404(b) is positively associated with financial performance.*

The inclusion of this second hypothesis adds to the reliability of this research as it can be investigated upon more easily based on hard data, whereas the operational efficiency is harder to capture. This is all further explained in the methodology (section 3.) below.

If the above hypotheses are proven to be correct this could help justify the costs associated with SOX 404(b) compliance. If they are proven to be incorrect, this would provide evidence that the implementation of SOX 404(b) provided too few benefits to compensate for the costs associated with it.

3. Methodology

3.1. Measuring operational efficiency

The operational efficiency is determined using data envelopment analysis (DEA). DEA uses an efficiency frontier to capture it's given output variable relative to other observations (firms in this case). In this paper, the DEA analysis performed closely follows the one in the paper of Cheng et al (2018). Sales revenue is taken as the output variable, as it is a primary source of the of the earnings and cash flows generated from a firm's operating activities (Demerjian et al., 2012; Cheng et al., 2018). Seven variables are then used as the input variables for the frontier, following the methodology in the paper of Cheng. These variables are Property, Plant & Equipment, Net Inventory, Cost of Goods Sold, Selling and Administrative Expenses, Research and Development Expenditures, Net Goodwill and Net Intangible Assets. This DEA constructs a measure for the relative operational efficiency of a firm compared to other firms in the sample. It does so by scoring each observation based on the weighted sum of outputs (in this case only sales revenue) divided by the weighted sum of inputs. The observation with the highest amount of sales revenue relative to its inputs is given an efficiency score of one, and the observation with the lowest amount is given a zero. It is possible that multiple observations receive either a zero or a one if they are deemed equally efficient by the DEA. The other observations receive a score that is between zero and one. This DEA approach is used for the same reason as it is used by Cheng: "DEA has advantages over a conventional efficiency approach, as it is non-parametric and does not require a functional relationship between inputs and outputs or a prior weighting of inputs". For a more elaborate description of this method, please refer to the paper of Chen et al. (2018).

3.2. Sample selection

3.2.1. Data sources and obtaining public float

The sample that is used in this research contains U.S. companies with fiscal year ends 2009 until 2019. All variables needed to determine both the operational efficiency and the financial performance, as well as the control variable are obtained from Compustat. The data about a company's public float is not available in any accessible datasets. In order to obtain it, the data needed is hand collected from the SEC website. The SEC provides files containing all information from companies' 10-K filings that are XBRL tagged for each filing the SEC received starting from 2009, combined per quarter. XBRL stand for eXtensible Business Reporting Language and is used in order to match values reported in 10-K filings to certain tags. The tag that contains a company's public float is 'EnitityPublicFloat', thus that is the tag that is filtered for in this research. After downloading all available quarterly tsv files, an R script is used to combine them, sort them in an understandable manner and filter them for the needed data. The SEC website only contains these files starting from 2009, hence why that will be taken as the starting year. Furthermore, 2019 is picked as the final year for this data sample because after that, amendments were made to the SOX 404 regulations, so including years after that could skew and/or bias the outcomes.

A description of how the data is obtained exactly is included in appendix B. Special thanks to Professor Roy Schmardebeck (co-writer on the McCallen (2020) paper) for providing additional guidance on how they obtained this data for the paper. Important to note is that in their paper they follow the above method first, but after that they scrape additional data from the files were the float is not XBRL tagged. A small summary of how they do so is as follows: they scrape for words that describe the public float and then they look for number that are preceded by a dollar sign and certain word combinations. Due to the complexity of this process, for this research only the first part is followed. This causes a loss of around 15% of the observations of the initial sample, see the following part of the sample selection procedure.

3.2.2. Sample selection procedure

The sample selection procedure is outlined in table 1. Starting with obtaining the public float data that as it is a critical variable for this research, the sample starts with 74,241 unique 10-K filings that were filing to the SEC in the years 2009-2019. At first, 10,940 observations are lost because the company's public float it is not XBRL tagged and thus not obtained when using the method described above. 53,374 variables are dropped because the float \ falls outside of the bandwidth (25 to 75 million USD). After this, 4,508 observations are excluded as not all necessary variables are available at Compustat Fundamentals Annual. Another 4,265 are dropped as not all needed variables are available at Compustat Financial Ratios. Lastly, 25 more observations are excluded because the company's compliance status data is not available on Audit Analytics. This leaves 1129 observations left for the main analyses of this research.

[Insert Table 1 here]

3.3. Research design

The following part will explain the methods used to investigate whether the hypotheses are correct. This research makes use of statistical data analyses (i.e., it is an empirical archival research). The dependent concepts used are the operational efficiency and the financial performance. Constructing these are the first step for performing the later analyses. Operational efficiency is already defined in section 3.1.. The financial performance is measured with a fairly simple construct using the Return On Equity (ROE), Return On Assets (ROA) and the Tobin's Q, similar as in Bouri (2018). The data for the ROE and ROA is publicly available. The data for the Tobin's Q however, has to be calculated. To do this the Tobin's Q is defined same as in the original paper of Brainard and Tobin (1968) as the ratio of the market value of equity plus the book value of debt to the book value of assets and it is calculated as follows:

$$\underline{Tobin's\ Q = (Total\ assets + Market\ value - Common\ equity)/(Total\ assets)\ (1)}$$

Here, Tobin's Q logically represents the Tobin's Q, Total assets refers to the total value of assets on their balance sheet, Market value refers to the worth of the company on the stock market, i.e. the share price times the amount of outstanding shares and Common equity refers to the value of the stocks held by common shareholders, i.e. the share price times the amount of shares held by common shareholders.

This research uses a two-stage least square regression (2SLS) design, were in the first stage SOX 404(b) compliance is predicted and is then used in the second stage as the instrument variable of a Regression Discontinuity Design (RDD). These stages and their equations are further explained below.

3.3.1. Predicting the compliance instrument (stage 1)

In the first stage, an instrument variable is created to predict compliance with SOX 404(b). This instrument variable is based on the float of a company, i.e., the value of the shares that are owned by public stockholders, excluding the value of shares that are for example locked-in by the company. The following represents the ordinary least squares regression used to predict compliance:

$$\underline{Compliant = \beta 0 + \beta 1 * float75 + \ \beta 2 * float50 + \ \beta x * controls + \ \varepsilon\ (2)}$$

Here, Compliant is and indicator for whether a company received and audit under SOX 404(b) (according to Audit Analytics), float75 is a dummy that equals 1 if the company float of the respective year is larger than $75 million, float50 is the same except that the threshold is $50 million.

Even though an advantage of a RDD design is that controls are less important for unbiased results, some are implemented as they are correlated with the dependent variables. For example, because a company's float is highly related to how large it is, this needs to be controlled for.

In the paper of Iliev (2010), a RDD design is also used looking at the effect of SOX 404(b). Therefore, based on that paper, the following control variables are included: a log of the market size (LogMarket), a log of the assets (LogAssets), the leverage (Leverage) and the receivables scaled by total assets (RecScaled). These control variables are the same for both equation 2, 3 and 4.

The cut-off is set at a float of $75 million, as this is a strong predictor for a company's compliance (McCallen et al., 2020). As stated by McCallen, companies with a public float above $75 million generally must comply with SOX 404(b), while companies below this cut-off generally do not. This means that for all companies with a float of $75 million or higher, they are predicted to be complying with SOX 404(b). To add on this companies that complied in the previous year are likely to remain complying unless their float falls below 50 million dollars in their current year (according to the SEC, 2005). Therefore, companies that had a float of over $75 million in the previous period and keep a float of $50 million of more in the current period are predicted to be complying as well. This instrument variable will be included as a dummy, meaning that it will have a value of 1 for companies that are predicted to be complying and 0 otherwise.

Important to note is that according to McCallen et al. (2020), there is no theoretical reason to believe that differences in the outcomes between the companies predicted to be or not be complying are caused by other factors than their SOX 404(b) compliance. This is important because for an instrument variable to be valid, it may only affect the dependent variable through its impact on the treatment (the compliance in this case) and not through other factors (e.g., other changes in characteristics that would happen at similar cut-offs). Further empirical evidence that the compliance predicted by float is a valid instrument is provided in the paper of McCallen, under the sections 3.3.2 till 3.3.4.

3.3.2. Regression discontinuity design model (stage 2)

Equation 2 to predict the compliance is consequentially used in two second regressions to investigate the effect of compliance on operational performance and separately financial performance. In this second stage of the 2SLS model a Regression Discontinuity Design (RDD) is used. This is used as it is a quasi-experimental design that offers a high internal validity. By keeping only the observations that lay within a certain range around the cut-off point of $75 million float there can be assumed that the companies will have similar underlying characteristics. Following the research of McCallen, a bandwidth of $100 million will be chosen. This means that only companies with a float between 25 and 125 million dollar float will be kept. The regression formulas for the second part of the two-stage least square regression are as follows:

$$Operational\ efficiency = \beta 0 + \beta 1 * \widehat{Compliant} + \beta x * controls + \varepsilon \ (3)$$

Here Operational efficiency represents the operational efficiency as measured using the data envelopment analysis, $\widehat{Compliant}$ represents the company's compliance status as predicted using equation 2 and controls represents the control variables same as in equation 2.

$$Finacial\ performance = \beta 0 + \beta 1 * \widehat{Compliant} + \beta x * controls + \varepsilon \ (4)$$

Here Financial performance represent the three different aspects that proxy for a company's financial performance and is substituted with all three of them separately in the analyses on financial performance (the three proxies being ROE, ROA and Tobin's Q) and the other variables are the same as in equation 3.

4. Results

4.1. Descriptive statistics

To get a grasp of the dataset that is used in this paper the descriptive statistics are displayed in Table 2 (section 8.1). The statistics are shown for all variables that are used in the further RDD analysis. Furthermore, Table 3 shows the data distribution by year. The mean values for each year are not significantly different and the years with a higher average float also have higher compliance rate, as would be logically expected. At first sight, there does not seem to be a clear connection between years with a higher compliance rate and their respective efficiencies. This does not point towards the direction of the first hypothesis being correct, but of course no conclusions can be drawn from this. Furthermore, it is clearly visible that the number of observations is not evenly distributed among the years. This is not expected and there is no logical explanation for this. This forms a weakness for this research, as the results might be biased towards certain years.

[Insert Table 2 here]

[Insert Table 3 here]

4.2. Strength of the instrument

To test the strength of the instrument variables (Float75 and Float50) the OLS to predict compliance is performed (see the methodology). The results are shown in Table 4, the first column without and the second with the control variables added. As visible in the table, both coefficients are positive (and large) and highly significant, with p-values of 2*10^-16 and 2.15*10^-13 respectively. This shows that both of the float variables are very solid instrument variables to predict compliance.

[Insert Table 4 here]

4.3. DEA frontier

Using R Studio, the DEA analysis is performed as described in the methodology. This results in each firm-year observation receiving a relative efficiency score between 0 and 1. Firms with a score of one generate the most output (sales revenue) relative to the input they use to do so. Firms with a score of 0 are the most inefficient in doing so. A significant part of the observations received a value of zero. This is possible and valid when using DEA analysis (Sarkis, 2007) as it does not mean that those firms were not able to create any output or that all their inputs are completely waisted, it only means that the input resources are a lot less efficiently used compared to others when taking only the output variable into consideration. The mean value of the efficiency is 0.309 (see descriptive statistics). Furthermore, the efficiency seems to be steady over time when looking at the yearly distribution, where significantly different values are only found in years where this can be attributed to the low number of observations in that specific year and thus more randomness.

4.4. OLS

Before moving to the second part of the two-stages least squares regression some simple OLS regressions are performed to get a better grip of the data. First, an OLS of compliance on efficiency is performed, both with and without the control variables. The results of this are shown in Table 5 (column 1 and 2 without and with the control variables respectively). The table shows that in this simplified model compliance has a significantly negative effect on efficiency. This could be explained by compliers for example having increased administrative expenses (being one of the input variables included in the DEA analysis) to ensure compliance.

[Insert Table 5 here]

A similar OLS regression is performed on the financial performance indicators, them being the return on equity (ROE), return on assets (ROA) and the Tobin's Q. The results are shown in Table 6, with column 1 and 2 representing the effect of compliance on ROE, columns 3 and 4 the effect of compliance on ROA and columns 5 and 6 representing the effect of compliance on the Tobin's Q, all without and with control variables respectively. As visible in the table, complying with SOX 404(b) seems to have a significant negative impact on both the ROE and ROA of the complying company. This suggest that, in contrast with the second hypothesis, complying with SOX 404(b) has a negative impact on the financial performance of a company. Of course, it must be taken into account that these simple OLS regressions do not have a strong validity, as there are a lot of endogeneity concerns that could bias the outcomes that are not attended to in these regressions, hence why in the next part the two-stage least squares regression design is used with a RDD approach, which has a way higher internal validity.

[Insert Table 6 here]

4.5. RDD H1: SOX compliance on operational efficiency

Using the iv_robust function of the estimatr package in R studio, the two stages of the two-stage least square regression model using an RDD design are simultaneously performed. The results presented in Table 7 show the result of the effect of compliance with SOX404(b) on operational efficiency through both the first and second stage of the two-stage least square regression model, with compliance being predicted by Float75 and Float50. The results on the first column are with and in the second stage without control variables.

As visible in the table, the variable of interest, being Compliant, has a positive value when the control variables are included. This is in line with the first conclusion. However, the effect is not statistically significant, as the p-value with control variables included is 0.17. This is not statistically significant as that would require a p-value of 0.05 or lower. So, although this regression does suggest that a positive effect of compliance on operational efficiency is in place, it cannot be stated that this is a causal effect.

[Insert Table 7 here]

4.6. RDD H2: compliance on financial performance

In the next three sub-section the results of the two-stage least squares regression models are given with the dependent variables being the proxies for the financial performance of a company: the return on equity, return on assets and the Tobin's Q respectively. For all these tests, the same methods and codes are used as above in section 4.5.

4.6.1. SOX compliance on return on equity

Table 8 shows the results of the two-stage least squares regression model with the dependent variable being the return on equity (ROE). As visible in column one, the compliance seems to have a positive impact on the return on equity that is significant on the 10% level. However, when the control variables are included, the coefficient turns negative and is not significant anymore (not even close to significant as the p-value of Complaint is 0.39 in that test).

Thus, this test fails to provide any evidence supporting the second hypothesis, as no results are found suggesting a significant positive effect of the compliance on the return of equity.

[Insert Table 8 here]

4.6.2. SOX compliance on return on assets

Table 9 shows the effect of the two-stage least squares regression model with the return on assets (ROA) included as the dependent variable. Again, the first column shows the results without the control variables and the second column the results with the control variables included.

The table shows that compliance has a negative coefficient that is statistically significant at the 1% level (with the p-value being approximately 0.0052). This provides evidence of a causal negative effect of compliance with SOX 404(b) and the return on assets of a company. This means that in contrast with the second hypothesis, compliance seems to have a negative impact on a company's financial performance. This implies that the costs of complying with SOX 404(b) outweigh the benefits, at least from the perspective of the company.

[Insert Table 9 here]

4.6.3. SOX compliance on Tobin's Q

The final financial performance indicator that is tested upon is the Tobin's Q. Table 10 shows the results of the two-stage least squares regression model with the Tobin's Q included as the dependent variable. As visible in the table, no significant results are found for the variable of interest. This means that no evidence is found that compliance with SOX 404(b) has a significant impact on the complier's Tobin's Q. This is not online with the second hypothesis, based upon a positive effect was expected.

[Insert Table 10 here]

4.7. Alternative bandwidths

The results found above in the sections 4.5 till 4.6.3. could potentially have been impacted by the selected bandwidth. In all those models, the bandwidth was set to 100 million dollars based on prior research. To test if this choice had an impact on the found results, the tests are reperformed with bandwidths of 80 and 60 million dollars (so a float between 35 to 115 and 45 to 105 respectively). A summary of these tests is shown in Table 11. Here, column 1 till 4 represent the outcomes of the two-stage least square regressions on operational efficiency, return on equity, return on assets and the Tobin's Q respectively, while using the smaller bandwidth of 80 million dollars. Column 5 till 8 represent the same, except for that they show the results for the bandwidth of 60 million dollars. As visible in the table, the coefficients for both bandwidths remain in the same direction as in the tests with the original bandwidths. The coefficients that where significant remain so and vice versa. An exception is the effect of compliance on ROE when using the bandwidth of 60 million dollars. Here the direction remains the same, but in contrast with the bandwidths of 80 and 100 million, it is significant. This provides further evidence against the second hypothesis and is in line with expectations based on the previous results. As the patterns found previously remain similar in these alternative analyses, this suggests that the initial results are not biased or invalid because of the choice of bandwidth.

[Insert Table 11 here]

5. Conclusion

This research tries to answer the question whether the benefits of the implementations of section 404 of the Sarbanes-Oxley Act outweighed the costs. Specifically, it investigates if SOX 404(b) has any direct benefits for the company that is complying in the form of either an increased operational efficiency or an increased financial performance. Here, the operational efficiency is proxied for using a DEA design and the financial performance is proxied for by three separate indicators: the return on equity, the return on assets and the Tobin's Q. To perform the analysis, a fuzzy RDD design is used, with a company's public float serving as the instrumental variable for compliance. Even though it was hypothesized that compliance with SOX 404(b) would have a positive effect on both the operational efficiency and the financial performance, there is no clear evidence found to support this. There are some coefficients found with the expected direction, mainly compliance on operation efficiency, but these effects are not statistically significant. To the contrary, there is statistically significant evidence provided that compliance actually has a negative effect on a company's financial performance, reflected in its return on assets, as well as in its return on equity when using the alternative bandwidth. Altogether, this means that both hypotheses need to be rejected, meaning that compliance with SOX 404(b) provides no significant benefits for either the complier's operational efficiency nor its financial performance. It might even have a negative impact, but this is outside the scope of this paper and thus left to future research.

These results can be combined with those of the paper of McCallen et al. (2020) to suggest that the overall costs of compliance outweigh the benefits of SOX 404(b), meaning that the implementation probably was not economically justifiable. A limitation of this paper is the time frame of the data used. While SOX went into effect in 2004, this data of this paper only starts from 2009. Prior research suggests that the effect of SOX where more pronounced in the earlier years after its implementation, meaning that significant results might be able to be found there. Also, there are some other data related issues in this research that are elaborated upon in the introduction. Furthermore, it would be interesting for future research to examine whether the changes made to the conditions to be exempted from complying (effective in 2020) have made any changes to the costs and/or benefits of the regulation. Especially since these changes were based on the idea to lower the costs for companies for whom the burden would be unreasonably high.

References

Ashbough-Skaife H., Collins D., Kinney Jr. W. Lafond R. 2009. *The Effect of SOX Internal Control Deficiencies on Firm Risk and Cost of Equity.* Journal of Accounting Research.

Alexander C., Bauguess G., Bernile Y., Lee H., Mariette-Westberg J. 2013. *Economic effects of SOX Section 404 compliance: a corporate insider perspective*. Journal of Accounting Research.

Barth M., Landsman J., Schroeder H., Taylor D. 2019. *Re: Amendments to the Accelerated Filer and Large Accelerated Filer Definitions.* Available at: https://www.sec.gov/comments/s7-06-19/s70619-5802113-187069.pdf

Bauer A. 2016. *Tax avoidance and the implications of weak internal controls.* Contemporary Accounting Research.

Brainard W.C. and Tobin J. 1968. *Pitfalls in financial model building.* American Economic Review.

Bouri A., Ajili H. 2018. *Corporate governance quality of Islamic banks: measurement and effect on financial performance.* International Journal of Islamic and Middle Eastern Finance and Management.

Cassar G and Gibson B. 2008. *Budgets, internal reports, and manager forecast accuracy.* Contemporary Accounting Research

Cheng Q., Goh B.W., Kim J.B. 2018. *Internal control and operational efficiency.* Contemporary Accounting Research.

Chen Y., Lu B., Zhou G. 2020. *Can internal audit functions improve firm operational efficiency? evidence from China.* Managerial Auditing Journal.

Chhaochharia V., Grinstein Y. 2007. *Corporate governance and firm value: The impact of the 2002 governance rules.* Journal of Finance.

Clay, C., Kim D. 2017. Sarbanes-Oxley: 15 Years of Successes and Challenges. Accounting Today.

Coates J., Srinivasan S. 2014. *SOX After Ten Years: A Multidisciplinary Review.* Accounting Horizons

DeFond M., Lennox C. 2017. *Do PCAOB Inspections Improve the Quality of Internal Control Audits?* Journal of Accounting Research.

Demerjian P., Lev B., McVay S. 2012. *Quantifying managerial ability: a new measure and validity tests.* Management Science.

Doogar R., Sivadasan P., Solomon I. 2010. *The regulation of public company auditing: evidence from the transition to AS5.* Journal of Accounting Research.

Feng M., Li C., McVay. 2009. *Internal control and management guidance.* Journal of Accounting and Economics.

Feng M., Li C., McVay S., Skaife H. 2015. *Does ineffective internal control over financial reporting affect a firm's operations? Evidence from firms' inventory management.* The Accounting Review.

Gallemore J, Labro E. 2015. *The importance of the internal information environment for tax avoidance.* Journal of Accounting and Economics.

Honingsberg C., Rajgopal S. 2019. *Re: Amendments to the accelerated filer and large accelerated filer definitions; proposed rules, request for comments.* Available at: https://www.sec.gov/comments/s7-06-19/s70619-5849740-188567.pdf

Lui Z., Cheng Y., Lui R. 2020. *Why Non-accelerated Filers Voluntarily Comply with SOX 404b?* International Journal of Accounting & Finance Review.

Iliev, P. 2010. *The Effect of SOX Section 404: Costs, Earnings Quality, and Stock Prices.* The Journal of Finance.

McCallen J., Scharmebeck R., Shipman J., Whited R. 2020. *Have the Effects of SOX Section 404(b) Compliance Changed Over Time?* Unpublished working paper, available at: https://ssrn.com/abstract=3420787.

Powell S. 2005. *Costs of Sarbanes-Oxley are out of control*. letters to the editor. Wall Street Journal.

Romano R. 2005. *The Sarbanes-Oxley Act and the making of quack corporate governance*. Yale Law Journal.

Sarkis J. (2007). *Preparing Your Data for DEA.* CH.17 in Zhu J., Cook WD. Modeling Data Irregularities and Structural Complexities in Data Envelopment Analysis. Springer, Boston.

Schroeder J., Shephardson M. 2016. *Do SOX 404 Control Audits and Management Assessments Improve Overall Internal Control System Quality?* The Accounting Review.

SEC. 2005. *Revisions To Accelerated Filer Definition and Accelerated Deadlines for Filing Periodic Reports*. https://www.sec.gov/rules/final/33-8644.pdf.

Stoel M., Muhanna W. 2011. *IT internal control weaknesses and firm performance: and organizational liability lens.* International Journal of Accounting Information Systems.

Appendix A – variable definitions

Variable:	Definition:
Common equity	The total value of existing common equity of a company, as found on Compustat.
Compliant	An indicator variable that is equal to one if a company is compliant with SOX404, as found on Audit Analytics and zero otherwise.
$\widehat{Compliant}$	An indicator variable that is equal to one if a company is predicted to be compliant with SOX404, as explained in section 3.3.1..
Efficiency	The relative operational efficiency on a scale of zero to one of a company compared to the others based on DEA analysis as described in section 3.1.
Financial performance	The financial performance of a company, proxied for by the return on equity, return on assets and Tobin's Q, see section 3.3. for further explanation.
Float(mil)	The public float of a company as found in their 10-K report, measured in millions.
Float50	An indicator variable that is equal to one if the Float(mil) of a company is 50 of higher and zero otherwise.
Float75	An indicator variable that is equal to one if the Float(mil) of a company is 75 of higher and zero otherwise.
Leverage	The total amount of debt of a company divided by its equity, as found on Compustat.
LogAssets	The logarithm of the total assets of the respective company, as found on Compustat.
LogMarket	The logarithm of the total market value of a company, as found on Compustat.
Market value	The total market value of a company, as found on Compustat.
Operational efficiency	The operational efficiency of a company as determined using data envelopment analysis, see section 3.1.
RecScaled	The total receivables of the company divided by its total assets, as found on Compustat.
ROA	The return on assets of a company, as found on Compustat.
ROE	The return on equity of a company, as found on Compustat.
Tobin's Q	The Tobin's Q of a company calculated as described in equation 1.
Total assets	The total assets of a company, as found on Compustat.

Appendix B – obtaining PublicFloat from the SEC

In order to obtain the PublicFloat data (or other variables that can be found in 10-K reports and that are XBRL tagged), the first step is to visit the website below:

https://www.sec.gov/dera/data/financial-statement-and-notes-data-set.html

Here, on the bottom of the page, the following list of downloads can be found:

Data Downloads

File	Format	Size
2022 05 *Monthly*	ZIP	289.63 MB
2022 04 *Monthly*	ZIP	188.28 MB
2022 03 *Monthly*	ZIP	358.08 MB
2022 02 *Monthly*	ZIP	361.39 MB
2022 01 *Monthly*	ZIP	45.86 MB
2021 12 *Monthly*	ZIP	70.9 MB
2021 11 *Monthly*	ZIP	393.48 MB

The ZIP files contain packages with multiple datasets containing different packages. These packages contain different information about every company that filed with the SEC in the time period displayed in the name of the ZIP file (e.g., 2017 Q2 is contain information about companies that filed with the SEC in the second quarter of 2017). The ZIP file itself will look as follows:

s PC › Downloads › 2022_05_notes.zip

Search 2022_05_notes.zip

Name	Type	Compressed size	Password ...	Size	Ratio	Date modified
cal.tsv	TSV File	6.906 KB	No	69.463 KB	91%	2-6-2022 09:02
dim.tsv	TSV File	10.880 KB	No	42.875 KB	75%	2-6-2022 09:01
notes-metadata.json	JSON File	8 KB	No	67 KB	90%	2-6-2022 09:03
num.tsv	TSV File	66.688 KB	No	1.193.109 KB	95%	2-6-2022 09:02
pre.tsv	TSV File	45.590 KB	No	452.687 KB	90%	2-6-2022 09:02
readme.htm	Microsoft Edge HTML Do...	19 KB	No	262 KB	93%	2-6-2022 09:03
ren.tsv	TSV File	9.783 KB	No	76.497 KB	88%	2-6-2022 09:01
sub.tsv	TSV File	1.000 KB	No	4.112 KB	76%	2-6-2022 09:01
tag.tsv	TSV File	19.138 KB	No	104.839 KB	82%	2-6-2022 09:01
txt.tsv	TSV File	136.573 KB	No	641.977 KB	79%	2-6-2022 09:02

Further explanation of which data is stored where and how the tables can be linked can be found on the top of the previously said link, when pressing on the "The Financial Statements and Notes Data" PDF link. For this specific research, in order to obtain the PublicFloat data the sub.tsv and num.tsv files where needed. The sub.tsv file contains information about the filed report e.g., its date, time period it reflects, CIK number and the unique filing number 'adsh' that is used to link the content of the different files together. The num.tsv file contains a multitude of financial data for the respective company, including a lot of XBRL tagged items. This file contained the 'EntitiyPublicFloat' tag representing the respective company's value of public float that is used in this research. An example of one of these num.tsv files is shown below:

adsh	tag	version	ddate	qtrs	uom	dimh	iprx	value	
0000836658-15-000031	DeferredPolicyAcquisitionCostsAdditions	us-gaap/2014	20140930	3	USD	0x00000000	0	92	
0000836658-15-000031	DeferredPolicyAcquisitionCostsAdditions	us-gaap/2014	20150930	3	USD	0x00000000	0	77	
0000837010-15-000067	DeferredPolicyAcquisitionCostsAdditions	us-gaap/2014	20140930	3	USD	0x00000000	0	51	
0000837010-15-000067	DeferredPolicyAcquisitionCostsAdditions	us-gaap/2014	20150930	3	USD	0x00000000	0	56	
0000844059-15-000024	EntityPublicFloat	dei/2014	20150331	0	USD	0x00000000	0	224249452.0000	
0000832480-15-000022	EntityPublicFloat	dei/2014	20130630	0	USD	0x00000000	0	0.0000	0
0000833444-15-000091	EntityPublicFloat	dei/2014	20150331	0	USD	0x00000000	0	17741657194.0000	
0000831259-15-000055	EntityPublicFloat	dei/2014	20140630	0	USD	0x00000000	0	37300000000.0000	
0000814676-15-000045	EntityPublicFloat	dei/2014	20151031	0	USD	0x00000000	0	19900000.0000	
0000814547-15-000019	EntityPublicFloat	dei/2014	20150331	0	USD	0x00000000	0	2098972006.0000	
0000817979-15-000041	EntityPublicFloat	dei/2014	20140630	0	USD	0x00000000	0	0.0000	0
0000823277-15-000024	EntityPublicFloat	dei/2014	20150831	0	USD	0x00000000	0	0.0000	0
0000829224-15-000038	EntityPublicFloat	dei/2014	20150331	0	USD	0x00000000	0	69000000000.0000	
0000826675-15-000031	EntityPublicFloat	dei/2014	20140630	0	USD	0x00000000	0	448865442.0000	
0000825410-15-000056	EntityPublicFloat	dei/2014	20150331	0	USD	0x00000000	0	334000000.0000	
0000808450-15-000065	EntityPublicFloat	dei/2014	20150430	0	USD	0x00000000	0	1500000000.0000	
0000812128-15-000014	EntityPublicFloat	dei/2014	20150430	0	USD	0x00000000	0	1229955610.0000	
0000801898-15-000122	EntityPublicFloat	dei/2014	20150430	0	USD	0x00000000	0	4200000000.0000	
0000808326-15-000038	EntityPublicFloat	dei/2014	20150331	0	USD	0x00000000	0	156300000.0000	
0000807882-15-000024	EntityPublicFloat	dei/2014	20150331	0	USD	0x00000000	0	3500000000.0000	
0000785557-15-000042	EntityPublicFloat	dei/2014	20150331	0	USD	0x00000000	0	8419975.0000	

Note that the column definitions above are taken from the top of the document and added for more clarification. Here you can see the unique filing number 'adsh', along with the correct tag 'EntitiyPublicFloat' and the corresponding value in the ninth column.

First, I downloaded all separate ZIP files for the time period used in this research. I then placed all sub.tsv and num.tsv files together in the same map on my computer, while renaming them to follow a sequential order, i.e., sub1.tsv, sub2.tsv etcetera. This was all done by hand.

I then wrote an R script that loops over all the different tsv files and binds everything together. The code for this is quite simple:

```
```{r}
df <- data.frame()
for (i in 1:43){
#load id and float
num <- read_tsv(paste0("num", i, ".tsv"))
numfilt <- filter(num, tag == "EntityPublicFloat" )

#load id and cik
sub <- read_tsv(paste0("sub", i , ".tsv"))

#combine cik and float and filter
combi <- merge(numfilt, sub, by = "adsh")
combifilt <- subset(combi, select = c("adsh", "cik", "name", "tag", "value", "period"))

#add year based on period
combifiltyear <- mutate(combifilt, year = as.numeric(substr(combifilt$period, 1, 4)))

#add to previous
final <-  rbind(final, combifiltyear)
}
```

For the exact meaning of the columns/variables selected, please refer to the PDF file mentioned above. Special thanks to Professor Roy Schmardebeck (co-writer on the McCallen paper) for helping me out with the above mentioned website.
```

Tables

Table 1 – Sample selection

Sample selection process	
Unique 10-K filings of companies to the SEC between 2009-2019	74,241
Less: Those where the company public float is not XBRL tagged	(10,940)
Less: Those where the company public float is either below 25 million or above 75 million	(53,374)
Less: Those where for the firm year observations the needed variables are not available on Compustat Fundamentals Annual	(4,508)
Less: Those where for the firm year observations the needed variables are not available on Compustat Financial Ratios	(4,265)
Less: Those where for the firm year observations the complaint status is not available on Audit Analytics	(25)
Number of observations used in all analyses (except for alternative bandwidths)	1129

Table 2 – Descriptive statistics

Variable	N	Mean	Std.Dev.	Min	Pctl..25	Pctl..75	Max
Compliant	1129	0.425	0.498	0	0	1	2
Float(mil)	1129	66.687	27.966	25.090	43.757	88.642	125
Float75	1129	0.356	0.479	0	0	1	1
Float50	1129	0.657	0.475	0	0	1	1
Efficiency	1129	0.31	0.318	0	0	0.585	1
ROE	1129	-0.093	0.565	-8.071	-0.084	0.102	2.154
ROA	1129	0.012	0.223	-2.354	0.014	0.107	0.748
TobinsQ	1129	1.473	1.231	0.275	0.978	1.461	19.911
LogAssets	1129	14.49	1.368	11.017	13.476	15.715	19.413
LogMarket	1129	13.699	0.841	10.855	13.161	14.124	18.68
Leverage	1129	3.412	5.064	0.009	0.444	6.172	77.247
RecScaled	1129	0.297	0.26	0	0.091	0.562	0.928

This table shows descriptive statistics for all variables used in this research. The columns showcase the value that is shown in the first row.

Table 3 – Yearly distribution

Year	*N*	Mean *Compliant*	Mean *Float75*	Mean *Float 50*	Mean *Float (millions)*	Mean *Operational efficiency*
2009	4	0.75	0.75	0.75	79.554	0.112
2010	9	1	0.556	0.778	74.601	0.071
2011	296	0.395	0.331	0.615	64.484	0.275
2012	201	0.433	0.333	0.627	64.083	0.331
2013	152	0.414	0.322	0.671	65.879	0.355
2014	123	0.366	0.358	0.707	68.062	0.366
2015	114	0.491	0.421	0.719	70.742	0.369
2016	91	0.495	0.33	0.626	65.146	0.336
2017	68	0.368	0.382	0.676	69.956	0.255
2018	57	0.439	0.474	0.702	70.960	0.234
2019	14	0.357	0.357	0.714	74.748	0.206

This table shows the distribution of the sample per year, with each column showcasing the mean value of the mentioned variable in that specific year.

Table 4 – OLS: instrument strength: instruments on compliance

	Compliant		*Compliant*	
Variable	Coeff.	Std. error	Coeff.	Std. error
Float75	0.4948***	0.029	0.4401***	0.0291
Float50	0.2175***	0.0292	0.1169***	0.0292
LogAssets			0.0487***	0.0138)
LogMarket			0.0672***	0.0181
Leverage			-0.0019	0.0029
RecScaled			-0.3362***	0.0615
N	1129		1129	
R-squared	0.3744		0.4135	

This table report the results of the OLS regression of the instrument variables Float75 and Float50 on compliance. Float75(50) is an indicator variable equal to one for companies with a public float of $75(50) million or larger and zero otherwise. *, **, *** indicate the statistical significance of the found coefficient, they mean that the p-value of the coefficient is lower than 0.10, 0.05, 0.01 respectively. Variable definitions are included in the Appendix.

Table 5 – OLS: compliance on efficiency

	Operational efficiency		*Operational efficiency*	
Variable	Coeff.	Std. error	Coeff.	Std. error
Compliant	-0.107***	0.0187	-0.0456**	0.0175
LogAssets			-0.0144	0.009
LogMarket			-0.0199	0.012
Leverage			0.0044*	0.002
RecScaled			0.658***	0.0428
N	1129		1129	
R-squared	0.0275		0.3201	

This table report the results of the OLS regression of compliance on efficiency. Compliant is an indicator variable equal to one for companies that comply with SOX404(b) and zero otherwise. *, **, *** indicate the statistical significance of the found coefficient, they mean that the p-value of the coefficient is lower than 0.10, 0.05, 0.01 respectively. Variable definitions are included in the Appendix.

Table 6 – OLS: compliance on financial performance

	ROE		*ROE*		*ROA*		*ROA*		*Tobin's Q*		*Tobin's Q*	
Variable	Coeff.	Std. error	Coeff.	Std. error	Coeff.	Std. error	Coeff.	Std. error	Coeff.	Std .error	Coeff.	Std. error
Compliant	-0.0249	0.03379	-0.1148***	0.0168	-0.0073	0.0133	-0.0644***	0.014	0.0977	0.0735	-0.0446	0.0595
LogAssets			0.1921***	0.0178			0.0546***	0.0076			-0.0907***	0.0324
LogMarket			-0.0086	0.022			0.0364***	0.0096			1.0111***	0.041
Leverage			-0.0599***	0.0038			-0.0084***	0.0016			0.0657***	0.0069
RecScaled			0.4848***	0.08			0.0187	0.0342			1.1234***	0.1253
N	1129		1129		1129		1129		1129		1129	
R-squared	-0.0034		0.2481		-0.0069		0.1198		0.0067		0.4773	

This table report the results of the OLS regression of the compliance variable on financial performance. Compliant is an indicator variable equal to one for companies that comply with SOX404(b) and zero otherwise. The financial performance is proxied for using return on equity in the 1st and 2nd column, return on assets in the 3rd and 4th column and the Tobin's Q in the 5th and 6th column. *, **, *** indicate the statistical significance of the found coefficient, they mean that the p-value of the coefficient is lower than 0.10, 0.05, 0.01 respectively. Variable definitions are included in the Appendix.

Table 7 – 2SLS: compliance on efficiency

	Operational efficiency		*Operational efficiency*	
Variable	Coeff.	Std. error	Coeff.	Std. error
$\widehat{Compliant}$	-0.02594	0.0312	0.046	0.0335
LogAssets			-0.0204	0.0125
LogMarket			-0.0381***	-0.0141
Leverage			0.0045	0.0044
RecScaled			0.694***	0.0526
N	1129		1129	
R-squared	0.0111		0.3036	

This table report the results of the RDD regression of compliance on efficiency, with compliance being predicted through the instrument variables Float75 and Float50. Float75(50) is an indicator variable equal to one for companies with a public float of $75(50) million or larger and zero otherwise. *, **, *** indicate the statistical significance of the found coefficient, they mean that the p-value of the coefficient is lower than 0.10, 0.05, 0.01 respectively. Variable definitions are included in the Appendix.

Table 8 – 2SLS: compliance on ROE

	ROE		*ROE*	
Variable	Coeff.	Std. error	Coeff.	Std. error
$\widehat{Compliant}$	0.0958*	0.0586	-0.0705	0.0834
LogAssets			0.1891***	0.0487
LogMarket			-0.0175	0.0504
Leverage			-0.0598**	0.024
RecScaled			0.502***	0.1686
N	1129		1129	
R-squared	0.0111		0.2468	

This table report the results of the RDD regression of compliance on return on equity, with compliance being predicted through the instrument variables Float75 and Float50. Float75(50) is an indicator variable equal to one for companies with a public float of $75(50) million or larger and zero otherwise. *, **, *** indicate the statistical significance of the found coefficient, they mean that the p-value of the coefficient is lower than 0.10, 0.05, 0.01 respectively. Variable definitions are included in the Appendix.

Table 9 – 2SLS: compliance on ROA

	ROA		*ROA*	
Variable	Coeff.	Std. error	Coeff.	Std. error
$\widehat{Compliant}$	0.0252	0.0228	-0.0871***	0.0311
LogAssets			0.0561***	0.0145
LogMarket			0.0409***	0.0144
Leverage			-0.0085**	0.0042
RecScaled			0.0099	0.035
N	1129		1129	
R-squared	-0.0058		0.1177	

This table report the results of the RDD regression of compliance on return on assets, with compliance being predicted through the instrument variables Float75 and Float50. Float75(50) is an indicator variable equal to one for companies with a public float of $75(50) million or larger and zero otherwise. *, **, *** indicate the statistical significance of the found coefficient, they mean that the p-value of the coefficient is lower than 0.10, 0.05, 0.01 respectively. Variable definitions are included in the Appendix.

Table 10 – 2SLS: compliance on Tobin's Q

	Tobin's Q		*Tobin's Q*	
Variable	Coeff.	Std. error	Coeff.	Std. error
$\widehat{Compliant}$	0.3062	0.1308	0.017	0.1068
LogAssets			-0.9119***	0.0982
LogMarket			0.9988***	0.0973
Leverage			0.0658**	0.0321
RecScaled			1.1650***	0.2235
N	1129		1129	
R-squared	-0.0064		0.4768	

This table report the results of the RDD regression of compliance on return on Tobin's Q, with compliance being predicted through the instrument variables Float75 and Float50. Float75(50) is an indicator variable equal to one for companies with a public float of $75(50) million or larger and zero otherwise. *, **, *** indicate the statistical significance of the found coefficient, they mean that the p-value of the coefficient is lower than 0.10, 0.05, 0.01 respectively. Variable definitions are included in the Appendix.

Table 11 – 2SLS: alternative bandwidths

	Operational efficiency		*ROE*		*ROA*		*Tobin's Q*		*Operational efficiency*		*ROE*		*ROA*		*Tobin's Q*	
Variable	Coeff.	Std. error	Coeff.	Std. error	Coeff.	Std. error	Coeff.	Std. error	Coeff.	Std .error	Coeff.	Std. error	Coeff.	Std. error	Coeff.	Std. error
$\widehat{Compliant}$	0.0465	0.0395	-0.0973	0.0996	-0.0833**	0.0375	-0.0254	0.1323	0.0013	0.0506	-0.2334**	0.0926	-0.1079**	0.0426	-0.0284	0.1731
Controls	*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*	
N	900		900		900		900		667		667		667		667	
R-squared	0.3274		0.159		0.0968		00.4917		0.3643		0.1233		0.0781		0.4918	

This table report the results of the RDD regressions of the compliance variable on efficiency, return on equity, return on assets and the Tobin's Q, with compliance being predicted through the instrument variables Float75 and Float50. Float75(50) is an indicator variable equal to one for companies with a public float of $75(50) million or larger and zero otherwise. These regressions are performed using alternative bandwidths of the sample compared to the other regressions. In the columns 1 to 4(5 to 8) a bandwidth of public float of 80$(60$) million is used. The columns 1 and 5 represent the regressions on efficiency, 2 and 6 on return on equity, 3 and 7 on return on assets and 4 and 8 on Tobin's Q. *, **, *** indicate the statistical significance of the found coefficient, they mean that the p-value of the coefficient is lower than 0.10, 0.05, 0.01 respectively. Variable definitions are included in the Appendix.