ERASMUS UNIVERSITEIT ROTTERDAM
Erasmus School of Economics
Masterthesis Accounting, Auditing and Control

"Collecting more taxes than is absolutely necessary is legalized robbery."

Financial statement comparability & tax avoidance

The effect of more comparable financial statements on tax avoidance

Author: G.R.C. (Glenn) Mostert
Studentnumber: 406235

Supervisor: dr. J. Pierk
Second supervisor: drs. R. van der Wal RA

Date final version: **9 September 2018**

Front quote: J. Calvin Coolidge jr. (1872– 1933), 30th president of the United States of America from 1923 till 1929.

*** ABSTRACT ***

I examine whether the comparability of financial statements in a two-digit SIC industry affects corporate tax avoidance behavior in that industry. I determine financial statement comparability using a relatively new and innovative measure: XBRL-tags. My regression models show, after controlling for relevant factors based on prior literature, that industries where financial statements are more comparable are characterized by less corporate tax avoidance or evasion (so: financial statement comparability reduces tax avoidance). This suggests that financial statement comparability reduces managerial incentives to engage in tax avoidance, because better financial statement comparability helps tax authorities in detecting tax avoidance. Contrary to prior literature I find that the effect that financial statement comparability has on tax avoidance is less pronounced for firms with less transparent information environments. Only in case of a transparent information environment, the effect of financial statement comparability seems more pronounced. Multiple main regressions and a robustness analysis assure that these findings are robust.

Table of contents

1. Introduction, research question and motivation

Generally, nobody likes paying taxes. We rather pay less taxes than more. This is reflected by the front quote by the 30th president of the United States of America: J. Calvin Coolidge jr. It is also why a lot of companies try to avoid taxes, some more subtle than others (see for example Philips et. al. 2017). However, paying taxes is an integral and a very important part of society, as is pointed out by Cass Sunstein[1]: "*There is no liberty without dependency. That is why we should celebrate Tax Day.*". Sunstein argues that taxes provide the whole basis for our social structure and our judicial and political system. Without taxes, there would be no civilized society with politically and legally enforced social norms possible. Despite that, society sees and is used to seeing companies, even huge and worldwide known companies, avoiding and evading taxes (for example Apple, that avoids taxes on a very large scale using shell techniques in Ireland: Bowers 2017). These tax evasion and avoidance practices are harmful to the economic growth and the affordability of social benefits (GAO 2008). Hence, tax authorities focus on corporate tax compliance and seek to maximize tax revenue by balancing enforcement measures with the corporate tax payer morale (Filippin et al. 2013). Financial statement comparability is an important instrument for tax authorities to check (and enforce if needed) tax compliance. Comparability helps tax authorities better understand a firm's tax-relevant information and helps detecting anomalies, such as internal transfer prices which are not at arm's length, earlier and easier (De Franco et al. 2011). Therefore, financial statement comparability is expected to aid tax authorities in detecting corporate tax non-compliance. This might discourage managers and CEO's to engage in tax avoidance practices (Suk and Zhao 2017). As such, financial statement comparability also affects tax compliance. This effect is subject to discussion though. Others argue that financial statement comparability increases corporate tax avoidance practices (De Simone 2016). Since taxation and tax revenues are such essentials parts of the modern-day society, it is important to have a clear understanding of the factors that might affect tax revenues and tax avoidance. In this thesis therefore, I focus on the effect that financial statement comparability has on corporate tax avoidance behavior. More specifically, I investigate and answer the following question:

Research Question: Does a higher level of financial statement comparability induce less aggressive tax avoidance by corporations?

I also investigate whether a firm's information environment has a moderating effect on the relation between financial statement comparability and tax avoidance.

[1] Administrator of the White House Office of Information and Regulatory Affairs under the 2009-2012 Obama Administration, current Robert Walmsley University Professor at Harvard Law School and Honorary Docter at Copenhagen Business School.

Research to the correlation between comparability and tax avoidance, and the moderating variables in that relation, is scarce, despite the importance. Only a few studies have examined this relation before. As said however, the results of these studies, performed by De Simone (2016) and Suk and Zhao (2017), are contradictory. It is thus still unclear whether and how comparability affects tax avoidance behavior. Suk and Zhao (2017) argue that managerial opportunism, such as accounting manipulations aimed at avoiding taxes, can be limited with more comparable financial statements. De Simone argues the opposite. She argues that more comparable financial statements result in a wider range of plausible arm's length transfer prices, which would increase the amount of profit managers could (justifiably) shift or manage.

I investigate the relation between tax avoidance and financial statement comparability by regressing tax avoidance proxies (I use two effective tax rate measures and two sheltering probabilities) on a proxy for financial statement comparability. To investigate the effect that the information environment has, I add an interaction effect. Financial statement comparability is a construct that is hard to operationalize (Gordon and Gallery 2012). In extant literature, a few different methods are used (see for example De Franco et al. 2011; Chen et al. 2017; Suk and Zhou 2017). However, Hoitash et al. (2017) recently developed a new and promising measure for financial statement comparability. They base their measure on the XBRL-tags reported in SEC filings. Hoitash et al. (2017) show that this method of measuring financial statement comparability is superior to other measures of comparability, especially with regard to financial reporting quality and managerial discretion, both important factors in relation to tax avoidance. I use an interpretation of this new measure to investigate the relation between financial statement comparability and tax avoidance.

My thesis shows, being the second study in a row, that corporate tax avoidance decreases as financial statement comparability increases. This result is both statistically and economically significant. This is most likely due to the fact that comparability makes it easier for tax authorities to detect tax avoidance and evasion, which in turn discourages managers from engaging in such practices. This indicates that it is beneficiary to design accounting standards with the goal of achieving more comparable financial statements. This also indicates that it is worthwhile to include the effect of accounting comparability into the designing of (anti-)tax (sheltering) models by tax authorities. The evidence of this study further suggests that the emphasis the FASB and the SEC place on the comparability of financial statements (Suk and Zhao 2017) is beneficial since this reduces corporate tax avoidance and evasion. As such, the results of this thesis are relevant for tax authorities and policy makers, regulators and standard setters.

Contrary to what I expected however, I find that a poor information environment has a mitigating effect on tax avoidance: the effect of financial statement comparability on tax avoidance is less pronounced when firms have a more opaque information environment. On the other hand, a transparent information environment does make the effect that financial statement comparability has on tax avoidance more pronounced. This suggests that financial statement comparability cannot completely offset the tax avoidance-incentive of an opaque information environment, while in a transparent information environment tax avoidance decreases even more. This indicates that tax authorities should focus extra on tax compliance in industries with a poor(er) information environment on average. Increasing the transparency of the information environment and increasing the financial statement comparability in such industries would result in a relatively big increase in tax revenues.

With regard to the literature on financial statement comparability, my thesis adds to the growing number of studies that focus on the consequences and benefits of comparability since the call of Schipper (2003), for example Brown and Kimbrough (2011), Kim et al. (2013), Francis et al. (2014), Choi et al. (2015), Chen et al. (2017) and Campbell and Yeung (2017). I specifically contribute to this research by showing that financial statement comparability is beneficial for tax authorities because it reduces tax avoidance and evasion.

With regard to the literature on corporate tax avoidance or evasion, my thesis adds to the academic discussion on the interaction between the information environment and corporate tax avoidance. Weisbach (2002) finds that companies forgo tax avoidance benefits for some reason. This is called the 'under-sheltering puzzle'. Hanlon and Heitzman (2010) explain that this might partly be due to the fact that corporate tax compliance is, among other things, determined by the firm's information environment. In their concluding remarks, Hanlon and Heitzman (2010) call, as does Maydew (2001), for research that further focusses on the role of information uncertainty, the information environment and the information quality in the avoidance of tax. I answer to this call by investigating what effect the information environment has on the relation between financial statement comparability and tax avoidance. My results indicate that a firm's information environment does indeed moderate the effect of financial statement comparability on tax avoidance, however not in the expected direction.

This paper continues as follows. Paragraph 2 gives a brief overview of the related literature on financial statement comparability and tax avoidance. Paragraph 3 develops the hypotheses. Paragraph 4 describes my sample and research design. Paragraph 5 describes the empirical analysis. Paragraph 6 shows the results of the statistical analysis of the data. Paragraph 7 provides an extra robustness check for these results. Lastly, paragraph 8 concludes this thesis.

2. Related key literature and background information

2.1. Tax Avoidance

2.1.1. Definitions

Taxes can be defined as a “compulsory levy by the government on the people’s income or wealth without a direct quid pro quo” (Song and Yarbrough 1978). Taxes are typically levied on profits, wages and other types of income (Payne and Raiborn 2018).

Tax avoidance is broadly defined as the reduction of taxes. However, a reduction of taxes can be achieved with tax planning, ranging from completely legal structures to blunt tax noncompliance, evasion and sheltering (Hanlon and Heitzman 2010). More specifically therefore, tax avoidance is defined as using legal means to reduce the total amount of tax payable (Hansen et al. 1992). Tax evasion on the other hand is defined as any real action taken, not using legal means, to reduce or conceal taxable incomes amounts or increase tax deductions in order to reduce the total amount of tax payable. Examples are the deliberate delaying or accelerating of profit making and order taking or the intentional misrepresentation of facts on tax returns (Sikka 2010).

Sikka finds that tax evasion is seen as unethical and illegal and entails illegal deception and concealment (Sikka 2010). That as it may, reducing the tax burden to the lowest level possible by avoiding, not evading, taxes used to be an accepted and reasonable element of corporate tax planning (Hand 1934). Whether tax avoidance is still accepted and reasonable nowadays is debatable. Despite everything that can be said about it, this debate on the morality of tax avoidance falls outside the scope of this thesis. Focusing on the legality, tax avoidance is defined as being legal. However, that is not without any limitations. Tax avoidance is only considered legal when accountants have a good faith belief that the particular tax avoidance measure has a ‘realistic possibility of being sustained administratively or judicially on its merits’ (AICPA 2010). Legal tax issues sometimes create (legal) loopholes. Interpreting these loopholes to the benefit of the taxpayer is not necessarily illegal but is generally seen as more unethical (Holmes 1930). Therefore, taking advantage of these loopholes is sometimes referred to as aggressive tax avoidance (Payne and Raiborn 2018).

For the purpose of this thesis, it doesn’t matter whether a particular company engages in tax avoidance, aggressive tax avoidance or even tax evasion. I focus on the relation between the comparability of financial statements and multiple tax avoidance proxies that indicate whether a company engages in tax avoidance or evasion on general. The particular form of tax avoidance or evasion, the legality of that practice, whether a tax avoidance or evasion practice

is legally challenged or even whether or not the (il)legality of the tax avoidance or evasion practice is proven, is therefore not relevant. Moreover, the legality of a tax avoidance or evasion practice is mostly determined after the fact (ex post), when a particular tax position gets challenged for example. However, not all questionable tax positions get challenged and not all challenged tax positions are truly illegal evasion or avoidance (Wu et al. 2012). Picking only the cases where the IRS challenged tax avoidance practices and imposed some tax audit adjustments or disclosed tax sheltering positions to the public would therefore lead to a biased sample, only including the most aggressive forms of tax avoidance and evasion, that also happen to be detected. Also, IRS data is restricted. I base my thesis therefore on tax avoidance proxies that are based on publicly available information and capture a wide variety of tax avoidance practices.

2.1.2. Relevant factors that influence tax avoidance

Since the Enron accounting scandal in 2001, in which aggressive tax sheltering also played a major role, the United States Treasury Department has focused more and more on the detection, prosecution and punishing of illegal tax avoidance and evasion activities (Frank et al. 2009). Likewise, the public, journalists (Murray 2002) and citizens (McIntyre et al. 2011) focus more on (corporate) tax compliance. However, corporate tax avoidance is influenced by many factors more than only public and governmental scrutiny. Firm specific characteristics, the statutory tax rate, the probability of detection and punishment, the company's risk appetite and the awareness of the corporate civic duty are just a few other important factors (Allignham and Sandmo 1972; Hanlon and Heitzman 2010; Kubick et al. 2017). Rego (2003) finds that the more a company is internationally operative, the more opportunities for tax avoidance are present. Rego shows that this on average results in lower effective tax rates. Separation of ownership and control also influences corporate tax compliance behavior, because of rent-seeking motives by the agent. An agent is more likely to try to aggressively reduce a company's tax liabilities (Crocker and Slemrod 2005; Desai et al. 2007). Other managerial incentives and corporate governance structures can also influence tax avoidance and evasion (Desai et al. 2008; Chen et al. 2010; Badertscher et al. 2013; Chi et al. 2017). Moreover, Hoi et al. (2013) show that socially irresponsible corporate activities correlate with more aggressive tax avoidance and behavior. Irresponsible corporate activities are defined as actions that are widely regarded as damaging to corporate governance, employee relations, communities, public health, human rights, diversity and the environment. Tax avoidance is also facilitated by incentive compensation (Desai and Dharmapala 2006; Armstrong et al. 2015), investor-level taxes (Amiram et al. 2017) and individual top executives (Dyreng et al. 2010). Besides, tax avoidance can be influenced by the geographic location of earnings disclosures (e.g. Hope et al. 2013; Leung and Verriest 2014).

Zimmerman (1983) argues that firm size is positively related with the effective tax rate, because of the higher political costs for bigger firms. Zimmerman argues that bigger companies look more like monopolists and therefore are likely targets of self-interested politicians. These politicians are hypothesized to place 'political costs' on these bigger companies in a (for their own) wealth pursuing effort, such as taxes and regulations. And indeed, Zimmerman shows that larger companies have a higher effective tax rate on average. This indicates that these companies participate less in tax avoidance practices. On the other hand, Wilson (2009) shows that tax sheltering firms are on average larger than the other firms in the same industry, which is consistent with the political power hypothesis. This political power hypothesis posits that bigger firms have more impact on their (social) environment, are better able to lobby the government in their favor and have access to better tax planning resources. Therefore, these bigger firms are (better) able to reduce their effective tax rate (Siegfried, 1972). My results support the political cost hypothesis.

Some literature focusses on the relation between financial reporting and tax-planning (see for example Shackelford and Shevlin 2001). This relation is interesting because financial reporting incentives and incentives regarding the reporting of financial information for tax purposes may collide. Firms with incentives to manage their earnings are found to have greater differences between their book and taxable income (Mills and Newberry 2001). Also, a strong positive relation is found between aggressive tax avoidance and aggressive financial reporting (Frank et al. 2009). This indicates that firms that manage earnings and aggressively avoid taxes report book income upward and taxable income downward. This increases information uncertainty and results in a poorer information environment (Frank et al. 2009; Suk and Zhao 2017). Others argue that information uncertainty and the quality of the information environment are not only affected by, but also affect corporate tax avoidance. That is why multiple studies call for further research on the role of information uncertainty and the information environment in tax avoidance (Maydew 2001; Hanlon and Heitzman 2010). For example, Gallemore and Labro (2015) look at the importance of the *internal* information environment in tax avoidance. They argue that a high quality of internal information enables the more effective and better identification of transactions that easily generate tax revenues. The effect that the information environment has on the relation between financial statement comparability and tax avoidance is, as said in the introduction, relatively unexplored.

Despite the academic interest in in tax avoidance and evasion, conceptualizing and measuring these abstracts is not easy (Gordon and Gallery 2012). Two widely used proxies for tax avoidance are the effective tax rate (ETR) and the Wilson (2009) tax sheltering probability model. The effective tax rate, which captures many aspects of tax avoidance activities, is used

by for example Gupta and Newberry (1997); Dyreng et al. (2008); Rego and Wilson (2012); Armstrong et al. (2015); Kubick et al. (2016) and Suk and Zhao (2017). The effective tax rate measures what percentage of its (book) income a company effectively pays as taxes. By comparing this effective tax rate with the statutory tax rate, one might get an indication of whether and by how much a company engages in tax avoidance practices. An effective tax rate that is much lower than the statutory tax rate could indicate that a company avoids taxes.

Using a company's effective tax rate(s) as proxy for tax sheltering is, as I already mentioned, a useful and arguably better alternative to focusing only on the cases where the IRS imposed financial statement adjustments or disclosed tax sheltering positions to the public. There is also no need for restricted data with the tax sheltering probability model developed by Wilson in 2009. This model only requires publicly available financial statement information. With this model the use of tax sheltering techniques by a particular company can be predicted. Wilson's model is used by for example Kim et al. (2011); Rego and Wilson (2012); Hoi et al. (2013); Hasan et al. (2017) and Suk and Zhao (2017). Wilson (2009) and Lisowsky (2010) show that firms that engage in tax sheltering have larger book-tax differences, more subsidiaries in tax havens, higher effective tax rates, more foreign operations, greater litigation costs and losses, as well as less leverage. Among others, these characteristics are part of the develop tax sheltering probability model.

2.2. Financial statement comparability

Financial statement comparability is seen as an aspect of useful financial information (FASB 2010)[2]. Financial statement comparability is found to improve the external information environment of a firm. This enables users to make better inferences with regard to the economic similarities and differences between firms, which aids investors and managers in investment decisions and policy making (FASB 1980; De Franco et al. 2011; Kim et al. 2013; Choi et al. 2015; Kim et al. 2016). Financial statement comparability for example positively correlates with forecast accuracy (De Franco et al. 2011) and reduces uncertainty about the pricing of firms' credit risk (Kim et al. 2013). Choi et al. (2015) find that more comparable financial statements allow investors to better anticipate and forecast future firm performance. Further, Kim et al. (2016) find that better corporate comparability encourages managers to share not only 'the good news' but also 'the bad news' with the firm's stakeholders: comparability discourages bad news hoarding. This reduces the investors' unfair perceptions of the firm's future earnings and risk. Also, firms make more profitable mergers and acquisitions when the financial statements of the target firm are more comparable (Chen et al.

[2] Other characteristics of useful financial information are timeliness, verifiability and understandability (FASB 2010).

2017). As such, these studies support the hypothesis that financial statement comparability lowers the cost of information acquisition and information processing, by providing more objective information and improving the quality (and the quantity) of available information. This of course also applies to tax authorities.

Several studies have tried to develop a good way to measure financial statement comparability. This research can be divided into two groups. The first group are the output-based measures: measures that are based on firms' financial statement information and the market reaction to that information. De Franco et al. (2011) investigate how stock prices (and returns) react to earnings information. Barth et al. (2012) add (prior) stock returns and cash flows to the tests of De Franco et al. (2011). Francis et al. (2014) use a firm's earnings covariation and the closeness of its accruals to industry peers to measure the comparability of financial statements. Such output-based measures are not based on some subjective selection or on the weighing of accounting inputs. Therefore, they are less subject to endogeneity issues (De Franco et al. 2011). They do require the use of time-series earnings data and stock returns data however (Hoitash et al. 2017). Nevertheless, Gross and Perotti (2017) find four advantages of output-based comparability measures: 1) they're more relevant for users, 2) they're more objective, 3) they're easier to implement, 4) they're potentially more accurate. The other group of measures for financial statement comparability are the input-based measures: measures that focus on the similarity of accounting methods and policies of a firm and its peers (for example: DeFond and Hung 2003; Bradshaw et al. 2009; Kim et al. 2013). Peterson et al. (2015) were the first to test a statistical model based on this kind of policy inputs. They focus on textual similarities in accounting policy disclosures and find that financial statement comparability is likely to be correlated with other earnings characteristics, such as persistence, predictability, and accrual quality. Input-based measures do require the (subjective) selection and weighing of data (Hoitash and Hoitash 2017).

Recently however, a new and arguably better method to measure financial statement comparability has been found. This is an XBRL-based measure, developed by Hoitash et al. (2017). XBRL is a way of standardizing accounting terminology, using tags that are prescribed by US GAAP. The tags required for this measure are mandatorily reported to the SEC every year since 2009. The XBRL-based measure is objective and comprehensive: it is based only but on all the monetary accounting information reported in the financial statements and does not require any market data (Hoitash et al. 2017). This new measure is thus essentially an input-based measure. However, it is an improvement compared to the Peterson et al. (2015) model since it isn't sensitive to semantics (XBRL is a standardized terminology) and, as opposed to the other input-based comparability measures, the location of the accounting

disclosures in a financial statement does not matter. Furthermore, no subjective selection and weighing of the data is needed (Hoitash et al. 2017). The evidence it provides on the implications of financial statement comparability and financial reporting behavior (e.g. tax shifting) is therefore clearer (Hoitash and Hoitash 2017 and Hoitash et al. 2017). Hoitash and Hoitash (2017) and Hoitash et al. (2017) further state that the XBRL-based measure allows for bigger samples to be analyzed and that it is developed with the explicit goal of determining the effect of financial statement comparability on managerial financial reporting behavior and tax avoidance. In that context Hoitash and Hoitash (2017) find that this measure is superior to the other existing input- and output-based measures. I therefore use my own version of this arguably better measure of financial statement comparability to analyze the relation between financial statement comparability and tax avoidance, which was not measured before using a XBRL-based comparability measure.

3. Hypothesis Development

(Aggressive) tax avoiding strategies often entail an elaborate and complex structure of transactions (Goh et al., 2016). Furthermore, disclosures in filed financial statements with regard to taxable income or payable tax are often inadequate to detect legally punishable tax avoidance practices (Bozanic et al., 2017). Hence there is great amount of information asymmetry between tax authorities and firm managers with regard to transactions structured to avoid taxes (Mills and Plesko, 2003). This tax-relevant information deficit means the role of financial statements, particularly the financial statements of peer companies, in detecting tax avoidance or clarifying private information by tax authorities is big (Bozanic et al., 2017). Financial statement information from comparable firms can namely serve as a (partly) substitute for, or as a supplement to, a peer company's own financial statement information. This increases the amount of tax relevant information the tax authorities have (De Franco et al. 2011). Better comparability of financial statements thus enlarges the amount of tax-relevant information tax authorities have. That enables the tax authorities to better understand and evaluate the tax-relevant transactions of a particular company (Bozanic et al., 2017).

For example: the internal transfer prices a company charges to its subsidiaries, parent or sister firm(s) need to be very well understood by tax authorities. These transfer prices are fiscally required to be 'at arm's length' (see Article 9 of the OECD Model Tax Convention), since these transfer prices would otherwise be an easy way to shift and manage profits (Choi et al., 2017). With the right information, the tax authorities could base the allowed transfer prices partly on peer information. The more comparable the financial statements are, the easier it is for tax authorities to evaluate which companies are truly peers and the easier it thus is for tax

authorities to get an indication of a reasonable at arm's length transfer price. In general: the more comparable the financial statements are, the easier it is for tax authorities to evaluate the financial and tax-relevant information of companies. The comparability of financial statements therefore also results in lower costs of (tax-relevant) information acquiring and processing by tax authorities (Suk and Zhao, 2017). This facilitates tax authorities in detecting tax evasion and avoidance practices and makes it more difficult for managers to hide and/or justify structures and transactions aimed at tax avoidance and evasion. Put in other words: better comparability of financial statements increases the risk of tax avoidance practices being detected. The (expected) costs of tax avoidance therefore increase. Basic behavioral-economic theories argue that managers would anticipate this, which would result in managers of firms with more comparable financial statements being less inclined to (aggressively) avoid taxes. Financial statement comparability therefore would result in the decrease of tax avoidance. This is the reasoning of Suk & Zhao (2017).

The reasoning and logic in the study of De Simone (2015) is different. De Simone argues that more comparable financial statement information results in a bigger bandwidth of, for example, reasonable arm's length transfer prices, since more information of peers is available. Sticking to that example: the comparability then results in an increased amount of leeway a manager has in determining the arm's length transfer prices and thereby shift profits and avoid taxes. The detection of tax avoidance or evasion practices also decreases at the same time because of the bigger bandwidth of reasonable arm's length transfer prices. Overall, this results in an increased incentive for managers to engage in tax avoidance en evasion practices. De Simone thus argues that financial statement comparability results in the increase of tax avoidance.

De Simone bases her hypotheses on the assumption that the transfer prices used in the financial statements of peers are very different and lie far apart from each other. Improved financial statement comparability would then lead to the tax authorities taking more of those (very) different transfer prices into account, which in turn would lead to the argued increased bandwidth of reasonable arm's length prices. I cannot find any (scientific) basis for this assumption, neither for the opposite by the way. Basic economic theory cannot help here either, because tax avoidance is, as I described, influenced by many factors, not only risk aversion for example. I can thus only assume, given the many factors that influence tax avoidance (incentives), that companies on average try to prevent tax avoidance or the corresponding incentive to a certain level, out of fear of detection. Hoopes et al. (2012) indeed find that managers engage less in tax avoidance practices if the chance of an IRS audit is higher. Kubick et al. (2016) show that firms indeed decrease their tax avoidance practices after they received a comment from the SEC related to their tax avoidance. I therefore assume that

managers, if the incentive to avoid taxes is present, try to avoid taxes as inconspicuously as reasonable. To do this, they might want to choose a transfer price that allows for (a bit of) tax avoidance but that does not lie very far apart from the 'true' at arm's length transfer price, so it remains (partly) justifiable, with variations depending on all the factors that influence tax avoidance and the chance that the tax avoidance or evasion is detected. This would create a so-called 'gray-area' of more and less justifiable transfer prices, with of course a few outliers. If financial statements then become more comparable, I expect the gray-area to get narrower. That is, the improved comparability allows tax authorities to take more used transfer prices into account and so pinpoint the most commonly used but also the 'true' at arm's length transfer price(s) more easily. That decreases the width of the 'gray-area' with justifiable transfer prices. Managers know this. The managers who were at the outer spectrum of the 'gray-area' might, because of the increased chance and fear of detection, decide to decrease the difference between the transfer price they use and the 'true' at arm's length transfer price, ergo: choose to decrease their tax avoidance and evasion practices. I therefore think the reasoning of Suk and Zhao (2017) is more compelling. Hence, I expect that financial statement comparability has a reducing effect on (aggressive) tax avoidance or evasion. My first hypothesis thus is:
H1*: Financial statement comparability deters a firm from engaging in aggressive tax avoidance, holding all else constant.*

Answering to the called for research on the relation between the information environment and tax avoidance, I examine the effect the information environment has on the correlation between financial statement comparability and tax avoidance. A small increase in comparability can namely be moderated by a firm's information environment. In case of very transparent public information, more financial information of comparable firms may not be that important and/or needed, tax-relevant anomalies, for example very deviant internal transfer prices, could also be detect with the available public information (Suk and Zhao 2017). Transparent public information can be achieved by for example 1) better, more or extra financial statement information filed or published by the firm itself, or 2) if the firm is cross-listed on multiple capital markets or 3) when the firm is followed by more (fiscal) analysts who share their findings (Fernandes and Ferreira 2008; Amel-Zadeh and Della Bina 2017). However, vice versa: the less transparent the information environment is, the more valuable the financial information of comparable firms. This opaque information environment might therefore pronounce the effect financial statement comparability has on tax avoidance. Taken all this into account, I predict that the decreasing effect of financial statement comparability on tax avoidance will be greater for firms with a less transparent (more opaque) information environment. Hence, my second hypothesis:

***H2**: Financial statement comparability has a more discouraging effect on aggressive tax avoidance in the case of a poor information environment.*

4. Research design

4.1. Tax avoidance

Following prior literature (see for example Suk and Zhao 2017), my proxies for tax avoidance are a companies' effective tax rates and the tax avoidance probability estimations of Wilson's (2009) tax avoidance probability model. Using both these measures as a proxy for (aggressive) tax avoidance, I show my findings are to some extent robust.

Effective tax rates

The effective tax rates I use are the GAAP effective tax rate (GAAP_ETR) and cash effective tax rate (CASH_ETR).[3] The GAAP_ETR is defined as the total tax expense (TXT) per dollar of book income before tax (PI): $GAAP\ ETR = \frac{TXT}{PI}$. The CASH_ETR is defined as the paid tax expense (TXPD), divided by the book income before tax (PI) minus special items (SPI): $Cash\ ETR = \frac{TXPD}{(PI-SPI)}$. Special items are deducted from the book income before tax because they can introduce unwanted volatility in the effective tax rates due to their possible (large) magnitude. This problem is most pressing when the numerator is the paid tax expense (Suk and Zhao 2017), so special items are only deducted in that calculation. I then winsorize the effective tax rates to [0, 1], because this is standard practice in most literature. Also, negative effective tax rates and tax rates above 1 are non-meaningful and could obscure the inferences about a firm's aggressive tax avoidance behavior (Hoi et al. 2013; Gallemore and Labro 2015; Suk and Zhao 2017).

Wilson's (2009) tax sheltering probability estimations

Wilson's (2009) tax sheltering probability model is a model that estimates the probability that a particular firm participates in tax avoidance, based on certain business characteristics of that particular firm. Wilson (2009) divided this model into two different equations, that I for robustness purposes will both use in this thesis. The first model developed by Wilson (2009) is reflected by the following equation:

$$P_{shelter,\ 1} = -4.3 + 6.63BTD_{it} - 1.72LEV_{it} + 0.66\log(AT)_{it} + 2.26ROA_{it} + 1.62FOREIGN_{it} + 1.56R\&D_{it}$$

[3] Two instead of one, again for robustness purposes.

Prior literature found discretionary accruals as derived with the modified Jones Model (as used by Dechow et al. 1995) to be positively related to aggressive tax avoidance (Frank et al. 2009). So Wilson (2009) added these discretionary accruals as a control variable in his second probability equation:

$$P_{shelter,\ 2} = -4.86 + 5.20BTD_{it} + 4.08DA_{it} - 1.41LEV_{it} + 0.76\log(AT)_{it} + 3.51ROA_{it} + 1.72FOREIGN_{it} + 2.43R\&D_{it}$$

BTD is calculated as the book income before tax (PI) minus taxable income, scaled by the companies lagged assets. Taxable income is defined as the sum of current federal tax expense (TXFED) and current foreign tax expense (TXFO), divided by the statutory tax rate (STR), after which the change in net operating loss carryforwards (TLCF) is subtracted.[4] If current federal tax expense is not available, the taxable income is calculated by subtracting deferred taxes (TXDI), state income taxes (TXS) and the other income taxes (TXO) from total income taxes (TXT).[5] DA are the discretionary accruals derived from the modified Jones Model as used by Dechow et al. (1995), scaled by lagged assets. LEV is the leverage of the firm; log (AT) is the logarithm of total assets; ROA are the return on assets, calculated as the pretax earnings scaled by total assets; FOREIGN is a dummy variable that equals 1 for firm-years that show foreign income and 0 otherwise and at last R&D are the research & development costs scaled by total assets.

After calculating both these tax sheltering probabilities, the aggressive tax avoidance proxy is obtained the same way as in prior studies (Rego and Wilson 2012; Hoi et al. 2013; Hasan et al. 2017). I rank $P_{shelter,\ a}$ by year and create an indicator (dummy) variable that is 1 for the firms that have the highest sheltering probability. That is, this indicator variable is 1 if a firm's estimated $P_{shelter,\ a}$ is in the top quintile of the sheltering probability distribution of that year. This binary measure means that the model estimating the effect of financial statement comparability on (aggressive) tax avoidance must be a logistic.

4.2. Financial statement comparability

As proxy for financial statement comparability, I use an interpretation of the aforementioned XBRL-based measure of Hoitash et al. (2017). I construct my XBRL-based measure as follows. First, I sort all reported XBRL-tags by year, not including the restatements of financial statements. Then, I delete all duplicate firm-XBRL-tag observations. This yields a yearly

[4] $BTD = \frac{PI-\left(\left(\frac{TXFED+TXFO}{STR}\right)-TLCF\right)}{lagged\ assets}$

[5] $BTD = \frac{PI-(TXT-(TXDI+TXS+TXO))}{lagged\ assets}$

database of all the XBRL-tags all the companies individually reported. Then, in each fiscal year and within each two-digit SIC industry, I determine how many times all the different distinguishable XBRL-tags have been reported in the two-digit SIC industry as a whole. The number of times a single XBRL-tag is reported in that two-digit SIC industry is the '*Count*'. If the tag is only reported by one company, the *Count* is 1. If all X companies in the two-digit SIC industry reported the XBRL-tag, the *Count* is X. With that *Count*, I calculate the 'T*ag Value'*. The *Tag Value* is calculated by dividing 1 by the *Count*. The *Tag Value* therefore differs for the same tag in different two-digit SIC industries. A really low value of the *Tag Value* indicates that the particular tag was reported many times in that two-digit SIC industry, a very high value the opposite. A *Tag Value* of one means that only 1 company reported the particular tag. Each XBRL-tag in the database is then assigned its own tag value. After that, I sum all the tag values of the individual companies in that two-digit SIC industry, yielding a '*Total Tag Value*' per individual company. A low *Total Tag Value* means that many of the tags that a single company reported were also reported by the other companies in that two-digit SIC industry. Two-digit SIC industries where the *Total Tag Value* is on average low therefore suggest more comparable financial statements, and vice versa of course. However, the total amount of tags each individual company reported also needs to be taken into account. [6] To do this I divide the *Total Tag Value* of the individual company by the number of tags that company reported, resulting in the '*Weighted Total Tag Value'*. At last, to make the results more intuitive, I divide 1 by the *Weighted Total Tag Value*, resulting in the '*Financial Statement Comparability Score' (FSC-score)*. A higher (lower) FSC-score implies that the particular company reported more (less) comparable financial statements with respect to its two-digit SIC industry peers. Intuitively, a two-digit SIC industry where the FSC-score scores are (on average) higher is more comparable than a two-digit SIC industry where the FSC-score scores are (on average) lower. The absolute value of FSC-score is not informative nor relevant. The information is provided by the relative difference between the absolute values: whether one value is higher or lower than another and by how much lower or higher.

One other important thing to take into consideration is the fact that the FSC-score might be endogenous, which would lead to endogeneity concerns with respect to the regression results. The choice of financial statement disclosures and the disclosure format, relative to industry peers, need not to be random for a company. Therefore, the FSC-score as well as other confounding factors might affect both the financial statement comparability and the amount of

[6] If a particular company only reports 5 tags and if those tags were also reported by all the other companies in the relevant two-digit SIC industry, the company that only reported 5 tags would have a very low (maybe the lowest) financial statement comparability score. By also dividing the total tag value by the number of tags each company reported (5 in this particular case), I take into account that reporting fewer tags would otherwise result in a lower total tag value, which would falsely suggest more comparable reported tags of that particular company.

(aggressive) tax avoidance a firm participates in. Also, aggressive tax avoidance might result in a more opaque external information environment because of the complexity of the tax avoidance and evasions related activities (Balakrishnan et al., 2018). This could influence the effect that financial statement comparability has on tax avoidance. Although there is controlled for a lot of different factors in the regression equations given below, and year and industry fixed effects are added, this might not cover the whole problem. I therefore also try to eliminate all (these) possible endogeneity concerns by using a two-stage regression analysis, with the geographical distance to industry peers as the instrumental variable for each firm. More about this two-stage regression analysis in paragraph 7.

4.3 Sample selection and data collection

I base my empirical analysis on the XBRL-tags that companies report in their SEC filings. I use the historical archives of the SEC to gather these XBRL-tags. These downloadable databases contain, unchanged, 'as filed' and on a quarterly basis, all filings the SEC received (e.g.: yearly (10-K), quarterly (10-Q), restatements (10-K/A and 10-Q/A) etc.).[7] I only focus on 10-K filings, because these are the most comprehensive overview of a firm's financial positions and results. Furthermore, in a 10-K filing all the reported XBRL-tags cover the same fiscal period: one fiscal year. I do not include financial restatements because these mostly only occur to fix material errors in earlier reportings and mostly result in just the changing of financial figures (Segal et al. 2006). Hence, I assume that restatements would only rarely result in adding or removing particular XBRL-tags. Not including the restatements does therefore not significantly result in biases regarding the reported XBRL-tags.[8] Further, I drop all companies that do not have their main business address in the United States, to prevent biases from different country regulations, (fiscal) exceptions and other possibly interacting effects. This allows for clearer intercompany comparisons (less noise).

The sample starts with data from the second quarter of 2009 and ends with the first quarter of 2018. The sample starts in 2009 because the SEC adopted rules on the 30th of January 2009 that required companies to provide their financial statements in an XBRL-format to the SEC as of April 13, 2009 (SEC 2009). Moreover, before the second quarter of 2009, reporting the XBRL-tags wasn't mandatory, only optional. Including also those optional reportings in the sample could lead to major endogeneity concerns.

[7] See *The Financial Statement and Notes Data Sets* downloadable at https://www.sec.gov/dera/data/financial-statement-and-notes-data-set.html.
[8] So I assume that in case of a restatement, the amount and specifically reported XBRL tags will not change, only the given values.

After this sampling, my XBRL-tag database contains more than 52 million reported tags.[9] Following prior literature on tax avoidance I exclude utility companies (SIC codes 4900 – 4949) and companies in the financial sector (SIC codes 6000 – 6999), because these firms likely face more regulatory scrutiny than other firms and these firms might also face different reporting incentives.[10] For each tax avoidance measure (CASH_ETR, GAAP_ETR, SHEL1 and SHEL2) I use an own sample, as did Suk & Zhao (2017). This way, I keep as much firm-year observations as possible for each proxy. This also results in the proxies being unbiased and independent robustness tests for each other. After requiring all regressions variables to be available, the sample contains 5.850 firm-year observations for the CASH_ETR-sample, 7.385 firm-year observations for the GAAP_ETR-sample, 7.456 firm-year observations for the SHEl1-sample and 7.464 firm-year observations for the SHEl2-sample. The only variations in firm-year observations per proxy across the calculations are caused by the adding of extra variables (such as the bid-ask spread), which are not defined for all the firm-year observations that the samples contain. The samples further do not differ from the samples used in the analysis of hypothesis 1. I calculate the proxies for aggressive tax avoidance (the effective tax rates and Wilson's (2009) sheltering probabilities) using Compustat and CRSP (monthly) data. The control variables are also calculated using Compustat and CRSP (monthly) data. I winsorize all the continuous variables at the top 1% and bottom 1% to account for outliers.

5. Empirical Analysis

Hypothesis 1

Regarding the first hypothesis, I test the relation between financial statement comparability and aggressive tax avoidance. I argue that financial statement comparability and aggressive tax avoidance are negatively correlated. I test this by using the following regression model (based on Suk & Zhao 2017):

Model 1

$$TA_{it} = a_1 + \beta_1 FSC_{it-1} + \beta_2 \Delta TLCF_{it} + \beta_3 NOL_{it-1} + \beta_4 ROA_{it-1} + \beta_5 LEV_{it-1} + \beta_6 PPE_{it-1} + \beta_7 INTAN_{it-1} + \beta_8 R\&D_{it-1} + \beta_9 ESUB_{it-1} + \beta_{10} PIFO_{it-1} + \beta_{11} MB_{it-1} + \beta_{12} \Delta SALE_{it} + \beta_{13} SIZE_{it-1} + \beta_{14} EMP_{it-1} + \beta_{15} CASH_{it-1} + \beta_{16} DA_{it-1} + \beta_{17} TA_{it-1} + Year\ Fixed + Industry\ Fixed + \varepsilon_{it}$$

TA is one of the aforementioned proxies for tax avoidance: GAAP_ETR, CASH_ETR, SHEL1 or SHEL2. I use an OLS regression when (aggressive) tax avoidance is proxied by GAAP_ETR

[9] 52.617.014 XBRL-tag observations to be exact.
[10] This results in dropping 8576 firm-year observations.

or CASH_ETR. I use a logit regression when I measure aggressive tax avoidance with the Wilson's (2009) tax sheltering probability scores.[11]

For H1 I expect the FSC-score coefficient to be positive when (aggressive) tax avoidance is measured with GAAP_ETR and CASH_ETR. I expect the FSC-score coefficient to be negative however when (aggressive) tax avoidance is measured using the two Wilson (2009) sheltering probability scores. A higher effective tax rate and a lower tax sheltering probability namely both indicate less aggressive tax avoidance.

In this model, I include control variables based on Zimmerman (1983); Gupta and Newberry (1997); Phillips et al. (2003); Graham and Tucker (2006); Wilson (2009); Frank et al. (2009); Hoi et al. (2013) and Hasan et al. (2017). I control for the change in loss carry forward (ΔTCLF), combined with a loss carry-forward indicator (NOL). I further control for return on assets (ROA) since less taxable income tends to result in lower tax-avoidance incentives. Leverage (LEV) is also included since interest payments work as a tax shields and therefore result in less aggressive tax avoidance incentives. I include property, plant and equipment (PPE), because firms that are more capital intensive (have a higher PPE) tend to have more depreciation expenses, which leads to a lower tax burden and a higher book-tax difference[12], which in turn results in less aggressive tax avoidance incentives. I also include intangible assets (INTAN as well as R&D) and equity income of subsidiaries (ESUB), because the difference in tax treatment of intangible assets and equity income between the equity and cost method might result in tax deficiencies, which would lead to less meaningful tax avoidance proxies. I also control for internationally operating firms (PIFO), because they have more opportunities to shift their income between countries and jurisdictions (with different tax rates). I further control for growth opportunities (measured by the market to book ratio (MB) and the change in sales (ΔSALE)), since (rapidly) growing firms tend invest more in tax-favored assets. I also account for the number of employees (EMP) of a firm and the firm's size (SIZE), and I control for the holding ratio (CASH) following Hasan et al. (2017). I further include the discretionary accruals based on the modified jones-model (DA), because these discretionary accruals have a protrusive effect on aggressive tax avoidance (Phillips 2003). I also add last year's tax avoidance, because the tax avoidance of a company seems to be consistent over time, meaning the lagged and current tax avoidance measures are correlated (Suk & Zhao, 2017). At last I include year and industry fixed effects, so that these account for macroeconomic effects due to regulations (tax or disclosure-related) and differences in tax avoidance between industries. For detailed definitions of all the control variables, **see Appendix B**.

[11] Because I divide the sample in quintiles and assign only a one to the top quintile of observations. This makes these proxies of tax avoidance binary, which requires a logistic regression.

[12] A bigger difference between the accounting income and the (estimated) taxable income.

Hypothesis 2

With my second hypothesis I consider the effect the information environment of a firm has on the relation between financial statement comparability and (aggressive) tax avoidance. I predict that the comparability of financial statements has a bigger impact for firms that have an opaque information environment, ergo: tax avoidance decreases further for firms in an opaque information environment then for firms in a transparent information environment given a certain increase in financial statement comparability. For this second hypothesis I use two measures to capture the transparency of a firm's information environment (IA). Larger firms are more likely to have better information environments, as they are more mature, have more established policies and are generally more heavily scrutinized (Vermaelen 1981; Ozkan and Ozkan 2004; Drobetz et al. 2010). So firm size is one of the proxies for the quality of the information environment, with a bigger size indicating a more transparent information environment. Second, I use the bid-ask spread as proxy for the quality of the information environment. I calculate the bid-ask spread as the average of the monthly bid-ask spreads over the past 12 months. Research shows that firms with higher bid-ask spreads are characterized by a bigger information asymmetry and thus poorer (more opaque) information environments (Armstrong et al. 2011; Huddart and Ke 2007). All the firm-year observations are, by year, classified in 5 groups based on these proxies. Firms in the first quintile regarding the size (the smallest firms) are assigned a 1, which means that these firms have a high information opacity: a poor information environment (IA = 1). The bigger firms in the other 4 quintiles are assigned a 0, meaning these firms are more information transparent (IA = 0). The same is done for the bid-ask-spread. Firms in the fifth quintile of the bid-ask spreads are assigned a 1, meaning these firms have a high information opacity: a poor information environment (IA = 1). The other firms are assigned a 0, meaning these firms are more information transparent (IA = 0). I test the second hypothesis with the following regression model (based on Suk & Zhao 2017):

Model 2

$$TA_{it} = a_1 + \beta_1 FSC_{it-1} + \beta_2 \mathrm{IA}_{it-1} + \beta_3 IA_{it-1} * FSC_{it-1} + \sum \beta_i Controls_{it-1} + Year\ Fixed + Industry\ Fixed + \varepsilon_{it}$$

By design, IA equals 1 if the firm has a poor (opaque) information environment and zero otherwise. The controls in this regression are the same as in the regression used in the analysis of the first hypothesis. I expect the coefficient of the interaction variable (IA*FSC-score) to be positive for the effective tax rate based measures (GAAP_ETR and CASH_ETR) and negative for the Wilson's (2009) tax sheltering probabilities, because I argue that an

opaque information environment results in a bigger (more pronounced) effect of the comparability of financial statements on tax avoidance: a bigger increase in the effective tax rate and a bigger decrease in the tax sheltering probability.

I refer to the part of the analysis of hypothesis 2 where I use firm size as a proxy for the information environment as hypothesis 2a. For the part of the analysis where I use the bid-ask spread as proxy for the information environment, I merge the different samples of all the tax avoidance proxies with the CRSP database to get the required data. This results in a different (lower) number of available firm-year observations for the second part of the analysis. I therefore refer to this part of the analysis as hypothesis 2b.

6. Results

In tables 1 to 4 I provide the summary statistics of the variables used in the analysis of hypothesis 1. In tables 6 to 9 I present the summary statistics of the variables used in the analysis of hypothesis 2a. In tables 12 – 15 I list the summary statistics of the variables used in the analysis of hypothesis 2b. In tables 3, 4, 8, 9, 14 and 15 I also present the shelter1 and shelter2 values, on which the SHEL1 and SHEL2 indicator variables are based. These summary statistics seem not to be contradicted by prior literature (De Simone 2016; Suk and Zhao 2017). The mean values for GAAP_ETR, CASH_ETR, SHEL1 and SHEL2 for hypothesis 1 are 0.192; 0.219; 0.235 and 0.234 respectively. For hypothesis 2a and 2b the mean values are 0.194; 0.221; 0.234; 0.232 and 0.213; 0.250; 0.181 and 0.183 respectively. The reported standard deviations indicate, when compared to the corresponding means, substantial but not unreasonable variations of all tax avoidance measures across all firm-year observations, and are in line with prior literature (Suk and Zhao 2017). I also report the summary statistics for the financial statement comparability measure (lag_lfsc being the FSC-score variable) and the control variables of the regressions. The value of the measure for financial statement comparability differs from the values of comparability found in prior literature. This is not surprising because I'm the first to compute the financial statement comparability using an interpretation of the new XBRL-based method. Moreover, the absolute value of the mean of the FSC-score variable is, as I already stated, not informative. I further report the lagged values of all control variables; the lagged values are after all used in the analysis. The not reported non-lagged variable means are all very similar to those reported by prior studies (De Simone 2016; Suk and Zhao 2017). The reported lagged variable means itself are also quite similar to the not reported non-lagged control variable means in this study and in prior studies. Actually only the mean value of equity income in earnings (ESUB; my variable: lag_esub) differs from, is bigger than, the mean values found in prior studies. I attribute this to the fact that I reported the lagged values of the variables. The not reported non-lagged mean value of equity income

in earnings is after all comparable to the mean values for this variable found in prior studies. Furthermore, lagged variables depend, even more than non-lagged variables, on data availability: not for all non-lagged variables is a lagged variable available.

6.1. H1: Financial statement comparability deters a firm from engaging in aggressive tax avoidance, holding all else constant.

The results of the regression analysis of Model 1 are shown in table 5. Columns 1 and 2 present the results for the regressions where GAAP_ETR and Cash_ETR are used as the proxy for financial statement comparability. Columns 3 and 4 present the results from the logistic regressions using SHEL1 and SHEL2 as financial statement comparability proxies.

For GAAP_ETR and Cash_ETR, the estimated coefficients of the financial statement comparability are positive and significant at the 1% level. The estimated coefficient of the FSC-score variable (lag_lfsc in the tables) for the GAAP_ETR proxy is 0.003 and for the Cash_ETR proxy 0.004.[13] These coefficients are economically significant. One standard deviation increase in the financial statement comparability increases the effective tax rate by 0.0080 and 0.0090 for Cash_ETR and GAAP_ETR respectively. This translates into an increase of the effective tax rate by 5,86% and 7.01%[14] on average, respectively, for a one standard deviation increase in financial statement comparability. The role that financial statement comparability plays in reducing aggressive tax avoidance is thus substantial.

When SHEL1 and SHEL2 are used as proxies for tax sheltering, the coefficients of financial statement comparability are both negative and statistically significant at the 1% respectively 5% level. The value of the FSC-score coefficient of the SHEL1 regression (-0.046[15]) translates into an odds ratio of 0.955[16]. This result indicates that for a one standard deviation increase in financial statement comparability, there would be an 11.3%[17] decrease in the odds that a firm attempts to shelter taxes. The FSC-score coefficient is also economically significant. The results indicate that the marginal effect that financial statement comparability has on the SHEL1 value is on average 1.09 percentage points. This indicates that a one standard deviation increase in financial statement comparability reduces the probability of a firm sheltering tax on average by 2,7%.[18] The same holds for the -0.041[19] coefficient for the FSC-

[13] 0.00314 and 0.00359 to be exact.
[14] NOT percentage points.
[15] -0.0463501 to be exact.
[16] e^-0.0463501; 0.95470766 to be exact.
[17] 1 – (0.95470766 ^ std dev of the FSC-score coefficient for the SHEL1 sample (2.595)).
[18] Percentage points * std dev of the FSC-score (2.595)
[19] -0.0408928 to be exact.

score variable in the SHEL2 regression. This translates into an odds ratio of 0.960[20]. This result indicates that for a one standard deviation increase in financial statement comparability, there would be a 10.1%[21] decrease in the odds that a firm attempts to shelter taxes. Moreover, these coefficients are also economically significant. The marginal effect of financial statement comparability on SHEL2 is on average 0.98 percentage points. This means that if the financial statement comparability increases by one standard deviation, the probability of sheltering for a given firm decreases by 2,5%. Graphs 1 and 2 visualize these effects.

The results with regard to the control variables are consistent with prior literature. Firm size is for example inversely related with financial statement comparability: when the effective tax rate is used as tax avoidance proxy, the FSC-score coefficient is positive and the effect of firm size is negative. When the sheltering probability indicator is used as proxy, the FSC-score coefficient is negative and the effect of firm size is positive. This means that bigger firms are less likely to shelter taxes and that my results are in line with the political cost hypothesis (see also Davidson & Heaney 2012), and thus support the reasoning of Zimmerman (1983).

6.2. H2: Financial statement comparability has a more discouraging effect on aggressive tax avoidance in case of a poor information environment.

Hypothesis 2a

This second regression model tests what impact a firm's information environment has on the relation between financial statement comparability and aggressive tax avoidance. I predict, based on prior literature, that the comparability of financial statements has a more pronounced effect on tax avoidance in case of a poor information environment. Table 10 shows the results of the regression analyses of the first sub model (H2a), where firm size is the proxy for the information environment. IA = 1 for firms in the quintile of smallest firms, indicating an opaque information environment (H2a(a)). The results of this regression analysis do not show a statistically significant more pronounced correlation between financial statement comparability and aggressive tax avoidance for small firms. The effect of the interaction variables (IA*FSC-score) are, contrary to what I hypothesize, for all columns except column 2 in the opposite direction as the FSC-score variables (0.004; 0.003; -0.048; -0.043 and -0.001; 0.001; 0.012; 0.001 are the coefficients of respectively the FSC-score variable and the interaction variable, for CAHS_ETR, GAAP_ETR, SHEL1 and SHEL2). The results therefore indicate that an opaque information environment results in a less pronounced effect of financial statement comparability: the interaction variable indicates a lower rise of the average tax rates and a

[20] e^-0.0408928; 0.959932029 to be exact.
[21] 1 – (0.959932029 ^ std dev of the FSC-score coefficient for the SHEL1 sample (2.597)).

smaller decrease of the sheltering probabilities for smaller firms if the FSC-score increases. The information environment thus has a mitigating effect on the relation between financial statement comparability and tax avoidance. However, none of the results are statistically significant. The evidence for this effect of the information environment is therefore weak.

Since these results are, regarding the sign, in the opposite direction of what I expected, I test the same model again, but now with IA = 1 for the quintile of biggest firms instead of for the smallest firms (H2a(b)). As I argued before, a bigger firm size proxies for a more transparent information environment. Table 11 presents the results of this regression. Indeed, all the IA*FSC-score coefficients are of the same sign as the FSC-score coefficient itself (0.003; 0.003; -0.042; -0.043 and 0.002; 0.002; -0.294; -0.173 are the coefficients of respectively the FSC-score variable and the interaction variable, for CAHS_ETR, GAAP_ETR, SHEL1 and SHEL2). This indicates that for bigger firms the information environment results in an even bigger rise in effective tax rates and reduces the tax sheltering probability of a company even more, relative to smaller firms. So, a transparent information environment does enhance the effect of financial statement comparability in a transparent information environment.

This also logically makes sense. Bigger firms are argued to have a better, meaning a more transparent, information environment. This results in the tax authorities being better able to 'see through' the company, analyze and evaluate the tax-relevant information and transactions and tax the company accordingly. As such, the incentive to shelter taxes would be lower (because of the fear of detection) and the effective tax rates would be higher. An increasing level of financial statement comparability then only further reduces the incentives to avoid or evade taxes. Moreover, this is in line with the mentioned political costs hypothesis and the results of the analysis of hypothesis 1.

Vice versa, smaller firms are argued to have a poorer and more opaque information environment. Tax authorities thus less easily understand and analyze the tax-relevant information of a firm. The chances of getting caught while avoiding or evading taxes are therefore slimmer in a poor information environment given a certain level of financial statement comparability. The results indeed indicate this: a relatively higher sheltering probability and relatively lower effective tax rate in case of a more opaque information environment. Very cautiously one might therefore say that the dampening effect of financial statement comparability on aggressive tax avoidance is less pronounced in firms with a less transparent information environment and that while financial statement comparability may reduce tax avoidance (see H1), it cannot completely offset the incentive-increasing effect with regard to tax avoidance of an opaque information environment. However, although all these results and

inferences are logically plausible, the results are not significant. As such, no strong conclusions can be drawn from these statistics and inferences.

Hypothesis 2b

I find more or less the same results when I use the bid-ask spread is as proxy for the information environment. See table 16. I start with IA = 1 for the biggest bid-ask spreads, indicating an opaque information environment (H2b(a)). The coefficients of the interaction variable are all insignificant. The effect of the IA*FSC-score variable is further for all regressions in the opposite direction of the FSC-score variables (0.002; 0.003; -0.043; -0.041 and -0.002; -0.001; 0.070; 0.045 are the coefficients of respectively the FSC-score variable and the interaction variable, for CAHS_ETR, GAAP_ETR, SHEL1 and SHEL2). The results thus indicate that a bigger bid-ask spread (more opaque information environment – more information asymmetry) mitigates the increasing effect with regard to the effective tax rate and the decreasing effect with regard to the sheltering probabilities of financial statement comparability. This can be explained with the same logic as above: bigger bid-ask spreads indicate a more opaque, less transparent information environment. Tax authorities thus have less information and less easily understand the company and its tax-relevant transactions. The chances of getting caught while avoiding taxes are therefore smaller compared to a more transparent information environment. This leads to an increased tax avoidance incentive. A more opaque information environment therefore results in the sheltering probabilities decreasing relatively less and the effective tax rates increasing relatively less given a certain increase financial statement comparability: financial statement comparability cannot completely offset the incentive-increasing effect with regard to tax avoidance of an opaque information environment.

Since these results are contrary to my expectations, I test the model again, now defining IA = 1 as the smallest bid-ask spread quintile (H2b(b)), as I have also done for H2a. See table 17. As I argued before, these firms have a more transparent information environment. The regression results show a similar pattern as before. All the interaction coefficients are of the same sign as the FSC-score coefficients and insignificant (0.001; 0.002; -0.013; -0.030 and 0.001; 0.001; -0.005; -0.003 are the coefficients of respectively the FSC-score variable and the interaction variable, for CAHS_ETR, GAAP_ETR, SHEL1 and SHEL2). This again indicates that a more transparent information environment results in even higher effective tax rates and even lower tax sheltering probabilities. This might be caused by the tax authorities being better able to evaluate a company's tax-relevant transactions, which increases the chances of getting caught avoiding taxes, which further reduces tax avoidance incentives.

However, all these results (for both H2b(a) and H2b(b)) are again insignificant, so no strong conclusions should be drawn from these results.

Overall, one might cautiously think that the deterring effect on tax avoidance of financial statement comparability is more pronounced in firms with a high information transparency. For firms with a less transparent information environment however, the effect of financial statement comparability on aggressive tax sheltering is mitigated. The decreasing effect financial statement comparability has on tax avoidance can thus not completely outweigh the incentive increasing effect with regard to tax avoidance of an opaque information environment. This can more easily be seen in the following summarizing tables, showing only the IA*FSC-score and FSC-score variables. All the IA*FSC-score coefficients that have an opposite sign with respect to the FSC-score variables are bold and blue. Again however, practically none of these results are significant. So the evidence for the inferred effects is weak.

Picture 1 – summary of important information of tables 10, 11, 16 and 17

H2a - firm size					
H2a(a)					
IA = 1 first quintile	**CASH_ETR**	**GAAP_ETR**	**SHEL1**	**SHEL2**	*Opaque information environment*
FSC	0.004***	0.003***	-0.048*	-0.043*	
IA*FSC-score	**-0.001**	0.001	**0.012**	**0.001**	
H2a(b)					
IA = 1 fifth quintile	**CASH_ETR**	**GAAP_ETR**	**SHEL1**	**SHEL2**	*Transparent information environment*
FSC-score	0.003***	0.003***	-0.042**	-0.043**	
IA*FSC-score	0.002	0.002	-0.294	-0.173	
H2b - bid-ask spread					
H2b(a)					
IA = 1 fifth quintile	**CASH_ETR**	**GAAP_ETR**	**SHEL1**	**SHEL2**	*Opaque information environment*
FSC-score	0.002*	0.003**	-0.043	-0.041	
IA*FSC-score	**-0.002**	**-0.001**	**0.070**	**0.045**	
H2b(b)					
IA = 1 first quintile	**CASH_ETR**	**GAAP_ETR**	**SHEL1**	**SHEL2**	*Transparent information environment*
FSC-score	0.001	0.002*	-0.013	-0.030	
IA*FSC-score	0.001	0.001	-0.005	-0.003	

This table presents a summary of the regression results presented in tables 10, 11, 16 and 17. This table presents only the FSC-score and IA*FSC-score, so the effect the information has on the relation between the financial statement comparability and tax avoidance can be seen at a glance.

In total, contrary to Suk & Zhou (2017) and other extant literature, both models for the second hypothesis do not support the hypothesis that financial statement comparability has a more pronounced effect on aggressive tax avoidance for firms with a poor information environment. If any, financial statement comparability has a less pronounced effect in case of a more opaque information environment.

7. Robustness test

Although I control for industry and time fixed effects, as well as for many other relevant factors following prior literature, it can be argued that some factors that are not included in the regression models and that enhance financial statement comparability, could also influence tax avoidance. It can for example not be ruled out that the level of financial statement comparability might partly be the result of some endogenous factors. If those factors also influence tax avoidance, the results might be subject to endogeneity concerns. For example: demanding investor relations and more specifically: some audit effort and high profits required by the investors. Those demands might affect both the financial statement comparability and tax avoidance. Kang et al. (2015) for example acknowledge that audit effort might affect financial statement comparability. Also, the results could be affected by reverse causality. Aggressive tax planning practices decrease the transparency of external information. This is due to the complexity of the activities related to tax planning (Balakrishnan et al., 2018). This could decrease financial statement comparability, which would lead to reverse causality.

To address these endogeneity concerns, I perform a robustness test using an instrumental variable two-stage regression approach.

For the instrumental variable to be a valid instrument in this case, it should be correlated with financial statement comparability but should not be correlated with aggressive tax avoidance (Angrist and Krueger 2001). Therefore, I use firms' their geographic locations as the instrument. More specifically, I use the firms' proximity to its industry peers. My instrument is defined as the mean value of the distance between the city where a firm's headquarter is situated, and the cities where the industry peers[22] of that particular firm have their headquarters situated. Of this mean distance, I take the natural logarithm following Suk and Zhao (2017). The cities where the companies in my sample have their headquarters situated can be found in Compustat. I calculate the distance between these cities as follows. I first obtain the geographical coordinates of all the cities where the companies in my database have their

[22] Industry peers being peers within the same two-digit SIC industry.

headquarters through a Stata module.[23] Then I calculated the distance between all these coordinates using another Stata module.[24] Lastly, I calculate for every company in a two-digit SIC industry the mean distance to its industry peers.

The reasoning behind using the geographical distance between industry peers as instrumental variable is the following. The distance between a firm and its industry peers influences the geographical and legal environments the firms are situated in as well as the firm's interactions with its peers, neighbors and auditors. These factors influence the accounting practices of the firms, and that in turn results in a certain (change in) financial statement comparability (Kedia and Rajgopal 2009, Choi et al. 2012). However, there is no theoretical nor practical foundation for the claim that the (mean) geographic distance between industry peers might have an effect on tax avoidance (Suk and Zhao 2017). Suk and Zhou (2017) also performed multiple (weak-)instrument tests to test if the *'geographic distance to industry peers*'-instrument is indeed valid. They find that their tests reject the null hypothesis of there being no correlation between the instrumental variable and comparability (significance at 1%). Also, the Cragg-Donald F-Statistic for IV indicates that the geographic proximity to industry peers is a valid instrument.

In the first stage of this two-stage analysis I regress my measure of financial statement comparability on the mean geographic proximity to industry peers as well as the control variables of the main regression models in this thesis. With these regressions I obtain the fitted values for financial statement comparability. Then in the second stage, I replicate my main analysis models, only now using the fitted values for financial statement comparability obtained from the first stage. I do not report the results of this first-stage regression. I present the results of the second stage analysis in tables 18 to 22. I present a summary of tables 18 to 22, showing only the IA*FSC-score and FSC-score variables, below, so the results of these regressions can be seen at a glance. All the IA*FSC-score coefficients that have an opposite sign with respect to the FSC-score variables are bold and blue.

[23] OpenCage Geo, https://opencagedata.com.
[24] Vincenty, Direct and inverse solutions of geodesics on the ellipsoid with application of nested equations, https://www.ngs.noaa.gov/PUBS_LIB/inverse.pdf.

Picture 2 - summary of important information of tables 18 to 22

H1	**CASH_ETR**	**GAAP_ETR**	**SHEL1**	**SHEL2**	
FSC	0.104***	0.158***	-1.117	-0.913	
H2a - firm size					
IA = 1 first quintile	**CASH_ETR**	**GAAP_ETR**	**SHEL1**	**SHEL2**	*Opaque information environment*
FSC	0.107***	0.161***	-1.317	-0.941	
IA * FSC	**-0.010**	**-0.015**	**0.270**	**0.043**	
IA = 1 fifth quintile	**CASH_ETR**	**GAAP_ETR**	**SHEL1**	**SHEL2**	
FSC	0.095***	0.147**	-1.323	-0.966	*Transparent information environment*
IA * FSC	0.033**	0.039***	**0.591**	-0.164	
H2b - bid-ask spread					
IA = 1 fifth quintile	**CASH_ETR**	**GAAP_ETR**	**SHEL1**	**SHEL2**	*Opaque information environment*
FSC	0.095**	0.072	-0.223	1.270***	
IA * FSC	0.023**	0.023**	-0.289	**-0.306*** **	
IA = 1 first quintile	**CASH_ETR**	**GAAP_ETR**	**SHEL1**	**SHEL2**	*Transparent information environment*
FSC	0.095**	0.066	-0.342	1.144	
IA * FSC	0.009	**-0.019**	-0.157	**-0.064**	

This table presents a summary of the regression results presented in tables 18 to 22. This table presents only the FSC-score and IA*FSC-score, so the effect the information has on the relation between the financial statement comparability and tax avoidance can be seen at a glance.

The results of this robustness tests mostly confirm the results of my main regression analyses. For H1 the FSC-score is again positive for GAAP_ETR and CASH_ETR and negative for SHEL1 and SHEL2. However, it is no longer significant for SHEL1 and SHEL2, so the evidence is somewhat less strong. Since the sign of the SFC-score coefficients is the same as in my main analysis however, I assume these results do not contradict the results of my main analysis of hypothesis 1. If any, they confirm my analyses.

With regard to the second hypothesis and focusing on H2a, the analysis where firm size proxies for the quality of the information environment, the robustness test results again confirm my main analysis. When IA = 1 for the smallest firms, indicating an opaque information environment, the IA*FSC-score coefficient is of the opposite direction to the FSC-score coefficient for every regression. When IA = 1 for the biggest firms, meaning the information environment is transparent, the IA*FSC-score coefficient is for all but one variable in the same

direction as FSC-score coefficient. In my main analysis I found that when a firm has a poor information environment, the reduction in tax avoidance (characterized by decreasing sheltering probabilities and increasing average tax rates) as a result of increased financial statement comparability is mitigated by the information environment. When a firm has a transparent information environment, the reduction in tax avoidance by financial statement comparability is increased by the information environment. The results of this robustness test indicate mostly the same. This confirms therefore, with regard to the sign of the different coefficients, most of my main analyses of hypothesis H2a. Most of the results are however still insignificant. So it is still not possible to draw any strong conclusions from these results.

With regard to the second hypothesis and looking to H2b, the results are less clear. When IA = 1 for the quintile of biggest bid-ask spreads, indicating an opaque information environment, only for SHEL2 the IA*FSC-score variable is opposite to the FSC-score variable. The other IA*FSC-score variables are, contrary to the results of my main analyses and the robustness analysis of H2a, in the same direction as the FSC-score variable. Also, the FSC-score variable of SHEL2 is not the expected sign. When IA = 1 for the quintile of smallest bid-ask spreads, indicating a transparent information environment, the results are mixed 50/50, and again almost all insignificant. For GAAP_ETR and SHEL2 the signs of the IA*FSC-score variables are opposite to the signs of the FSC-score variables themselves, while for CASH_ETR and SHEL2 the signs of the IA*FSC-score variables are in the same direction as the signs of the FSC-score variable themselves. Only the latter are in line with results of my main analyses. The analysis with regard to H2b thus provides little to no evidence for the mitigating effect of an opaque information environment on the relation between financial statement comparability and tax avoidance, nor for the increasing effect in case of transparent information environment. However, most of these results are again insignificant.

In total, almost all results of these robustness tests are in line with my main analyses, looking at the sign (positive or negative) of the coefficients. Only the robustness analysis of H2b yields in inconclusive results. Again however, almost all the results of these robustness analyses are insignificant, as were the results of my main H2 analyses (H2a and H2b). I therefore (can only) conclude that these robustness tests do not show my main analysis, the implications and my drawn conclusions to be false. If any, the results of these robustness tests show that the results of my main analysis hold, at least with regard to the direction of almost all the effects.

8. Conclusion – Summary and implications

In this thesis I examined what effect the comparability of financial statements has on corporate tax avoidance. I predict, based on prior literature, that better financial statement comparability has a positive effect on the quality and the quantity of corporate (financial) information. This enables tax authorities to better detect, and reduces the cost of detecting, tax evasion and tax avoidance. I hypothesize that this cost-reducing detection-increasing effect decreases the incentive of managers and CEO's to avoid or evade taxes. Based on multiple proxies of tax avoidance, all widely used in extant literature, I find that financial statement comparability indeed decreases tax avoidance. I further hypothesize that for firms with a poor information environment the effect of financial statement comparability would be bigger, because the comparability would then be more important for tax authorities in detecting tax avoidance. I found this last hypothesis, contrary to what I expected, not to be true. The effect seems, although not significant, even the opposite, contrary to my hypothesis and to prior literature. In a poor information environment, the tax-avoidance decreasing effect of financial statement comparability cannot completely outweigh the increasing tax-avoidance-incentive effect of a poor information environment. In a transparent information environment, the information environment does make the effect of financial statement comparability more pronounced.

The implications of these results are as follows. More comparable financial statements decrease corporate incentives to (aggressively) avoid taxes. Corporate tax avoidance therefore differs depending on the level of financial statement comparability, as does it depending on the quality of the information environment. This should be considered in academic discussions. Secondly, these results are useful for tax authorities and policy makers, for example in setting taxation policies. The effect of financial statement comparability and the information environment should be considered in the process of establishing tax sheltering models to increase the quality of tax auditing. As such, this thesis is a reaction to the called for research on the effect of the information environment on tax avoidance and a tie breaker with regard to the research to the effect of financial statement comparability on tax avoidance. Based on the implications, this thesis is relevant for tax authorities and policy makers, regulators and standard setters.

Future research might try to explain the true effect of the information environment on the relation between financial statement comparability and tax avoidance, since the results with regard to this effect are now contradicting (although the results in this thesis are insignificant). Future research might also focus on other factors, next to the information environment, that influence the effect that financial statement comparability has on tax avoidance.

Appendix A - References

Allingham, M., A. Sandmo. 1972. Income tax evasion: a theoretical analysis. *Journal of Public Economics* 1 (3-4): 323–338.

Angrist, J. D., A. B. Kreuger. 2001. Instrumental variables and the search for identification: from supply and demand to natural experiments. *Journal of Economic Perspectives* 15 (4): 69-85.

American Institute of Certified Public Accounts (AICPA). 2010. *Statements on Standards for Tax Services. No. 1. Tax Return Positions. Section 5.* Durham, NC: AICPA. Available at: https://www.aicpa.org/interestareas/tax/resources/standardsethics/statementsonstandardsfortaxservices/downloadabledocuments/ssts,%20effective%20january%201,%202010.pdf

Amiram, D., A. Bauwer, M. Frank. 2017. *Tax avoidance at public corporations driven by shareholder taxes: evidence from changes in divided tax policy.* Darden Business School Working Paper No. 2111467. Available at SSRN: https://ssrn.com/abstract=2111467.

Amel-Zadeh, A., A. C. Della Bina. 2018. *Do U.S. Analysts Improve the Local Information Environment of Cross-Listed Stocks? Evidence from Recommendation Revisions.* Saïd Business School WP 2016-28. Available at SSRN: https://ssrn.com/abstract=2801109.

Armstrong, C., J. Core, D. Taylor, R. Verrecchia. 2010. When does information asymmetry affect the cost of capital? *Journal of Accounting Research* 49 (1): 1-40.

Armstrong, C., J. Blouin, A. Jagolinzer, D. Larcker. 2015. Corporate governance, incentives, and tax avoidance. *Journal of Accounting and Economics* 60 (1): 1–17.

Badertscher, B., S. Katz, S. Rego. 2013. The separation of ownership and control and corporate tax avoidance. *Journal of Accounting and Economics* 56 (2): 228–250.

Balakrishnan, K., J. Blouin, W. Guay. 2018. Tax aggressiveness and corporate transparency. *Accounting Review,* forthcoming. ISSN 0001-4826.

Barth, M., W. Landsman, M. Lang, C. Williams. 2012. Are IFRS–based and US GAAP–based accounting amounts comparable? *Journal of Accounting and Economics* 54 (1): 68–93.

Bowers, S., J. Drucker. 2017. Apple's cash mountain, how it avoids tax, and the Irish link. *The Irish Times*. Available at https://www.irishtimes.com/business/apple-s-cash-mountain-how-it-avoids-tax-and-the-irish-link-1.3281734.

Bradshaw, M. T., G. S. Miller, G. Serafeim. 2009. *Accounting method heterogeneity and analysts' forecasts*. University of Chicago, University of Michigan and Harvard Business School Working Paper. Available at: https://www.researchgate.net/publication/255611106_Accounting_Method_Heterogeneity_and_Analysts'_Forecasts.

Brown, N., M. Kimbrough. 2011. Intangible investment and the importance of firm–specific factors in determination of earnings. *Review of Accounting Studies* 16 (3): 539–573.

Bozanic, Z., J. Hoopes, J. Thornock, B. Williams. 2017. IRS attention. *Journal of Accounting Research* 55 (1): 79–114.

Boer, C. de. 2012. The emerging requirement for taxonomy standardization and management. *Interactive Business Reporting* 2 (4): 9-11.

Campbell, J., P. Yeung. 2017. Earnings comparability, accounting similarities, and stock returns: Evidence from peer firms' earnings restatements. *Journal of Accounting, Auditing & Finance* 32 (4): 480–509.

Chen, S., X. Chen, Q. Cheng, T. Shevlin. 2010. Are family firms more tax aggressive than non–family firms? *Journal of Financial Economics* 95 (1): 41–61.

Chen, C., D. W. Collins, T. D. Kravet, R. D. Mergenthaler. 2017. Financial statement comparability and the efficiency of acquisition decisions. *Contemporary Accounting Research* 35 (1): 203-210.

Chi, S., S. Huang, J. Sanchez. 2017. CEO inside debt incentives and corporate tax sheltering. *Journal of Accounting Research* 55 (4): 837–876.

Choi, J., S. Choi, L. A. Myers, D. Ziebart. 2018. Financial statement comparability and the informativeness of stock prices about future earnings. *Contemporary Accounting Research*. Accepted Author Manuscript.

Choi, J.P., T. Furusawa, J. Ishikawa. 2017. *Transfer pricing and the arm's length principle under imperfect competition.* Michigan State University and Hitotsubashi University Working Paper. Available at: http://www7.econ.hitu.ac.jp/cces/trade_conference_2017/paper/jota_ishikawa.pdf
Crocker, K., J. Slemrod. 2005. Corporate tax evasion with agency costs. *Journal of Public Economics* 89 (9-10): 1593–1610.
Dechow, P., R. Sloan, A. Sweeney. 1995. Detecting earnings management. *The Accounting Review* 70 (2): 193-225.
DeFond, M., M. Hung. 2003. An empirical analysis of analysts' cash flow forecasts. *Journal of Accounting and Economics* 35 (1): 73–100.
Davidson, S., R. Heaney. 2012. Effective tax rates and the political cost hypothesis: A re-evaluation of Australian evidence. *Australian Tax Forum* 27 (1); 79-105.
Desai, M. A., D. Dharmapala. 2006. Corporate tax avoidance and high powered incentives. *Journal of Financial Economics* 79 (1): 145-179.
Desai, M., A. Dyck, L. Zingales. 2007. Theft and taxes. *Journal of Financial Economics* 84 (3): 1–39.
Desai, M., C. Foley, J. Hines. 2008. Capital structure with risky foreign investment. *Journal of Financial Economics* 88 (3): 534–553.
Drobetz, W., M. C. Gruninger, S. Hirschvogl. 2010. Information asymmetry and value of cash. *Journal of Banking & Finance* 34 (9): 2168–2184.
Dyreng, S., M. Hanlon, E. Maydew. 2008. Long–run corporate tax avoidance. *The Accounting Review* 83 (1): 61–82.
Dyreng, S., M. Hanlon, E. Maydew. 2010. The effects of executives on corporate tax avoidance. *The Accounting Review* 85 (4): 1163-1189.
Fernandes, N., M. A. Ferreira. 2008. Does international cross-listing improve the information environment? *Journal of Financial Economics* 88 (2): 216-244.
Filippin, A., C. V. Fiorio, E. Viviano. 2013. The effect of tax enforcement on tax morale. *European Journal of Political Economy* 32 (3): 320-331.
Francis, J., M. Pinnuck, O. Watanabe. 2014. Auditor style and financial statement comparability. *The Accounting* Review 89 (2): 605–633.
De Franco, G., S. Kothari, R. Verdi. 2011. The benefits of financial statement comparability. *Journal of Accounting Research* 49 (4): 895–931.
Frank, M., L. Lynch, S. Rego. 2009. Tax reporting aggressiveness and its relation to aggressive financial reporting. *The Accounting Review* 84 (2): 467–96.
Financial Accounting Standards Board (FASB). 1980. *Statement of Financial Accounting Concepts No. 2. Qualitative characteristics of accounting information*. Norwalk, CT: FASB. Available at: http://www.fasb.org/cs/BlobServer?blobcol=urldata&blobtable=MungoBlobs&blobkey=id&blobwhere=1175820900526&blobheader=application%2Fpdf
Financial Accounting Standards Board (FASB). 2010. *Statement of Financial Accounting Concepts No. 8. Conceptual framework for financial reporting.* Norwalk, CT: FASB. Available at: https://www.fasb.org/jsp/FASB/Document_C/DocumentPage?cid=1176157498129&acceptedDisclaimer=true
Gallemore, J., E. Labro. 2015. The importance of the internal information environment for tax avoidance. *Journal of Accounting and Economics* 60 (1): 149-167.
Goh, B. W, J. Lee, C. Y. Lim, T. Shevlin. 2016. The effect of corporate tax avoidance on the cost of equity. *The Accounting Review* 91 (6): 1647–1670.
Gupta, S., K. Newberry. 1997. Determinants of the variability in corporate effective tax rates: Evidence from longitudinal data. *Journal of Accounting and Public Policy* 16 (1): 1–34.
Graham, J., A. Tucker. 2006. Tax shelters and corporate debt policy. *Journal of Financial Economics* 81 (3): 563–594.
Gross, C., P. Perotti. 2017. Output-based measurement of accounting comparability: A survey of empirical proxies. *Journal of Accounting Literature* 39 (3): 1-22.
Gordon, I., N. Gallery. 2012. Assessing financial reporting comparability across institutional settings: the case of pension accounting. *The British Accounting Review* 44 (1): 11-20.

Hand, L. 1934. Opinion in the United States Supreme Court decision in *Gregory v. Helvering, 69F.2d 809 (2d Cir.).* 293 U.S. 465.
Hanlon, M., S. Heitzman. 2010. A review of tax research. *Journal of Accounting and Economics* 50 (2-3): 127–178.
Hansen, D., R. Crosser, D. Laufer. 1992. Moral ethics v. tax ethics: The case of transfer pricing among multinational corporations. *Journal of Business Ethics* 11 (9): 679–686.
Hasan, I., C. Hoi, Q. Wu, H. Zhang. 2017. Does social capital matter in corporate decisions? Evidence from corporate tax avoidance. *Journal of Accounting Research* 55 (3): 629–668.
Hoi, C., Q. Wu, H. Zhang. 2013. Is Corporate Social Responsibility (CSR) associated with tax avoidance? Evidence from irresponsible CSR activities. *The Accounting Review* 88 (6): 2025–2059.
Hoitash, R., U. Hoitash, A. Kurt, R. Verdi., 2017*. An XBRL-Based Measure of Financial Statement Comparability.* Bentley University, Northeastern University, Suffolk University and MIT Sloan School of Management Working Paper No. 3208928. Available at SSRN: https://ssrn.com/abstract=3208928.
Hoitash, R., U. Hoitash. 2018. Measuring Accounting Reporting Complexity with XBRL. *The Accounting Review* 93 (1): 259-287.
Huddart, S., B. Ke. 2007. Information asymmetry and cross-sectional variation in insider trading. *Contemporary Accounting Research* 24 (1): 195-232.
Hope, O., M. Ma, W. Thomas. 2013. Tax avoidance and geographic earnings disclosure. *Journal of Accounting and Economics* 56 (2): 170–189.
Holmes, O., Jr. (1930). Opinion in Iowa Supreme Court decision in *Superior oil co. v. state of Mississippi, 50 S. Ct 169, 72 L.Ed.* 504. 280 U.S. 390.
Hoopes, J., D. Mescall, J. Pittman. 2012. Do IRS audits deter corporate tax avoidance? *The Accounting Review* 87 (5): 1603–1639.
Internal Revenue Service (IRS). 2016. *Publication 542. Corporations.*
Kang, M., J. W. Kim, H. Lee, M. Lee. 2015. Financial statement comparability and audit efficiency: evidence from South Korea. *Applied Economics* 47 (4): 358-373.
Kedia, S., S. Rajgopal. 2009. Neighborhood matters: The impact of location on broad based stockoptions plans. *Journal of Financial Economics* 92 (1): 109–127.
Kim, J., Y. Li, L. Zhang. 2011. Corporate tax avoidance and stock price crash risk: Firm-level analysis. *Journal of Financial Economics* 100 (3): 639–662.
Kim, S., P. Kraft, S. Ryan. Financial statement comparability and credit risk. *Review of Accounting Studies* 18 (3): 783–823.
Kim, J., L. Li, L. Lu, Y. Yu. 2016. Financial statement comparability and expected crash risk. *Journal of Accounting and Economics* 61 (2-3): 294–312.
Kubick, T., D. Lynch, M. Mayberry, T. Omer. 2016. The effects of regulatory scrutiny on tax avoidance: An examination of SEC comment letters. *The Accounting Review* 91 (6): 1751–1780.
Kubick, T., G. Lockhart, L. Mills, J. Robinson. 2017. IRS and corporate taxpayer effects of geographic proximity. *Journal of Accounting and Economics* 63 (2): 428-453.
Leung, E., A. Verriest. 2015. The Impact of IFRS 8 on Geographical Segment Information. *Journal of Business Finance & Accounting* 42 (3): 273-309.
Lisowsky, P. 2010. Seeking shelter: empirically modeling tax shelters using financial statement information. *The Accounting Review* 85 (5): 1693–1720.
Maydew, E. 2001. Empirical tax research in accounting: A discussion. *Journal of Accounting and Economics* 31 (1-3), 389–403.
Mclntyre, S., M. Gardner, R. Wilkins, R. Phillips. 2011. *Corporate tax payers and corporate tax dodgers 2008–10.* A joint project of Citizens for Tax Justice and the Institute on Taxation and Economic Policy. Available at https://www.ctj.org/wp-content/uploads/2017/11/CorporateTaxDodgers Report.pdf
Mills, L., K. Newberry. 2001. The influence of tax and non–tax costs on book–tax reporting differences: public and private firms. *Journal of the American Taxation Association* 23 (1): 1–19.

Murray, A. 2002. *Narrowing tax gap should be priority of next Congress*. Wall Street Journal of 8 October 2002.
Ozkan, A., N. Ozkan. 2004. Corporate cash holdings: An empirical investigation of UK companies. *Journal of Banking & Finance* 28 (9): 2103–2134.
Payne, D. M., C. A. Raiborn. 2018. Aggressive tax avoidance; A Conundrum for Stakeholders, Governments and Morality. *Journal of Business Ethics* 147 (3): 469-487.
Peterson, K., R. Schmardebeck, T. J. Wilks. 2015. The earnings quality and information processing effects of accounting consistency. *The Accounting Review* 90 (6): 2483–2514.
Phillips, J., M. Pincus, S. Rego. 2003. Earnings management: New evidence based on deferred tax expense. *The Accounting Review* 78 (2): 491–521.
Philips, R., M. Gardner, A. Robins, M. Surka. 2017. *Offshore shell games, the use of offshore tax havens by fortune 500 companies.* A joint project of the U.S. PIRG Education Fund and the Institute on Taxation and Economic Policy. Available at: https://uspirgedfund.org/sites/pirg/files/reports/USP%20ShellGames%20Oct17%201.2.pdf.
Rego, S. 2003. Tax–avoidance activities of U.S. multinational corporations. *Contemporary Accounting Research* 20 (4): 805–33.
Rego, S., R. Wilson. 2012. Equity risk incentives and corporate tax aggressiveness. *Journal of Accounting Research* 50 (3): 775–810.
Securities and Exchange Commission (SEC). 2009. *Interactive data to improve financial reporting*. Washington, DC: SEC. Available at: https://www.sec.gov/rules/final/2009/33-9002.pdf
Schipper, K. 2003. Principle-based accounting standards. *Accounting Horizons* 17 (1): 61–72.
Segal, D., J. L. Callen, J. Livnat. 2006. Accounting Restatements: are they always bad news for investors? *Journal of Investing* 15 (3): 57-68.
Shackelford, D., T. Shevlin. 2001. Empirical tax research in accounting. *Journal of Accounting and Economics* 31 (1-3), 321–387.
Sikka, P. (2010). Smoke and mirrors: Corporate social responsibility and tax avoidance. *Accounting Forum* 34 (3-4): 153–168.
De Simone, L. 2016. Does a common set of accounting standards affect tax-motivated income shifting for multinational firms? *Journal of Accounting and Economics* 61 (1): 145-165.
Song, Y., T. Yarbrough. 1978. Tax ethics and taxpayer attitudes: A survey. *Public Administration Review* 38 (5): 442–452.
Suk, I., Y. Zhao. 2017. *Does Financial Statement Comparability Deter Aggressive Tax Avoidance?* Buffalo School of Management Working Paper No. 3065250. Available at: https://ssrn.com/abstract= 3065250.
United States Government Accountability Office. 2008. *U.S. Multinational Corporations: Effective tax rates are correlated with where income is reported.* A report to the committee on Finance, U.S. Senate. Available at: http://www.gao.gov/assets/280/279648.pdf
Vermaelen, T. 1981. Common stock repurchases and market signaling: An empirical study. *Journal of Financial Economics* 9 (2): 139–183.
Weisbach, D. A. 2002. Ten truths about tax shelters. *Tax Law Review* 55 (2): 65-104.
Wilson, R. 2009. An examination of corporate tax shelter participants. *The Accounting Review* 84 (3): 969–99.
Wu, Roung-Shiunn, C. S. Ou, H. Lin, S. Chang, D. Yen. 2012. Using data mining technique to enhance tax evasion detection performance. *Expert systems with applications* 39 (10): 8769-8777.
Zimmerman, J. 1983. Taxes and firm size. *Journal of Accounting and Economics* 5 (1): 119–149.

Appendix B - Variables

Tax avoidance measures (*TA*)

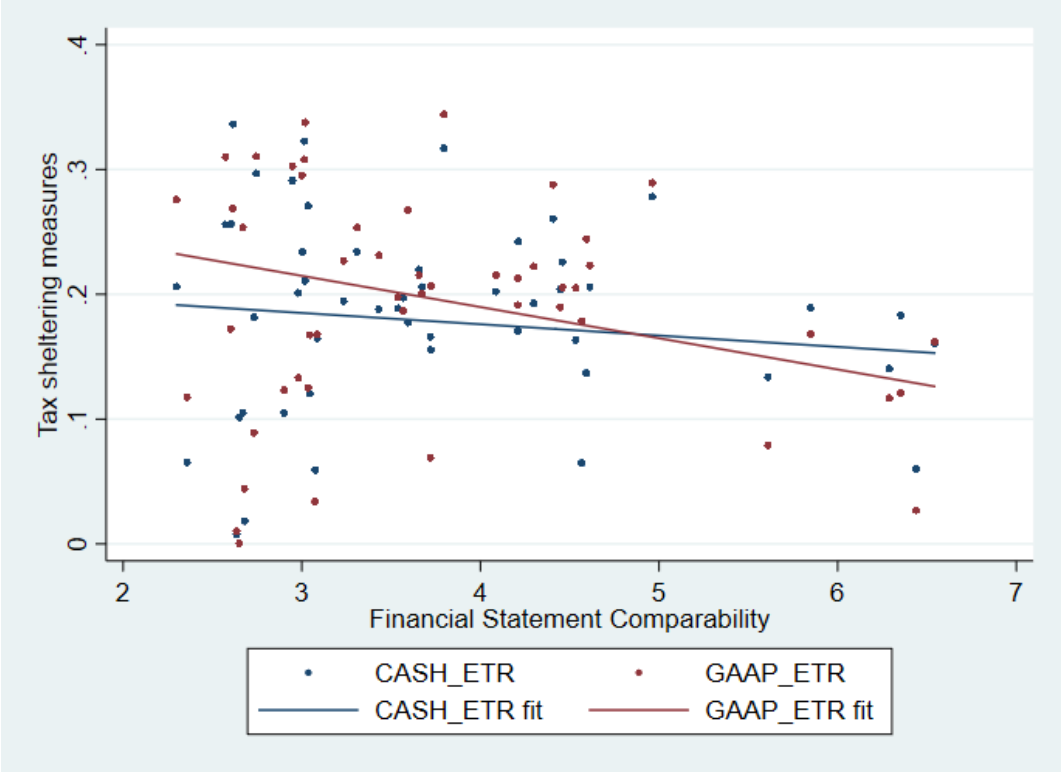
CASH_ETR	Total tax expense / (book income before tax – special items)
GAAP_ETR	Total tax expense / book income before tax
SHEL1	Tax sheltering score calculated with the Wilson (2009) sheltering probability model: shelter_1 = -4.86 + 5.20* BTD + 4.08*DA - 1.41*LEV + 0.76*log (AT) + 3.51*ROA + 1.72* FOREIGN INCOME +2.43*R&D. SHEL1 = 1 if the estimated shelter_1 is in the top quartile in that year, and 0 otherwise.
SHEL2	Tax sheltering score calculated with the Wilson (2009) sheltering probability model: shelter_2 = -4.30 + 6.63* BTD - 1.72*LEV + 0.66*log (AT) + 2.26*ROA + 1.62*FOREIGN INCOME +1.56*R&D. SHEL2 = 1 if the estimated shelter_2 is in the top quartile in that year, and 0 otherwise.

Other explanatory variables

Δtlcf	Change in loss carry forward (TLCF) scaled by lagged assets (AT).
nol	Equals 1 if loss carry forward (TLCF) is positive at the beginning of the year.
roa	Return on assets, measured as pretax income (PI) scaled by lagged assets (AT).
lev	Leverage for the firm, measured as long-term debt (DLTI) scaled by lagged assets (AT).
ppe	Property, plant, and equipment (PPENT) scaled by lagged assets (AT).
intan	Intangible assets (INTAN) scaled by lagged assets (AT).
r&d	Research and development expense ratio, measured as research and development expense (XRD) scaled by lagged assets (AT). Missing values in XRD are set to 0.
esub	Equity income in earnings (ESUB) scaled by lagged assets (AT).
pifo	Foreign income (PIFO) scaled by lagged assets (AT). Missing values in PIFO are set to 0.
mb	Market-to-book ratio at the beginning of the year, measured as the market value of equity (PRCC_F × CSHO) scaled by book value of equity (CEQ).
Δsale	Changes in sales (SALE) scaled by lagged sales.
size	Natural logarithm of the market value of equity (PRCC_F × CSHO) at the beginning of the year.
emp	The natural logarithm of the number of employees (EMP).
cash	Cash holding, defined as cash and marketable securities (CHE) divided by lagged assets (AT).
da	Discretionary accruals, computed by using the modified Jones model (Dechow et al., 1995) and including lagged ROA as an additional regressor, scaled by lagged assets.

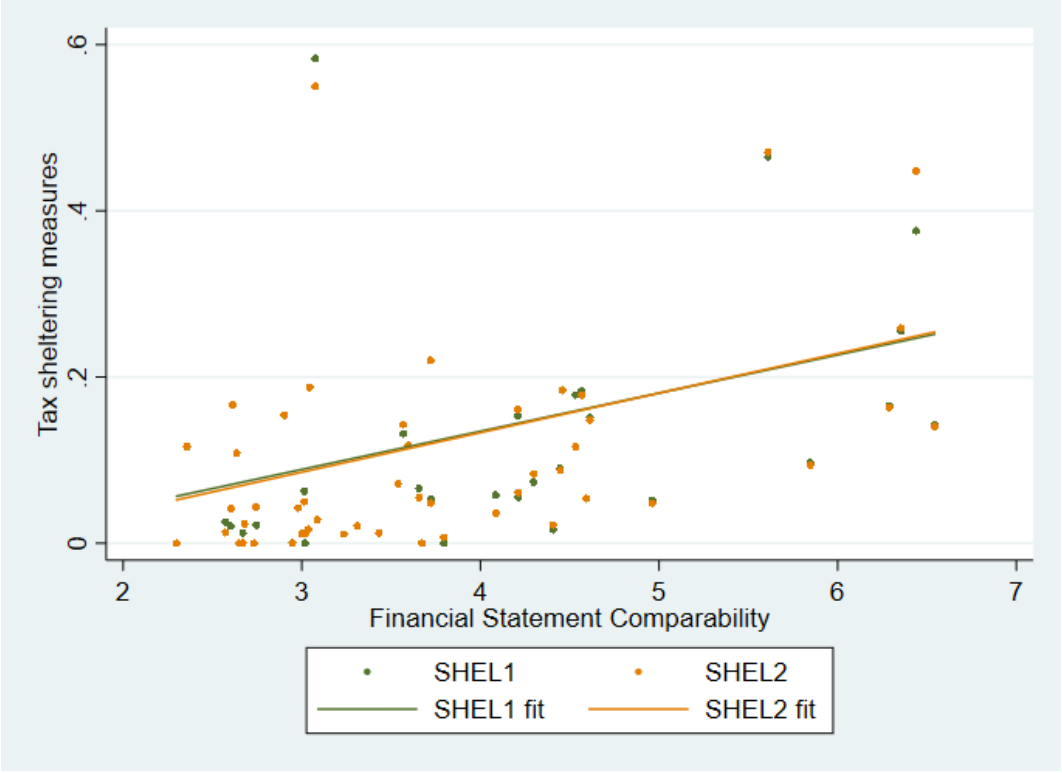
Graphs

Graph 1 - Correlation between the ETR's and financial statement comparability



This graph presents the relation between CASH_etr and financial statement comparability, as well as the relation between GAAP_ETR and financial statement comparability. All the dots are industry averages. Industries with a higher financial statement comparability on average are characterized by a lower average effective tax rate.

Graph 2 - Correlation between the SHEL's and financial statement comparability



This graph presents the relation between SHEL1 and financial statement comparability, as well as the relation between SHEL2 and financial statement comparability. All the dots are industry averages. Industries with a higher financial statement comparability on average are characterized by a higher average sheltering probability.

Tables

Table 1 – Sample CASH_ETR for H1

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
cash_etr	5,850	0.192	0.171	0.000	1.000
lag_cash_etr	5,850	0.175	0.235	-0.659	1.291
changetlcf	5,850	0.271	1.864	-2.642	19.250
lag_nol	5,850	0.844	0.363	0.000	1.000
lag_lev	5,850	0.170	0.355	0.000	2.420
lag_roa	5,850	-0.311	2.252	-20.404	0.446
lag_rand	5,850	0.059	0.171	0.000	1.488
lag_ppe	5,850	0.257	0.263	0.000	1.303
lag_intan	5,850	0.244	0.282	0.000	1.576
lag_esub	5,850	4.964	26.341	-22.862	210.700
lag_pifo	5,850	0.017	0.039	-0.145	0.171
lag_mb	5,850	2.833	11.693	-85.885	71.645
changesale	5,850	0.068	0.269	-1.008	1.639
lag_size	5,850	6.543	2.352	-0.066	11.508
lag_ln_emp	5,850	0.432	2.535	-6.215	5.159
lag_cash	5,850	0.231	0.351	0.000	2.882
lag_da	5,850	-0.077	1.030	-9.831	3.169
lag_lfsc	5,850	5.012	2.522	1.787	15.376
Number of industries	249	249	249	249	249

This table provides the descriptive statistics of the sample of CASH_ETR used in the analysis of hypothesis 1. This sample includes 5.850 firm-year observations over the period 2009 – 2017. The definition of the variables is reported in Appendix B. All continous variables are winsorized at the top 1% and bottem 1%.

Table 2 – Sample GAAP_ETR for H1

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
gaap_etr	7,385	0.219	0.190	0.000	1.000
lag_gaap_etr	7,385	0.157	0.393	-2.359	1.319
changetlcf	7,385	0.439	2.123	-2.642	19.250
lag_nol	7,385	0.863	0.344	0.000	1.000
lag_lev	7,385	0.166	0.368	0.000	2.420
lag_roa	7,385	-0.528	2.533	-20.404	0.446
lag_rand	7,385	0.113	0.249	0.000	1.488
lag_ppe	7,385	0.241	0.269	0.000	1.303
lag_intan	7,385	0.221	0.285	0.000	1.576
lag_esub	7,385	4.169	24.298	-22.862	210.700
lag_pifo	7,385	0.011	0.039	-0.145	0.171
lag_mb	7,385	2.927	13.357	-85.885	71.645
changesale	7,385	0.057	0.285	-1.008	1.639
lag_size	7,385	6.087	2.420	-0.066	11.508
lag_ln_emp	7,385	-0.274	2.734	-6.215	5.159
lag_cash	7,385	0.322	0.480	0.000	2.882
lag_da	7,385	-0.108	1.173	-9.831	3.169
lag_lfsc	7,385	5.104	2.528	1.787	15.376
Number of industries	250	250	250	250	250

This table provides the descriptive statistics of the sample of GAAP_ETR used in the analysis of hypothesis 1. This sample includes 7.385 firm-year observations over the period 2009 – 2017. The definition of the variables is reported in Appendix B. All continous variables are winsorized at the top 1% and bottem 1%.

Table 3 – Sample SHEL1 for H1

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
shel1	7,456	0.235	0.424	0.000	1.000
lag_shel1	7,456	0.230	0.421	0.000	1.000
changetlcf	7,456	0.464	2.152	-2.642	19.250
lag_nol	7,456	0.903	0.295	0.000	1.000
lag_lev	7,456	0.160	0.369	0.000	2.420
lag_roa	7,456	-0.570	2.547	-20.404	0.446
lag_rand	7,456	0.129	0.252	0.000	1.488
lag_ppe	7,456	0.213	0.256	0.000	1.303
lag_intan	7,456	0.228	0.294	0.000	1.576
lag_esub	7,456	3.408	22.388	-22.862	210.700
lag_pifo	7,456	0.009	0.040	-0.145	0.171
lag_mb	7,456	3.227	13.720	-85.885	71.645
changesale	7,456	0.050	0.293	-1.008	1.639
lag_size	7,456	5.847	2.383	-0.066	11.508
lag_ln_emp	7,456	-0.565	2.600	-6.215	5.159
lag_cash	7,456	0.352	0.492	0.000	2.882
lag_da	7,456	-0.112	1.197	-9.831	3.169
shelter1	7,456	3.663	57.709	-108.709	607.939
lag_lfsc	7,456	5.516	2.595	1.787	15.376
Number of industries	144	144	144	144	144

This table provides the descriptive statistics of the sample of SHEL1 used in the analysis of hypothesis 1. This sample includes 7.456 firm-year observations over the period 2009 – 2017. The definition of the variables is reported in Appendix B. All continuous variables are winsorized at the top 1% and bottom 1%.

Table 4 – Sample SHEL2 for H1

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
shel2	7,464	0.234	0.423	0.000	1.000
lag_shel2	7,464	0.227	0.419	0.000	1.000
changetlcf	7,464	0.466	2.161	-2.642	19.250
lag_nol	7,464	0.903	0.296	0.000	1.000
lag_lev	7,464	0.161	0.370	0.000	2.420
lag_roa	7,464	-0.572	2.556	-20.404	0.446
lag_rand	7,464	0.129	0.252	0.000	1.488
lag_ppe	7,464	0.214	0.258	0.000	1.303
lag_intan	7,464	0.229	0.294	0.000	1.576
lag_esub	7,464	3.426	22.504	-22.862	210.700
lag_pifo	7,464	0.009	0.040	-0.145	0.171
lag_mb	7,464	3.241	13.737	-85.885	71.645
changesale	7,464	0.050	0.292	-1.008	1.639
lag_size	7,464	5.849	2.382	-0.066	11.508
lag_ln_emp	7,464	-0.560	2.602	-6.215	5.159
lag_cash	7,464	0.351	0.491	0.000	2.882
lag_da	7,464	-0.112	1.197	-9.831	3.169
shelter2	7,464	4.128	59.601	-86.283	607.487
lag_lfsc	7,464	5.511	2.597	1.787	15.376
Number of industries	144	144	144	144	144

This table provides the descriptive statistics of the sample of SHEL2 used in the analysis of hypothesis 1. This sample includes 7.464 firm-year observations over the period 2009 – 2017. The definition of the variables is reported in Appendix B. All continuous variables are winsorized at the top 1% and bottom 1%.

Table 5 – The effects of financial statement comparability on tax avoidance - H1

	(1)	(2)	(3)	(4)
VARIABLES	CASH_ETR	GAAP_ETR	SHEL1	SHEL2
lag_lfsc	0.003***	0.004***	-0.046***	-0.041**
	(3.460)	(4.337)	(-2.629)	(-2.351)
changetlcf	-0.002	-0.003***	-0.230***	-0.268***
	(-1.288)	(-2.878)	(-11.135)	(-11.871)
lag_nol	-0.075***	-0.059***	1.052***	1.044***
	(-13.273)	(-10.696)	(5.094)	(5.073)
lag_roa	0.002	0.002*	0.030	0.001
	(1.041)	(1.658)	(1.281)	(0.022)
lag_lev	-0.008	-0.002	0.116	0.194*
	(-1.352)	(-0.346)	(1.039)	(1.787)
lag_ppe	-0.036***	0.004	-0.124	-0.184
	(-3.511)	(0.459)	(-0.569)	(-0.864)
lag_intan	-0.004	-0.001	-0.284*	-0.359**
	(-0.450)	(-0.114)	(-1.713)	(-2.249)
lag_rand	-0.065***	-0.068***	1.012***	1.188***
	(-4.607)	(-6.809)	(4.957)	(5.751)
lag_esub	-0.000	-0.000***	-0.038*	-0.025
	(-1.276)	(-3.918)	(-1.688)	(-1.262)
lag_pifo	0.142**	0.024	-6.358***	-6.329***
	(2.568)	(0.471)	(-4.260)	(-4.377)
lag_mb	0.000	0.000	0.004	0.003
	(0.375)	(0.392)	(1.567)	(1.244)
changesale	-0.010	0.014**	-0.909***	-0.714***
	(-1.352)	(2.124)	(-7.111)	(-5.642)
lag_size	-0.000	0.006***	-0.157***	-0.165***
	(-0.213)	(3.960)	(-4.173)	(-4.488)
lag_ln_emp	0.018***	0.022***	-0.589***	-0.537***
	(10.330)	(13.822)	(-13.805)	(-13.029)
lag_cash	0.002	-0.027***	0.159	0.089
	(0.289)	(-5.392)	(1.557)	(0.888)
lag_da	-0.004*	-0.006***	-0.103***	-0.050
	(-1.665)	(-2.785)	(-2.630)	(-1.213)
lagged D.V.	0.194***	0.047***	2.585***	2.454***
	(22.150)	(9.773)	(24.722)	(23.964)
Year fixed effects	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes
Constant	0.214***	0.231***		
	(16.567)	(19.207)		
Observations	5,850	7,385	7,456	7,464
R-squared	0.233	0.241		
Pseudo R2			0.558	0.535
Number of industries	249	250	144	144

This table presents the results of the regression of hypothesis 1, which focusses on the effect that financial statement comparability has on tax avoidance. Lagged D.V. means lagged dependent variable. The definition of the other variables is reported in Appendix B. ***, ** and * respectively represent significance at the 1%, 5% and 10% levels (two-sided). The T-statistics (CASH_ETR and GAAP_ETR) and Z-statistics (SHEL1 and SHEL2) are in parentheses.

Table 6 – Sample CASH_ETR for H2a

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
cash_etr	5,746	0.194	0.171	0.000	1.000
lag_cash_etr	5,746	0.177	0.236	-0.659	1.291
changetlcf	5,746	0.245	1.754	-2.642	19.250
lag_nol	5,746	0.842	0.364	0.000	1.000
lag_lev	5,746	0.168	0.348	0.000	2.420
lag_roa	5,746	-0.284	2.181	-20.404	0.446
lag_rand	5,746	0.060	0.171	0.000	1.488
lag_ppe	5,746	0.260	0.263	0.000	1.303
lag_intan	5,746	0.247	0.282	0.000	1.576
lag_esub	5,746	4.990	26.507	-22.862	210.700
lag_pifo	5,746	0.017	0.039	-0.145	0.171
lag_mb	5,746	2.902	11.655	-85.885	71.645
changesale	5,746	0.067	0.264	-1.008	1.639
lag_size	5,746	6.608	2.274	0.029	11.508
lag_ln_emp	5,746	0.494	2.485	-6.215	5.159
lag_cash	5,746	0.228	0.341	0.000	2.882
lag_da	5,746	-0.091	1.012	-9.831	3.169
lag_lfsc	5,746	5.009	2.502	1.787	15.376
Number of industries	244	244	244	244	244

This table provides the descriptive statistics of the sample of CASH_ETR used in the analysis of hypothesis 2, part a. This sample includes 5.746 firm-year observations over the period 2009 – 2017. The sample differs from the sample used in the analysis of hypothesis 1, because lag_IA (firm size) is now required. This extra variable made the sample of firm-year observations for which all variables are available a bit smaller. The definition of the variables is reported in Appendix B. All continuous variables are winsorized at the top 1% and bottom 1%.

Table 7 – Sample GAAP_ETR for H2a

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
gaap_etr	7,268	0.221	0.190	0.000	1.000
lag_gaap_etr	7,268	0.160	0.388	-2.359	1.319
changetlcf	7,268	0.410	2.019	-2.642	19.250
lag_nol	7,268	0.862	0.345	0.000	1.000
lag_lev	7,268	0.164	0.362	0.000	2.420
lag_roa	7,268	-0.492	2.432	-20.404	0.446
lag_rand	7,268	0.114	0.249	0.000	1.488
lag_ppe	7,268	0.243	0.269	0.000	1.303
lag_intan	7,268	0.223	0.285	0.000	1.576
lag_esub	7,268	4.186	24.432	-22.862	210.700
lag_pifo	7,268	0.012	0.039	-0.145	0.171
lag_mb	7,268	2.992	13.361	-85.885	71.645
changesale	7,268	0.056	0.280	-1.008	1.639
lag_size	7,268	6.143	2.360	0.019	11.508
lag_ln_emp	7,268	-0.226	2.707	-6.215	5.159
lag_cash	7,268	0.320	0.476	0.000	2.882
lag_da	7,268	-0.112	1.127	-9.831	3.169
lag_lfsc	7,268	5.104	2.514	1.787	15.376
Number of industries	245	245	245	245	245

This table provides the descriptive statistics of the sample of GAAP_ETR used in the analysis of hypothesis 2, part a. This sample includes 7.268 firm-year observations over the period 2009 – 2017. The sample differs from the sample used in the analysis of hypothesis 1, because lag_IA (firm size) is now required. This extra variable made the sample of firm-year observations for which all variables are available a bit smaller. The definition of the variables is reported in Appendix B. All continuous variables are winsorized at the top 1% and bottom 1%.

Table 8 – Sample shel1 for H2a

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
shel1	7,357	0.234	0.423	0.000	1.000
lag_shel1	7,357	0.228	0.420	0.000	1.000
changetlcf	7,357	0.433	2.039	-2.642	19.250
lag_nol	7,357	0.902	0.297	0.000	1.000
lag_lev	7,357	0.159	0.362	0.000	2.420
lag_roa	7,357	-0.533	2.446	-20.404	0.446
lag_rand	7,357	0.129	0.251	0.000	1.488
lag_ppe	7,357	0.214	0.256	0.000	1.303
lag_intan	7,357	0.230	0.293	0.000	1.576
lag_esub	7,357	3.447	22.531	-22.862	210.700
lag_pifo	7,357	0.009	0.040	-0.145	0.171
lag_mb	7,357	3.286	13.691	-85.885	71.645
changesale	7,357	0.051	0.288	-1.008	1.639
lag_size	7,357	5.901	2.324	0.019	11.508
lag_ln_emp	7,357	-0.524	2.568	-6.215	5.159
lag_cash	7,357	0.350	0.487	0.000	2.882
lag_da	7,357	-0.118	1.151	-9.831	3.169
lag_lfsc	7,357	5.516	2.583	1.787	15.376
Number of industries	141	141	141	141	141

This table provides the descriptive statistics of the sample of SHEL1 used in the analysis of hypothesis 2, part a. This sample includes 7.357 firm-year observations over the period 2009 – 2017. The sample differs from the sample used in the analysis of hypothesis 1, because lag_IA (firm size) is now required. This extra variable made the sample of firm-year observations for which all variables are available a bit smaller. The definition of the variables is reported in Appendix B. All continuous variables are winsorized at the top 1% and bottom 1%.

Table 9 – Sample SHEL2 for H2a

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
shel2	7,359	0.232	0.422	0.000	1.000
lag_shel2	7,359	0.226	0.418	0.000	1.000
changetlcf	7,359	0.436	2.051	-2.642	19.250
lag_nol	7,359	0.902	0.297	0.000	1.000
lag_lev	7,359	0.160	0.364	0.000	2.420
lag_roa	7,359	-0.535	2.456	-20.404	0.446
lag_rand	7,359	0.129	0.251	0.000	1.488
lag_ppe	7,359	0.215	0.258	0.000	1.303
lag_intan	7,359	0.230	0.293	0.000	1.576
lag_esub	7,359	3.428	22.510	-22.862	210.700
lag_pifo	7,359	0.009	0.040	-0.145	0.171
lag_mb	7,359	3.294	13.729	-85.885	71.645
changesale	7,359	0.051	0.288	-1.008	1.639
lag_size	7,359	5.905	2.325	0.019	11.508
lag_ln_emp	7,359	-0.516	2.570	-6.215	5.159
lag_cash	7,35	0.350	0.486	0.000	2.882
lag_da	7,359	-0.118	1.151	-9.831	3.169
lag_lfsc	7,359	5.511	2.585	1.787	15.376
Number of industries	140	140	140	140	140

This table provides the descriptive statistics of the sample of SHEL2 used in the analysis of hypothesis 2, part a. This sample includes 7.359 firm-year observations over the period 2009 – 2017 The sample differs from the sample used in the analysis of hypothesis 1, because lag_IA (firm size) is now required. This extra variable made the sample of firm-year observations for which all variables are available a bit smaller. The definition of the variables is reported in Appendix B. All continuous variables are winsorized at the top 1% and bottom 1%.

Table 10 – Hypothesis H2a(a); effects of the information environment - firm size

	(1)	(2)	(3)	(4)
VARIABLES	CASH_ETR	GAAP_ETR	SHEL1	SHEL2
lag_lfsc	0.004***	0.003***	-0.048*	-0.043*
	(3.304)	(3.439)	(-1.896)	(-1.745)
lag_IA	-0.009	-0.023**	-0.504**	-0.430*
	(-0.677)	(-1.981)	(-2.094)	(-1.819)
lag_IA_lag_lfsc	-0.001	0.001	0.012	0.001
	(-0.667)	(0.392)	(0.338)	(0.034)
changetlcf	-0.002	-0.003***	-0.221***	-0.256***
	(-1.168)	(-2.715)	(-10.292)	(-11.111)
lag_nol	-0.073***	-0.058***	1.091***	1.054***
	(-12.872)	(-10.362)	(5.228)	(5.084)
lag_roa	0.002	0.002	0.020	0.001
	(0.927)	(1.638)	(0.792)	(0.029)
lag_lev	-0.009	-0.003	0.149	0.216*
	(-1.494)	(-0.472)	(1.248)	(1.892)
lag_ppe	-0.035***	0.006	-0.220	-0.287
	(-3.356)	(0.647)	(-0.979)	(-1.315)
lag_intan	-0.003	0.001	-0.339**	-0.368**
	(-0.339)	(0.091)	(-1.978)	(-2.238)
lag_rand	-0.065***	-0.067***	0.850***	1.082***
	(-4.504)	(-6.474)	(4.053)	(5.113)
lag_esub	-0.000	-0.000***	-0.041*	-0.026
	(-0.964)	(-3.400)	(-1.692)	(-1.211)
lag_pifo	0.150***	0.033	-6.459***	-6.376***
	(2.694)	(0.647)	(-4.269)	(-4.376)
lag_mb	0.000	0.000	0.004	0.003
	(0.408)	(0.480)	(1.495)	(1.262)
changesale	-0.011	0.016**	-0.917***	-0.741***
	(-1.433)	(2.401)	(-6.919)	(-5.661)
lag_size	-0.002	0.004**	-0.251***	-0.250***
	(-1.133)	(2.080)	(-5.408)	(-5.541)
lag_ln_emp	0.017***	0.022***	-0.600***	-0.546***
	(9.894)	(13.436)	(-13.748)	(-12.995)
lag_cash	0.001	-0.029***	0.176*	0.105
	(0.159)	(-5.574)	(1.668)	(1.019)
lag_da	-0.004	-0.007***	-0.082*	-0.042
	(-1.565)	(-2.887)	(-1.878)	(-0.913)
lagged D.V.	0.193***	0.050***	2.630***	2.492***
	(21.857)	(10.258)	(24.302)	(23.611)
Year fixed effects	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes
Constant	0.225***	0.248***		
	(15.574)	(18.406)		
Observations	5,746	7,268	7,357	7,359
R-squared	0.226	0.235		
Pseudo R2			0.568	0.542
Number of industries	244	245	141	140

This table present the results of the regression of hypothesis 2, part a, sub a, focusing on the effect the information environment, proxied by firm size, has on the relation between financial statement comparability and tax avoidance. IA is an indicator variable, which equals 1 if a firm is in the first quintile of the firm size distribution, indicating an opaque information environment. Lagged D.V. means the lagged dependent variable. The definition of the variables is reported in Appendix B. ***, ** and * respectively represent significance at the 1%, 5% and 10% levels (two-sided). The T-statistics (CASH_ETR and GAAP_ETR) and Z-statistics (SHEL1 and SHEL2) are in parentheses.

Table 11 – Hypothesis H2a(b); effects of the information environment - firm size

	(1)	(2)	(3)	(4)
VARIABLES	CASH_ETR	GAAP_ETR	SHEL1	SHEL2
lag_lfsc	0.003***	0.003***	-0.042**	-0.043**
	(2.913)	(3.691)	(-2.299)	(-2.386)
lag_baspread	-0.034***	-0.064***	0.404	-0.190
	(-2.840)	(-5.270)	(0.328)	(-0.162)
lag_IA_lag_lfsc	0.002	0.002	-0.294	-0.173
	(0.914)	(0.984)	(-1.258)	(-0.814)
changetlcf	-0.002	-0.003**	-0.223***	-0.257***
	(-1.092)	(-2.569)	(-10.408)	(-11.147)
lag_nol	-0.073***	-0.057***	1.086***	1.056***
	(-12.926)	(-10.318)	(5.217)	(5.111)
lag_roa	0.001	0.002	0.019	-0.001
	(0.820)	(1.124)	(0.751)	(-0.029)
lag_lev	-0.009	-0.004	0.147	0.211*
	(-1.527)	(-0.716)	(1.234)	(1.844)
lag_ppe	-0.038***	-0.000	-0.241	-0.309
	(-3.669)	(-0.012)	(-1.072)	(-1.412)
lag_intan	-0.004	-0.001	-0.378**	-0.403**
	(-0.521)	(-0.113)	(-2.213)	(-2.458)
lag_rand	-0.064***	-0.066***	0.890***	1.114***
	(-4.405)	(-6.417)	(4.228)	(5.251)
lag_esub	-0.000	-0.000***	-0.041	-0.026
	(-0.772)	(-2.614)	(-1.636)	(-1.140)
lag_pifo	0.164***	0.063	-6.565***	-6.383***
	(2.939)	(1.242)	(-4.286)	(-4.332)
lag_mb	0.000	0.000	0.004	0.003
	(0.326)	(0.478)	(1.507)	(1.244)
changesale	-0.013*	0.013*	-0.955***	-0.776***
	(-1.651)	(1.952)	(-7.228)	(-5.951)
lag_size	0.003	0.011***	-0.153***	-0.154***
	(1.372)	(6.621)	(-3.865)	(-3.966)
lag_ln_emp	0.018***	0.022***	-0.586***	-0.531***
	(10.020)	(13.871)	(-13.415)	(-12.655)
lag_cash	-0.000	-0.032***	0.146	0.076
	(-0.010)	(-6.272)	(1.382)	(0.743)
lag_da	-0.004	-0.006***	-0.080*	-0.039
	(-1.539)	(-2.624)	(-1.828)	(-0.856)
lagged D.V.	0.192***	0.049***	2.640***	2.503***
	(21.787)	(10.022)	(24.442)	(23.762)
Year fixed effects	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes
Constant	0.203***	0.212***		
	(14.677)	(16.808)		
Observations	5,746	7,268	7,357	7,359
R-squared	0.227	0.242		
Pseudo R2			0.568	0.541
Number of industries	244	245	141	140

This table present the results of the regression of hypothesis 2, part a, sub b, focusing on the effect the information environment, proxied by firm size, has on the relation between financial statement comparability and tax avoidance. IA is an indicator variable, which equals 1 if a firm is in the fifth quintile of the firm size distribution, indicating a transparent information environment. Lagged D.V. means the lagged dependent variable. The definition of the variables is reported in Appendix B. ***, ** and * respectively represent significance at the 1%, 5% and 10% levels (two-sided). The T-statistics (CASH_ETR and GAAP_ETR) and Z-statistics (SHEL1 and SHEL2) are in parentheses.

Table 12 – Sample CASH_ETR for H2b

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
cash_etr	4,829	0.213	0.167	0.000	1.000
lag_cash_etr	4,829	0.196	0.235	-0.659	1.291
changetlcf	4,829	0.033	0.476	-2.642	17.654
lag_nol	4,829	0.840	0.367	0.000	1.000
lag_lev	4,829	0.156	0.307	0.000	2.420
lag_roa	4,829	0.052	0.285	-12.167	0.446
lag_rand	4,829	0.044	0.103	0.000	1.488
lag_ppe	4,829	0.257	0.246	0.000	1.303
lag_intan	4,829	0.251	0.269	0.000	1.576
lag_esub	4,829	5.427	27.673	-22.862	210.700
lag_pifo	4,829	0.020	0.040	-0.145	0.171
lag_mb	4,829	3.308	9.075	-85.885	71.645
changesale	4,829	0.062	0.211	-1.008	1.639
lag_size	4,829	7.074	1.904	0.996	11.508
lag_ln_emp	4,829	1.011	2.024	-6.215	5.159
lag_cash	4,829	0.207	0.261	0.000	2.882
lag_da	4,829	0.002	0.066	-0.673	3.169
lag_lfsc	4,829	5.060	2.426	1.787	15.376
Number of industries	222	222	222	222	222

This table provides the descriptive statistics of the sample of CASH_ETR used in the analysis of hypothesis 2, part b. This sample includes 4.829 firm-year observations over the period 2009 – 2017. The sample differs from the sample used in the analysis of hypothesis 1 and 2a, because lag_IA (bid-ask spread) is now required. This extra variable made the sample of firm-year observations for which all variables are available a bit smaller. The definition of the variables is reported in Appendix B. All continuous variables are winsorized at the top 1% and bottom 1%.

Table 13 – Sample GAAP_ETR for H2b

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
gaap_etr	5,774	0.250	0.182	0.000	1.000
lag_gaap_etr	5,774	0.187	0.397	-2.359	1.319
changetlcf	5,774	0.125	0.704	-2.642	19.250
lag_nol	5,774	0.856	0.351	0.000	1.000
lag_lev	5,774	0.149	0.312	0.000	2.420
lag_roa	5,774	-0.062	0.601	-20.404	0.446
lag_rand	5,774	0.090	0.191	0.000	1.488
lag_ppe	5,774	0.242	0.253	0.000	1.303
lag_intan	5,774	0.228	0.270	0.000	1.576
lag_esub	5,774	4.791	26.090	-22.862	210.700
lag_pifo	5,774	0.014	0.041	-0.145	0.171
lag_mb	5,774	3.394	10.596	-85.885	71.645
changesale	5,774	0.054	0.221	-1.008	1.639
lag_size	5,774	6.733	2.003	1.175	11.508
lag_ln_emp	5,774	0.432	2.305	-6.215	5.159
lag_cash	5,774	0.293	0.408	0.000	2.882
lag_da	5,774	-0.001	0.154	-9.831	3.169
lag_lfsc	5,774	5.175	2.427	1.787	15.376
Number of industries	224	224	224	224	224

This table provides the descriptive statistics of the sample of GAAP_ETR used in the analysis of hypothesis 2, part b. This sample includes 5.774 firm-year observations over the period 2009 – 2017. The sample differs from the sample used in the analysis of hypothesis 1 and 2a, because lag_IA (bid-ask spread) is now required. This extra variable made the sample of firm-year observations for which all variables are available a bit smaller. The definition of the variables is reported in Appendix B. All continuous variables are winsorized at the top 1% and bottom 1%.

Table 14 – Sample SHEL1 for H2b

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
shel1	5,172	0.181	0.385	0.000	1.000
lag_shel1	5,172	0.170	0.376	0.000	1.000
changetlcf	5,172	0.168	0.794	-2.642	19.250
lag_nol	5,172	0.912	0.283	0.000	1.000
lag_lev	5,172	0.139	0.300	0.000	2.420
lag_roa	5,172	-0.125	0.695	-20.404	0.446
lag_rand	5,172	0.119	0.204	0.000	1.488
lag_ppe	5,172	0.201	0.239	0.000	1.303
lag_intan	5,172	0.240	0.285	0.000	1.576
lag_esub	5,172	3.357	22.281	-22.862	210.700
lag_pifo	5,172	0.011	0.043	-0.145	0.171
lag_mb	5,172	3.771	11.335	-85.885	71.645
changesale	5,172	0.052	0.219	-1.008	1.639
lag_size	5,172	6.417	2.030	0.996	11.508
lag_ln_emp	5,172	-0.047	2.211	-6.215	5.159
lag_cash	5,172	0.350	0.441	0.000	2.882
lag_da	5,172	-0.003	0.213	-9.831	3.169
lag_lfsc	5,172	5.820	2.512	1.787	15.376
Number of industries	86	86	86	86	86

This table provides the descriptive statistics of the sample of SHEL1 used in the analysis of hypothesis 2, part b. This sample includes 5.172 firm-year observations over the period 2009 – 2017. The sample differs from the sample used in the analysis of hypothesis 1 and 2a, because lag_IA (bid-ask spread) is now required. This extra variable made the sample of firm-year observations for which all variables are available a bit smaller. The definition of the variables is reported in Appendix B. All continuous variables are winsorized at the top 1% and bottom 1%.

Table 15 – Sample SHEL2 for H2b

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
shel2	5,150	0.183	0.386	0.000	1.000
lag_shel2	5,150	0.170	0.376	0.000	1.000
changetlcf	5,150	0.169	0.794	-2.642	19.250
lag_nol	5,150	0.911	0.284	0.000	1.000
lag_lev	5,150	0.140	0.302	0.000	2.420
lag_roa	5,150	-0.126	0.697	-20.404	0.446
lag_rand	5,150	0.120	0.204	0.000	1.488
lag_ppe	5,150	0.202	0.242	0.000	1.303
lag_intan	5,150	0.241	0.285	0.000	1.576
lag_esub	5,150	3.315	22.129	-22.862	210.700
lag_pifo	5,150	0.011	0.043	-0.145	0.171
lag_mb	5,150	3.801	11.452	-85.885	71.645
changesale	5,150	0.053	0.220	-1.008	1.639
lag_size	5,150	6.416	2.027	0.996	11.508
lag_ln_emp	5,150	-0.048	2.214	-6.215	5.159
lag_cash	5,150	0.351	0.441	0.000	2.882
lag_da	5,150	-0.003	0.213	-9.831	3.169
lag_lfsc	5,150	5.813	2.522	1.787	15.376
Number of industries	86	86	86	86	86

This table provides the descriptive statistics of the sample of SHEL2 used in the analysis of hypothesis 2, part b. This sample includes 5.150 firm-year observations over the period 2009 – 2017. The sample differs from the sample used in the analysis of hypothesis 1 and 2a, because lag_IA (bid-ask spread) is now required. This extra variable made the sample of firm-year observations for which all variables are available a bit smaller. The definition of the variables is reported in Appendix B. All continuous variables are winsorized at the top 1% and bottom 1%.

Table 16 – Hypothesis H2b(a); effects of the information environment - bid-ask spread

	(1)	(2)	(3)	(4)
VARIABLES	CASH_ETR	GAAP_ETR	SHEL1	SHEL2
lag_lfsc	0.002*	0.003**	-0.043	-0.041
	(1.658)	(2.085)	(-1.278)	(-1.222)
lag_baspread	0.025**	0.038***	-0.653*	-0.555
	(2.007)	(3.180)	(-1.747)	(-1.523)
lag_IA_lag_lfsc	-0.002	-0.001	0.070	0.045
	(-1.148)	(-0.681)	(1.321)	(0.865)
changetlcf	-0.003	-0.015***	0.244***	0.196***
	(-0.599)	(-4.747)	(3.517)	(2.690)
lag_nol	-0.060***	-0.034***	0.984***	0.677**
	(-9.217)	(-5.341)	(3.040)	(2.108)
lag_roa	0.067***	0.022***	-0.277*	-0.372**
	(6.055)	(4.316)	(-1.776)	(-2.133)
lag_lev	-0.010	0.011	0.048	0.273
	(-1.324)	(1.594)	(0.236)	(1.421)
lag_ppe	-0.060***	0.007	-1.048***	-0.639*
	(-4.501)	(0.532)	(-2.768)	(-1.677)
lag_intan	0.004	0.002	-0.302	-0.542**
	(0.403)	(0.244)	(-1.154)	(-2.141)
lag_rand	-0.115***	-0.110***	1.445***	0.923**
	(-3.760)	(-6.214)	(3.334)	(2.256)
lag_esub	-0.000	-0.000***	0.000	0.002
	(-0.166)	(-3.258)	(0.054)	(0.261)
lag_pifo	0.049	-0.023	-8.726***	-1.715
	(0.789)	(-0.412)	(-3.891)	(-0.770)
lag_mb	-0.000	-0.000	0.007	0.009*
	(-0.780)	(-1.101)	(1.425)	(1.813)
changesale	-0.041***	0.012	-1.713***	-1.680***
	(-3.786)	(1.237)	(-6.824)	(-6.497)
lag_size	-0.004*	0.007***	-0.260***	-0.197***
	(-1.758)	(3.422)	(-4.353)	(-3.305)
lag_ln_emp	0.016***	0.015***	-0.461***	-0.396***
	(7.134)	(7.501)	(-7.221)	(-6.349)
lag_cash	0.015	-0.027***	-0.459***	-0.483***
	(1.322)	(-3.462)	(-2.615)	(-2.766)
lag_da	0.100**	-0.052***	-0.619	-0.609
	(2.541)	(-3.654)	(-1.084)	(-0.963)
lagged D.V.	0.181***	0.045***	2.812***	3.244***
	(18.580)	(8.402)	(16.948)	(16.514)
Year fixed effects	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes
Constant	0.243***	0.218***		
	(14.468)	(14.070)		
Observations	4,804	5,749	5,165	5,143
R-squared	0.172	0.166		
Pseudo R2			0.609	0.553
Number of industries	216	217	85	85

This table present the results of the regression of hypothesis 2, part b, sub a, focusing on the effect the information environment, proxied by bid-ask spread, has on the relation between financial statement comparability and tax avoidance. IA is an indicator variable, which equals 1 if a firm is in the fifth quintile of the bid-ask spread distribution, indicating an opaque information environment. Lagged D.V. means the lagged dependent variable. The definition of the variables is reported in Appendix B. ***, ** and * respectively represent significance at the 1%, 5% and 10% levels (two-sided). The T-statistics (CASH_ETR and GAAP_ETR) and Z-statistics (SHEL1 and SHEL2) are in parentheses.

Table 17 – Hypothesis H2b(b); effects of the information environment - bid-ask spread

	(1)	(2)	(3)	(4)
VARIABLES	CASH_ETR	GAAP_ETR	SHEL1	SHEL2
lag_lfsc	0.001	0.002*	-0.013	-0.030
	(1.192)	(1.900)	(-0.463)	(-1.073)
lag_baspread	-0.018	-0.037***	0.050	0.039
	(-1.257)	(-2.694)	(0.095)	(0.081)
lag_IA_lag_lfsc	0.001	0.001	-0.005	-0.003
	(0.288)	(0.422)	(-0.052)	(-0.032)
changetlcf	-0.003	-0.015***	0.199***	0.189***
	(-0.604)	(-4.760)	(2.816)	(2.702)
lag_nol	-0.061***	-0.038***	0.716**	0.711**
	(-9.550)	(-6.000)	(2.211)	(2.381)
lag_roa	0.067***	0.023***	-0.386**	-0.384**
	(6.096)	(4.493)	(-2.471)	(-2.474)
lag_lev	-0.011	0.010	0.189	0.297
	(-1.410)	(1.344)	(0.881)	(1.524)
lag_ppe	-0.059***	0.008	-0.555	-0.680*
	(-4.428)	(0.667)	(-1.381)	(-1.724)
lag_intan	0.003	-0.001	-0.455	-0.544**
	(0.285)	(-0.087)	(-1.569)	(-2.088)
lag_rand	-0.116***	-0.111***	0.732	0.962*
	(-3.791)	(-6.309)	(1.598)	(1.894)
lag_esub	-0.000	-0.000***	0.002	0.002
	(-0.141)	(-3.179)	(0.269)	(0.252)
lag_pifo	0.044	-0.041	-1.466	-1.830
	(0.714)	(-0.738)	(-0.819)	(1.095)
lag_mb	-0.000	-0.000	0.009*	0.009*
	(-0.886)	(-1.290)	(1.779)	(1.783)
changesale	-0.041***	0.013	-1.844***	-1.708***
	(-3.780)	(1.323)	(-6.894)	(-6.611)
lag_size	-0.004*	0.006***	-0.181***	-0.180***
	(-1.927)	(2.979)	(-2.994)	(-2.987)
lag_ln_emp	0.016***	0.016***	-0.428***	-0.393***
	(7.216)	(7.825)	(-6.666)	(-6.267)
lag_cash	0.015	-0.027***	-0.576***	-0.505***
	(1.300)	(-3.513)	(-3.336)	(-3.050)
lag_da	0.099**	-0.053***	-0.714	-0.596
	(2.509)	(-3.712)	(-1.253)	(-1.093)
lagged D.V.	0.181***	0.044***	3.553***	3.265***
	(18.701)	(8.328)	(17.595)	(17.460)
Year fixed effects	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes
Constant	0.255***	0.243***		
	(15.972)	(16.387)		
Observations	4,829	5,774	5,172	5,150
R-squared	0.172	0.166		
Pseudo R2			0.594	0.551
Number of industries	222	224	86	86

This table present the results of the regression of hypothesis 2, part b, sub b, focusing on the effect the information environment, proxied by bid-ask spread, has on the relation between financial statement comparability and tax avoidance. IA is an indicator variable, which equals 1 if a firm is in the first quintile of the bid-ask spread distribution, indicating a transparent information environment. Lagged D.V. means the lagged dependent variable. The definition of the variables is reported in Appendix B. ***, ** and * respectively represent significance at the 1%, 5% and 10% levels (two-sided). The T-statistics (CASH_ETR and GAAP_ETR) and Z-statistics (SHEL1 and SHEL2) are in parentheses.

Table 18 – Robustness test H1 - instrumental variable & two-stage regression

VARIABLES	(1) CASH_ETR	(2) GAAP_ETR	(3) SHEL1	(4) SHEL2
lag_lfsc	0.104***	0.158***	-1.117	-0.913
	(2.941)	(2.716)	(-0.415)	(-0.340)
changetlcf	0.004	0.005	-0.279**	-0.310**
	(1.628)	(1.499)	(-2.154)	(-2.258)
lag_nol	-0.020	0.011	0.635	0.693
	(-1.003)	(0.420)	(0.573)	(0.608)
lag_roa	0.016***	0.009***	0.016	-0.015
	(3.102)	(3.219)	(0.397)	(-0.304)
lag_lev	0.053**	0.079**	-0.324	-0.158
	(2.405)	(2.559)	(-0.292)	(-0.145)
lag_ppe	0.007	0.115***	-1.042	-0.934
	(0.378)	(2.676)	(-0.428)	(-0.386)
lag_intan	-0.039***	-0.015*	-0.276*	-0.351**
	(-2.633)	(-1.690)	(-1.650)	(-2.125)
lag_rand	0.056	0.105	0.215	0.530
	(1.265)	(1.600)	(0.114)	(0.271)
lag_esub	0.000	0.000	-0.040*	-0.027
	(1.479)	(1.120)	(-1.703)	(-1.277)
lag_pifo	0.026	-0.165*	-6.374***	-6.486***
	(0.389)	(-1.887)	(-4.177)	(-4.240)
lag_mb	-0.000	-0.000*	0.007	0.006
	(-0.227)	(-1.674)	(0.908)	(0.701)
changesale	-0.041***	-0.015	-0.559	-0.436
	(-3.211)	(-1.173)	(-0.667)	(-0.512)
lag_size	0.008**	0.010***	-0.220	-0.214
	(2.439)	(4.524)	(-1.250)	(-1.294)
lag_ln_emp	0.028***	0.042***	-0.769*	-0.683*
	(6.825)	(5.543)	(-1.786)	(-1.645)
lag_cash	-0.001	-0.060***	0.537	0.387
	(-0.165)	(-4.399)	(0.560)	(0.407)
lag_da	-0.057***	-0.055***	0.213	0.225
	(-3.103)	(-3.002)	(0.277)	(0.277)
lagged D.V.	0.171***	0.032***	1.673	1.783
	(14.936)	(3.836)	(0.743)	(0.890)
Year fixed effects	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes
Constant	-0.434*	-0.718**		
	(-1.912)	(-2.007)		
Observations	5,766	7,299	7,427	7,435
R-squared	0.229	0.239		
Pseudo R2			0.556	0.535
Number of industries	244	245	144	144

This table present the results of the second stage regression from the instrumental variable approach. The lfsc-value is the fitted value for financial statement comparability obtained from the first stage regression. The instrumental variable is the geographical distance between a firm and its industry peers. Lagged D.V. means the lagged dependent variable. The definition of the variables is reported in Appendix B. ***, ** and * respectively represent significance at the 1%, 5% and 10% levels (two-sided). The T-statistics (CASH_ETR and GAAP_ETR) and Z-statistics (SHEL1 and SHEL2) are in parentheses.

Table 19 – Robustness test H2a(a) - instrumental variable & two-stage regression

	(1)	(2)	(3)	(4)
VARIABLES	CASH_ETR	GAAP_ETR	SHEL1	SHEL2
lag_lfsc	0.107***	0.161***	-1.317	-0.941
	(2.968)	(2.734)	(-0.478)	(-0.346)
lag_IA	0.040	0.063	-1.968***	-0.669
	(0.833)	(1.462)	(-2.693)	(-0.917)
lag_IA_lag_lfsc	-0.010	-0.015*	0.270**	0.043
	(-1.185)	(-1.949)	(2.116)	(0.342)
changetlcf	0.003	0.004	-0.270**	-0.300**
	(1.422)	(1.298)	(-2.036)	(-2.149)
lag_nol	-0.019	0.013	0.654	0.702
	(-0.929)	(0.473)	(0.577)	(0.606)
lag_roa	0.015***	0.009***	0.008	-0.014
	(2.836)	(3.095)	(0.203)	(-0.271)
lag_lev	0.052**	0.077**	-0.334	-0.143
	(2.331)	(2.473)	(-0.293)	(-0.130)
lag_ppe	0.007	0.115***	-1.218	-1.042
	(0.378)	(2.663)	(-0.489)	(-0.424)
lag_intan	-0.039***	-0.013	-0.306*	-0.355**
	(-2.585)	(-1.498)	(-1.769)	(-2.082)
lag_rand	0.050	0.100	0.005	0.421
	(1.104)	(1.516)	(0.003)	(0.213)
lag_esub	0.000*	0.000	-0.043*	-0.027
	(1.721)	(1.325)	(-1.708)	(-1.227)
lag_pifo	0.034	-0.155*	-6.375***	-6.517***
	(0.501)	(-1.760)	(-4.110)	(-4.223)
lag_mb	-0.000	-0.000*	0.007	0.006
	(-0.270)	(-1.650)	(0.932)	(0.708)
changesale	-0.040***	-0.012	-0.559	-0.468
	(-3.133)	(-0.982)	(-0.652)	(-0.542)
lag_size	0.007*	0.008***	-0.316*	-0.298*
	(1.946)	(3.200)	(-1.739)	(-1.756)
lag_ln_emp	0.028***	0.041***	-0.799*	-0.694*
	(6.638)	(5.456)	(-1.813)	(-1.648)
lag_cash	-0.001	-0.060***	0.560	0.401
	(-0.120)	(-4.387)	(0.571)	(0.415)
lag_da	-0.054***	-0.054***	0.226	0.226
	(-2.868)	(-2.865)	(0.288)	(0.275)
lagged D.V.	0.170***	0.032***	1.652	1.815
	(14.759)	(3.836)	(0.717)	(0.892)
Year fixed effects	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes
Constant	-0.435*	-0.713**		
	(-1.889)	(-1.974)		
Observations	5,705	7,227	7,331	7,334
R-squared	0.222	0.232		
Pseudo R2			0.567	0.541
Number of industries	244	245	141	140

This table present the results of the second stage regression from the instrumental variable approach, testing hypothesis 2, part a, sub a (see also table 10). The lfsc-value is the fitted value for financial statement comparability obtained from the first stage regression. The instrumental variable is the geographical distance between a firm and its industry peers. IA is an indicator variable, which equals 1 if a firm is in the first quintile of the firm size distribution, indicating an opaque information environment. Lagged D.V. means the lagged dependent variable. The definition of the variables is reported in Appendix B. ***, ** and * respectively represent significance at the 1%, 5% and 10% levels (two-sided). The T-statistics (CASH_ETR and GAAP_ETR) and Z-statistics (SHEL1 and SHEL2) are in parentheses.

Table 20 – Robustness test H2a(b) - instrumental variable & two-stage regression

	(1)	(2)	(3)	(4)
VARIABLES	CASH_ETR	GAAP_ETR	SHEL1	SHEL2
lag_lfsc	0.095***	0.147**	-1.323	-0.966
	(2.625)	(2.508)	(-0.483)	(-0.355)
lag_baspread	-0.187***	-0.248***	-4.151	-0.291
	(-2.950)	(-3.822)	(-0.798)	(-0.056)
lag_IA_lag_lfsc	0.033**	0.039***	0.591	-0.164
	(2.544)	(2.998)	(0.584)	(-0.160)
changetlcf	0.003	0.004	-0.281**	-0.303**
	(1.450)	(1.377)	(-2.133)	(-2.175)
lag_nol	-0.022	0.011	0.585	0.685
	(-1.077)	(0.398)	(0.518)	(0.593)
lag_roa	0.014***	0.008***	0.001	-0.017
	(2.723)	(2.810)	(0.032)	(-0.334)
lag_lev	0.049**	0.073**	-0.385	-0.165
	(2.180)	(2.349)	(-0.340)	(-0.150)
lag_ppe	0.002	0.106**	-1.342	-1.105
	(0.124)	(2.454)	(-0.541)	(-0.451)
lag_intan	-0.040***	-0.015*	-0.366**	-0.390**
	(-2.629)	(-1.650)	(-2.126)	(-2.303)
lag_rand	0.046	0.095	-0.059	0.410
	(1.026)	(1.449)	(-0.031)	(0.207)
lag_esub	0.000**	0.000**	-0.045*	-0.027
	(2.322)	(1.971)	(-1.667)	(-1.154)
lag_pifo	0.055	-0.124	-6.536***	-6.480***
	(0.799)	(-1.413)	(-4.181)	(-4.166)
lag_mb	-0.000	-0.000	0.007	0.006
	(-0.277)	(-1.523)	(0.953)	(0.714)
changesale	-0.040***	-0.014	-0.550	-0.487
	(-3.088)	(-1.104)	(-0.643)	(-0.565)
lag_size	0.011***	0.015***	-0.231	-0.205
	(3.063)	(6.310)	(-1.282)	(-1.218)
lag_ln_emp	0.027***	0.041***	-0.798*	-0.685
	(6.541)	(5.447)	(-1.819)	(-1.629)
lag_cash	-0.002	-0.063***	0.596	0.390
	(-0.296)	(-4.600)	(0.610)	(0.404)
lag_da	-0.052***	-0.052***	0.296	0.247
	(-2.786)	(-2.822)	(0.378)	(0.301)
lagged D.V.	0.171***	0.031***	1.556	1.792
	(14.791)	(3.775)	(0.679)	(0.882)
Year fixed effects	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes
Constant	-0.390*	-0.671*		
	(-1.691)	(-1.867)		
Observations	5,705	7,227	7,331	7,334
R-squared	0.225	0.240		
Pseudo R2			0.566	0.541
Number of industries	244	245	141	140

This table present the results of the second stage regression from the instrumental variable approach, testing hypothesis 2, part a, sub b (see also table 11). The lfsc-value is the fitted value for financial statement comparability obtained from the first stage regression. The instrumental variable is the geographical distance between a firm and its industry peers. IA is an indicator variable, which equals 1 if a firm is in the fifth quintile of the firm size distribution, indicating a transparent information environment. Lagged D.V. means the lagged dependent variable. The definition of the variables is reported in Appendix B. ***, ** and * respectively represent significance at the 1%, 5% and 10% levels (two-sided). The T-statistics (CASH_ETR and GAAP_ETR) and Z-statistics (SHEL1 and SHEL2) are in parentheses.

Table 21 – Robustness test H2b(a) - instrumental variable & two-stage regression

	(1)	(2)	(3)	(4)
VARIABLES	CASH_ETR	GAAP_ETR	SHEL1	SHEL2
lag_lfsc	0.095**	0.072	-0.223	1.270*
	(2.417)	(1.107)	(-0.087)	(1.831)
lag_baspread	-0.110*	-0.096	1.430	1.465
	(-1.793)	(-1.536)	(0.845)	(1.558)
lag_IA_lag_lfsc	0.023**	0.023**	-0.289	-0.306***
	(2.020)	(2.036)	(-0.897)	(-10.659)
changetlcf	0.002	-0.011**	0.185	0.250***
	(0.315)	(-2.515)	(0.713)	(3.390)
lag_nol	-0.004	0.002	0.525	1.143**
	(-0.167)	(0.082)	(0.244)	(2.148)
lag_roa	0.084***	0.026***	-0.385	-0.365**
	(6.845)	(4.558)	(-0.789)	(-2.437)
lag_lev	0.050**	0.050	0.058	0.770*
	(2.011)	(1.445)	(0.047)	(1.931)
lag_ppe	-0.017	0.061	-0.763	0.453
	(-0.783)	(1.254)	(-1.404)	(1.528)
lag_intan	-0.029*	-0.004	-0.464	-0.524**
	(-1.660)	(-0.346)	(-1.633)	(-2.016)
lag_rand	0.011	-0.024	0.502	1.803**
	(0.190)	(-0.314)	(0.162)	(2.496)
lag_esub	0.000*	-0.000	0.002	0.005
	(1.820)	(-0.087)	(0.154)	(0.758)
lag_pifo	-0.073	-0.109	-1.458	-2.168
	(-0.969)	(-1.123)	(-0.040)	(-0.333)
lag_mb	-0.000	-0.000	0.009	0.005
	(-1.262)	(-1.569)	(0.476)	(0.948)
changesale	-0.071***	-0.001	-1.725***	-2.096***
	(-4.494)	(-0.059)	(-2.665)	(-4.852)
lag_size	0.005	0.009***	-0.213	-0.118
	(1.342)	(3.302)	(-0.725)	(-1.491)
lag_ln_emp	0.025***	0.024***	-0.473	-0.208
	(5.348)	(2.872)	(-0.971)	(-1.592)
lag_cash	0.014	-0.043***	-0.468	-0.936**
	(1.199)	(-2.639)	(-0.841)	(-2.360)
lag_da	0.063	-0.075***	-0.665	-0.987*
	(1.447)	(-3.048)	(-0.126)	(-1.733)
lagged D.V.	0.158***	0.036***	3.294**	4.154***
	(12.398)	(3.951)	(2.028)	(9.356)
Year fixed effects	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes
Constant	-0.369	-0.218		
	(-1.464)	(-0.544)		
Observations	4,773	5,719	5,154	5,124
R-squared	0.172	0.166		
Pseudo R2			0.594	0.552
Number of industries	216	217	85	84

This table present the results of the second stage regression from the instrumental variable approach, testing hypothesis 2, part b, sub a (see also table 16). The lfsc-value is the fitted value for financial statement comparability obtained from the first stage regression. The instrumental variable is the geographical distance between a firm and its industry peers. IA is an indicator variable, which equals 1 if a firm is in the fifth quintile of the bid-ask spread distribution, indicating an opaque information environment. Lagged D.V. means the lagged dependent variable. The definition of the variables is reported in Appendix B. ***, ** and * respectively represent significance at the 1%, 5% and 10% levels (two-sided). The T-statistics (CASH_ETR and GAAP_ETR) and Z-statistics (SHEL1 and SHEL2) are in parentheses.

Table 22 – Robustness test H2b(b) - instrumental variable & two-stage regression

	(1)	(2)	(3)	(4)
VARIABLES	CASH_ETR	GAAP_ETR	SHEL1	SHEL2
lag_lfsc	0.095**	0.066	-0.342	1.144
	(2.420)	(1.014)	(-0.730)	(1.071)
lag_baspread	-0.058	0.065	0.868	0.382
	(-0.835)	(1.055)	(0.872)	(0.234)
lag_IA_lag_lfsc	0.009	-0.019	-0.157	-0.064
	(0.618)	(-1.581)	(-1.775)	(-0.217)
changetlcf	0.001	-0.012***	0.183**	0.246***
	(0.149)	(-2.733)	(2.405)	(2.989)
lag_nol	-0.011	-0.009	0.610	1.232**
	(-0.478)	(-0.289)	(1.573)	(2.263)
lag_roa	0.084***	0.027***	-0.392**	-0.374**
	(6.859)	(4.585)	(-2.127)	(-2.063)
lag_lev	0.047*	0.043	0.027	0.738
	(1.878)	(1.231)	(0.083)	(1.489)
lag_ppe	-0.018	0.051	-0.859	0.364
	(-0.815)	(1.058)	(-1.645)	(0.379)
lag_intan	-0.030*	-0.007	-0.433	-0.498*
	(-1.727)	(-0.627)	(-1.612)	(-1.939)
lag_rand	0.006	-0.039	0.464	1.781**
	(0.109)	(-0.523)	(0.154)	(2.259)
lag_esub	0.000**	-0.000	0.001	0.005
	(2.023)	(-0.242)	(0.116)	(0.643)
lag_pifo	-0.072	-0.105	-1.475	-2.197
	(-0.945)	(-1.076)	(-0.427)	(-1.254)
lag_mb	-0.000	-0.000	0.010**	0.005
	(-1.287)	(-1.570)	(2.034)	(0.899)
changesale	-0.070***	0.002	-1.718***	-2.092***
	(-4.461)	(0.136)	(-5.345)	(-4.975)
lag_size	0.004	0.007***	-0.200***	-0.103
	(1.023)	(2.769)	(-3.031)	(-1.155)
lag_ln_emp	0.026***	0.023***	-0.492***	-0.222
	(5.432)	(2.766)	(-4.871)	(-1.162)
lag_cash	0.014	-0.041**	-0.451**	-0.919**
	(1.214)	(-2.534)	(-1.991)	(-2.480)
lag_da	0.063	-0.075***	-0.622	-0.936
	(1.459)	(-3.044)	(-1.035)	(-1.439)
lagged D.V.	0.159***	0.037***	3.244***	4.123***
	(12.479)	(4.061)	(7.989)	(4.955)
Year fixed effects	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes
Constant	-0.348	-0.148		
	(-1.380)	(-0.370)		
Observations	4,773	5,719	5,154	5,124
R-squared	0.172	0.166		
Pseudo R2			0.593	0.550
Number of industries	216	217	85	84

This table present the results of the second stage regression from the instrumental variable approach, testing hypothesis 2, part b, sub b (see also table 17). The lfsc-value is the fitted value for financial statement comparability obtained from the first stage regression. The instrumental variable is the geographical distance between a firm and its industry peers. IA is an indicator variable, which equals 1 if a firm is in the first quintile of the bid-ask spread distribution, indicating a transparent information environment. Lagged D.V. means the lagged dependent variable. The definition of the variables is reported in Appendix B. ***, ** and * respectively represent significance at the 1%, 5% and 10% levels (two-sided). The T-statistics (CASH_ETR and GAAP_ETR) and Z-statistics (SHEL1 and SHEL2) are in parentheses.