Does It Pay to Not Write Well? The Case of Tax Avoiders

Justin Hung Nguyen[a]
Victoria Business School
Victoria University of Wellington, New Zealand
Justin.Nguyen@vuw.ac.nz

Nam Tran[b]
Melbourne Business School
University of Melbourne, Australia
N.Tran@mbs.edu

This version: June 2019

Abstract
Most prior studies focus on adverse impacts of low-quality disclosures and largely overlook possible benefits firms might derive from providing low-quality disclosures. This paper examines whether managers of firms that engage in high levels of tax avoidance (TA) strategically reduce their financial statement readability (FSR) to mitigate consequences of their tax planning activities. Our panel regression results show a negative relation between TA and FSR, and our difference-in-differences analysis based on the "Check-the-Box" regulation in 1997 that exogenously increases tax planning opportunities, suggests the negative impact of TA on FSR is likely causal. This relationship primarily holds for firms with high levels of TA (i.e., above industry median TA) and firms faced with high likelihood of tax audit. In addition, more TA is associated with higher firm risk and higher cost of debt; however, these negative impacts of TA are mitigated for firms with lower FSR. The evidence adds to our understanding of the multidimensional effects of FSR.

JEL Classifications: G18; G21; H26; M41
Keywords: tax avoidance; financial statement readability; firm risk; cost of debt

Acknowledgement:
We greatly acknowledge helpful comments and suggestions from Terry Shevlin, Ryan Wilson, Brian Williams, Iftekhar Hasan, Paul Healy, Henk Berkman, Steven Cahan, and seminar participants at The University of Auckland, and Victoria University of Wellington. We thank Brian Miller for sharing financial statement readability (the Bog Index) data, Jeffrey Hoopes for sharing IRS audit probability data, Bill McDonald for sharing financial statement length (10-K file size and number of words) data, and Michael Roberts for sharing Dealscan-Compustat Link data. The usual disclaimers apply.

[a] Corresponding author. Email: Justin.Nguyen@vuw.ac.nz. Phone: +64 4 463 5758. Victoria Business School, Victoria University of Wellington, Wellington 6011, New Zealand.
[b] Email: N.Tran@mbs.edu. Phone: +61 3 9349 8217. Melbourne Business School, The University of Melbourne, VIC 3010, Australia.

1. Introduction

This study examines (1) the relationship between corporate tax avoidance (TA) and financial statement readability (FSR) and (2) how FSR moderates the impacts of TA on firm risk and cost of debt. Prior research documents that firms engage in a variety of strategies to avoid paying substantial amounts of taxes (Dyreng, Hanlon, and Maydew, 2008). This behavior is not surprising given that executives are provided with significant explicit and implicit incentives to engage in TA (Phillips, 2003; Armstrong, Blouin, and Larcker, 2012; Rego and Wilson, 2012; Gaertner, 2014; Khan, Srinivasan, and Tan, 2017). Other studies consider costs associated with TA activities, including back taxes, penalty, and reputational costs to both shareholders and managers that would arise if the firm's TA strategies are challenged and overturned by tax authorities (Hanlon and Slemrod, 2009; Graham, Hanlon, Shevlin, and Shroff, 2014). Arguably, managers have a strong incentive to minimize costs associated with their TA activities. Our study examines whether firms that engage in high levels of TA issue less readable financial statements in order to avoid revealing information that makes them more likely to be audited and challenged by tax authorities. Further, we evaluate the effectiveness of this information obfuscation strategy by examining whether having lower FSR helps mitigate the impacts of TA on firm risk and cost of debt.

Prior research suggests that tax authorities use public information such as firms' financial statements in addition to private information from firms' tax returns in the law enforcement process (Bozanic, Hoopes, Thornock, and Williams, 2017). Information in financial statements can be useful for tax authorities in identifying questionable items in firms' tax returns and potentially illegitimate tax avoidance activities (Mills and Sansing 2000; Erickson, Hanlon, and Maydew, 2004). Managers that engage in risky TA strategies, therefore, have an incentive to obfuscate information in order to reduce the risk of tax audit and, upon being audited, the risk that their tax strategies are deemed noncompliant. One effective way to obfuscate

information provided to financial statement users (tax authorities included) is to make the statements less readable.

However, ex ante it is not clear whether on average firms with more TA would issue less readable financial statements to reduce the likelihood of being challenged by tax authorities. There is a wide range of TA strategies with different levels of risk, ranging from benign strategies such as buying tax-exempt municipal bonds to more risky strategies such as tax-sheltering (Dyreng et al., 2008; Blouin, 2014).[1] To the extent that firms avoid taxes by utilizing benign favorable tax treatments that have low risks of being audited and challenged by tax authorities, making financial statements less readable would impose costs on the firm without much benefit.[2] Moreover, certain TA activities result in more operational complexity, which increases information asymmetry between managers and shareholders (Desai and Dharmapala, 2006). To compensate for the increased information asymmetry, firms that engage in more TA may disclose more information and/or make extra effort to improve the readability of their financial statements. Consistent with this argument, prior studies find evidence that firms with higher TA are more likely to provide tax-related disclosures in earnings releases (Schwab, 2009), in the Management Discussion and Analysis (MD&A) section of their 10-Ks (Balakrishnan, Blouin, and Guay, 2019), and in tax footnotes (Inger, Meckfessel, Zhou, and Fan, 2018). Similarly, Ehinger, Lee, Stomberg, and Towery (2017) find that income tax is more likely to be mentioned in conference calls by firms that have larger changes in annual effective tax rate (ETR). In sum, firms that engage in TA face an inherent trade-off. On the one hand is

[1] Firms may have incentives to obfuscate information about their TA activities provided to other stakeholders as well. For example, firms do not want investors or customers to view them as irresponsible corporate citizens that engage in noncompliant TA activities (Hanlon and Slemrod 2009).

[2] Prior studies find several potential costs of having less readable financial statements, including poorer information environment (Loughran and McDonald 2014), higher cost of debts (Bonsall and Miller 2017; Ertugrul et al. 2017), higher crash risk (Ertugrul et al. 2017), less efficient investments (Biddle, Hilary, and Verdi 2009), and lower firm value (Hwang and Kim 2017).

the incentive to reduce FSR to lower the risk of tax audit; on the other hand is the need to increase FSR to reduce information asymmetry between managers and shareholders.

Dyreng, Hanlon, and Maydew (2019) suggest that firms tend to focus on safe TA strategies first, and only utilize riskier TA strategies when safe TA strategies are no longer available. Building on this argument, we assume that the higher level of TA a firm engages in the more likely that it has to adopt high-risk TA strategies. From a disclosure perspective, for firms that engage in low levels of TA by utilizing low-risk TA strategies, the incentive to provide more transparent disclosure to reduce information asymmetry is likely stronger than the incentive to obfuscate information for/submitted to tax authorities. In contrast, since firms that engage in high levels of TA likely employ high-risk TA strategies, the incentive to obfuscate information to reduce the likelihood of being audited and challenged by tax authorities likely outweighs the incentive to improve transparency for investors. The above arguments together suggest that the relation between TA and FSR is likely nonlinear. Specifically, we expect a positive association between TA and FSR for firms with relatively low levels of TA, and a negative association between TA and FSR for firms with relatively high levels of TA.

To test this hypothesis, we utilize a new measure of FSR – the Bog Index. Bonsall, Leone, Miller, and Rennekamp (2017) show that the Bog Index, which is developed specifically for measuring readability in financial disclosure documents, captures readability more accurately than the widely used Fog Index. For TA, we employ two commonly used measures: GAAP ETR and cash ETR. We use the median industry-year ETR as the threshold to partition firms into high and low TA groups. We find that, on average, firms with more TA (i.e., lower GAAP ETR or cash ETR) have less readable financial statements (i.e., higher Bog Index).[3] Consistent

[3] Inger et al. (2018) use the Fog Index to measure readability in financial statements and find no significant relationship between TA and FSR. However, Loughran and McDonald (2014) show that the Fog Index is a poor

with our hypothesis, we find that the relationship between TA and FSR is negative for the subsample of firms with relatively high levels of TA (i.e, above the industry-year median TA), and positive for the subsample of firms with relatively low levels of TA (i.e, below the industry-year median TA). This finding is consistent with Inger et al. (2018), who find that firms with high TA have *less* readable tax footnotes than firms with medium TA, but that firms with medium TA have *more* readable tax footnotes than firms with low TA.[4]

We also find that the positive relation between TA and Bog Index for firms with medium-to-high TA remains highly significant after controlling for the length of financial statements (measured by the number of words in the 10-K filing or 10-K file size). Controlling for the length of financial statements helps mitigate the concern that the Bog Index might be correlated with the amount of disclosure. That is, firms with more TA have longer financial statements simply because they have more information to disclose, and longer statements are more difficult to read.[5] Our findings are consistent with managers making their financial statements less readable when they have an incentive to avoid revealing information that might expose their TA strategies. Thus, using a more accurate measure of FSR, we are able to document evidence that Inger et al. (2018) fail to find using the Fog Index.

To mitigate concerns about endogeneity, we conduct several tests. First, we show that the positive association between TA and Bog Index remains significant after controlling for

measure of readability in financial documents, since many words considered "complex" for calculating the Fog Index (i.e., words with three or more syllables) are in fact well-understood terms in business (e.g., corporation, company, agreement, management). This might explain the lack of a significant relationship between TA and FSR documented in Inger et al. (2018).

[4] Inger et al. (2018) argue that firms that are able to reduce taxes using low-risk strategies (medium tax avoiders) provide more disclosures to highlight their skills in utilizing favorable tax treatments. Meanwhile, those with aggressive TA likely use high-risk strategies and, therefore, want to obfuscate information to reduce the likelihood of exposing their TA strategies.

[5] In untabulated tests we find that firms with more TA have longer financial statements. This could be consistent with two opposing interpretations: (1) firms with more TA provide more disclosure to make their financial statements more readable (Balakrishnan et al. 2019) or (2) firms with more TA make their financial statements unnecessarily long to impede readability (Li 2008; Laughran and McDonald 2014; Inger et al. 2018). Our findings suggest that firms with more TA have not only longer but also less readable financial statements.

firm fixed effects. Second, we utilize the "Check-the-Box" regulation event in 1997, which exogenously offers new tax planning opportunities, to improve the causal interpretation of the negative impact of TA on FSR among firms with medium-to-high TA. Finally, consistent with the notion that firms reduce FSR to lower the probability of tax audit (Hoopes, Mescall, and Pittman 2012), we find evidence that the negative relation between TA and FSR among firms with medium-to-high TA is stronger as the probability of tax audit increases.

Having established that firms with high levels of TA have less readable financial statements than those with medium levels of TA, our second set of tests examine whether having less readable financial statements mitigates some negative consequences of high TA on firms. For these tests, we focus on the group of firms with medium-to-high TA, given our previous finding that a negative association between TA and FSR robustly holds for this group. First, we examine whether having less readable financial statements reduces firm risk associated with TA. Several studies suggest that engaging in TA can increase firm risk (Rego and Wilson 2012; Badertscher, Katz, and Rego 2013; Chi, Huang, and Sanchez 2017). Further, Dyreng et al. (2019) find a positive association between TA and tax-related uncertainty measured based on the balances of uncertain tax benefits. On the other hand, Guenther, Matsunaga, and William (2017) find no evidence that TA increases firm risk as measured by ETR persistence, ETR volatility, and stock return volatility. One possible explanation for these seemingly contradictory results is that, although TA does increase firm risk, on average firms that engage in TA are able to take actions that reduce this impact of TA (e.g., issuing less readable financial statements). Thus, we test whether having less readable financial statements moderates the relation between TA and firm risk. Following Guenther et al. (2017), we use three measures of firm risk: ETR persistence, future tax rate volatility, and future stock return volatility. We also consider a fourth measure that captures the likelihood that ETR will increase in the future. Consistent with Guenther et al. (2017), we find no significant relation between

TA and firm risk for the overall sample. However, we find a positive association between TA and firm risk among firms with medium-to-high levels of TA. More importantly, the positive association between TA and firm risk is weaker for firms with less readable financial statements, suggesting that lowering FSR helps mitigate the impact of TA on firm risk. This result holds for all four measures of firm risk.

Next, we examine whether having less readable financial statements moderates the impact of TA on cost of debt. Hasan, Hoi, Wu, and Zhang (2014) find that firms with more TA have higher loan spread and suggest that lenders view borrowers with more TA as riskier, and, hence demand higher interest rates. Thus, we examine whether the impact of TA on loan spread is moderated by FSR. If TA increases adverse cash flow risk, and having less readable financial statements mitigates this risk, then FSR should have a moderating effect on the relation between TA and cost of debt.[6] Consistent with Hasan et al. (2014), we find a strong positive association between TA and loan spread. Consistent with our hypothesis that FSR moderates the impact of TA, we find that the positive relation between TA and cost of debt is weaker for firms with less readable financial statements. We also find that FSR is on average negatively related to cost of debt, consistent with Bonsall and Miller (2017). Overall, our findings suggest that having less readable financial statements mitigates the adverse impacts of TA on firm risk and cost of debt.

Our study makes important contributions to the literature. First, we add to a growing body of literature that examines the relation between financial disclosures and TA, particularly the emerging stream of research that focuses on attributes of qualitative disclosures (Law and Mills 2015; Inger et al. 2018; Balakrishnan et al. 2019). Using a new measure of readability in

[6] Having less readable financial statements might increase lenders' perceptions about the firm's information risk and agency risk. This could increase cost of debt (Hasan et al. 2014; Bonsall and Miller 2017; Ertugrul et al. 2017). This main effect of readability on cost of debt is independent of TA.

financial statements – the Bog Index – we provide evidence that the relation between TA and FSR is negative for firms that engage in relatively high levels of TA, but positive for firms that engage in relatively low levels of TA. Prior studies that use the Fog Index to measure readability find no significant relationship between FSR and TA (e.g., Law and Mills 2015; Inger et al. 2018). Our findings suggest that this absence of evidence might be due to the inaccuracy of the Fog Index in measuring readability in financial statements. In a concurrent study, Bauselinck, Blanco, Dhole, and Lobo (2018) find that firms with aggressive TA (i.e., firms that have unusually low ETR, high numbers of tax havens, and tax litigations) have less readable financial statements. We show that the relationship between TA and FSR is more complex. It is negative for firms with medium-to-high TA while being positive for firms with low-to-medium TA.

Second, we provide evidence that having less readable financial statements helps mitigate the impacts of TA on firm risk and cost of debt. These findings help explain the evidence in Inger et al. (2018) that investors value high tax avoiders with less readable tax footnotes more than high tax avoiders with more readable tax footnotes. Bonsall and Miller (2017) find that having less readable financial statements increases cost of debt for companies in general. Our findings highlight the multidimensional impacts of FSR. While having less readable financial statemnents directly increases cost of debt, presumably through increased information asymmetry, it indirectly reduces the cost of debt caused by TA, presumably by reducing the probablity of negative cash flow shocks that would arise if the firm's tax position is challenged and overturned by tax authorities.

We structure our study as follows. Section 2 discusses prior literature and develops hypotheses. In section 3, we present our data and research design. Section 4 discusses empirical results. Section 5 comprises a conclusion.

2. Literature Review and Hypothesis Development

2.1 Corporate Tax Avoidance and Financial Statement Readability

TA is a risky investment in which firms trade off tax savings against potential costs that can arise if their TA strategies are challenged and overturned by tax authorities (Amstrong et al. 2015). This risk exists even for firms that utilize ex ante legitimate TA strategies, because "tax laws, judicial rulings and administrative enforcements do not always provide clear and consistent guidance" (Law and Mills 2015, p.805). The risk is higher for firms that avoid taxes by taking advantage of "grey areas" in tax codes, which have high levels of uncertainty as to whether the company will be able to justify their strategies as legitimate.[7] Managers should have a strong incentive to reduce this risk because being challenged and deemed noncompliant by tax authorities (e.g., Internal Revenue Service, or IRS) is costly for both shareholders and managers (Hanlon and Slemrod 2009; Graham et al. 2014). Graham et al. (2014) survey Tax Executives Institute members and find that the "risk of detection and challenge by the IRS" is an important factor affecting managers' decisions on whether to implement proposed tax planning strategies.

Given their limited resources, tax authorities have to select firms to audit. This decision is likely influcenced by tax authorities' assessment of the likelihood that firms have engaged in noncompliant TA activities. One important source of information for tax authorities is firms' financial statements. Although tax authorities receive private tax-related disclosures from firms, they can utilize public information in financial statements to verify and assess the credibility of their private information, and identify potential issues or questionable items on a firm's tax returns (Mills and Sansing 2000). Erickson, Hanlon, and Maydew (2004) suggest

[7] Certain TA activities may increase firms' underlying business risk. These activities include investing in R&D to take advantage of tax credit, and investing in high-risk countries that have favorable tax treatments (Guenther et al. 2017).

that tax authorities view large differences between net income reported in tax returns and net income reported in financial statements as indication of possible illigitimate tax avoidance. More recently, Bozanic et al. (2017) provide evidence that tax authorities use 10-K filings as a source of information, presumably to help them assess the ligitimacy of firms' tax planning activities. Managers that engage in risky TA strategies, therefore, are likely to be careful when disclosing information that might increase the risk of tax audit and, upon being audited, the likelihood that their tax strategies are deemed noncompliant. Consistent with this idea, Hope, Ma, and Thomas (2013) find that firms discontinue segment disclosure to avoid revealing the source of their TA activities. On the other hand, TA activities can lead to an increase in information asymmetry between managers and shareholders (Desai and Dharmapala 2006), which increases cost of capital (Shevlin et al. 2013; Hasan et al. 2014). To mitigate these effects, firms that engage in more TA have an incentive to provide more disclosures to investors. Evidence consistent with this argument can be found in prior studies. For example, Schwab (2009) shows that firms with greater TA are more likely to voluntarily disclose book-tax difference information in earnings releases. Ehinger et al. (2017) find that income tax is more likely to be mentioned in conference calls by firms that have larger changes in ETR. Balakrishnan et al. (2019) find that firms engaging in more aggressive tax planning include more tax-related words in the MD&A section of their 10-Ks and in conference calls. Inger et al. (2018) find that TA is positively related to tax footnote readability for firms with low levels of TA. In sum, firms that engage in TA have to weigh the benefits of disclosing more information to investors to reduce information asymmetry against the costs that arise because tax authorities will use the disclosed information against the firms (Mills, Robinson, and Sansing 2010). Depending on whether the benefits of providing information to investors outweigh the benefits of obfuscating the information from tax authorities, firms would choose to provide more or less information in financial statements.

In addition to changing the quantity of disclosure, managers can also alter qualitative attributes to make their disclosure more or less informative. Prior research finds that managers strategically alter qualitative attributes of disclosure to influence financial statement users' perceptions (e.g., Li 2008; Davis and Tama-Sweet 2012; Lo, Ramos, and Rogo 2017). Li (2008) and Lo et al. (2017) find that managers make financial statements less readable when they have incentives to obfuscate information about the firm's true underlying earnings. If managers believe that less readable financial statements reduce the likelihood that their firms are audited and successfully challenged by tax authorities, then we expect firms that engage in high levels of TA to strategically have lower FSR. Balakrishnan et al. (2019) provide evidence that firms that engage in more aggressive TA have less transparent information environment, as measued by analyst forecast error, forecast dispersion, and information asymmetry. It is possible that this lack of transparency is a consequence of firms providing less readable financial statements, which makes it more difficult for analysts and investors to understand the firm's underlying economic reality.

Dyreng et al. (2019) suggest a pecking order in which firms first utilise low-risk TA strategies, and only engage in riskier TA strategies when safe TA strategies are no longer available. One implication is that the riskiness of TA strategies employed increases with the level of TA. Consistent with this idea, Dyreng et al. (2019) find that firms with higher levels of TA employ more uncertain TA strategies. As discussed above, whether firms that engage in more TA provide more or less transparent disclosure depends on whether the benefits from lower information asymmetry are greater than the costs of being audited and challenged by tax authorities. Thus, we argue that the relation between TA and FSR is likely dependent on the existing level of TA. The incentive to disclose more information to investors is likely greater than the incentive to obfuscate information for firms that engage in low levels of TA, which employ low-risk TA strategies. As a result, we expect a positive relation between TA and FSR

for those firms, consistent with the findings in prior studies (Schwab 2009; Ehinger et al. 2017; Inger et al 2018; Balakrishnan et al. 2019). Conversely, for firms that engage in high levels of TA, the incentive to obfuscate information for tax authorities is likely stronger than the incentive to provide transparent information to investors, since those firms likely utilize substantially riskier strategies, which have a high probability of being challenged and overturned by tax authorities. Thus, we expect a negative association between TA and FSR among firms with relatively high levels of TA.

Inger et al. (2018) find some evidence consistent with our prediction. Specifically, the authors find a positive association between TA and tax footnote readability for firms with low TA, and a negative association between TA and tax footnote readability for firms with high TA. However, Inger et al. (2018) do not find a significant relation between TA and readability of the whole financial statements, possibly because their measure of financial statement readability – the Fog Index – is too noisy, which lowers the power of their tests.[8] Li, Ma, Omer, and Sun (2017) find that the relation between TA and corporate transparency is negative for firms with medium-to-high TA and positive for firms with low-to-medium TA, which is also consistent with our prediction. Taking the literature as a whole, we make the following hypothesis:

H1: The relation between tax avoidance and financial statement readability is negative (positive) for firms with high (low) levels of tax avoidance.

2.2 Corporate Tax Avoidance, Financial Statement Readability, and Firm Risk

As discussed above, TA activities are risky investments in which managers trade off the benefit of higher tax savings (and possibly higher reported earnings) against the risk that their

[8] Law and Mills (2015) also do not find a significant relation between the Fog Index and TA. As discussed previously, Fog Index is a noisy measure of readability in financial disclosure documents. We use a more advanced and accurate measure – the Bog Index – to provide evidence on this important question.

TA strategies are challenged and overturned by tax authorities, which could result in financial and reputational costs (Hanlon and Slemrod 2009; Armstrong et al. 2015). In other words, TA activities can increase adverse cash flow risk by increasing the probabilities of tax audits, penalties, and interest charges (Shevlin et al. 2013). In addition to financial and reputational risks, TA can also increase information risk (Hasan et al. 2014; Balakrishnan et al. 2019) and agency problem risk (Desai and Dharmapala 2006). Consistent with the argument that TA increases cash flow risk, prior studies find that firms in which managers hold more inside debts are less likely to engage in TA activities (Kubick, Lockhart, and Robinson 2014; Chi et al. 2017). Similarly, the level of TA is lower for firms in which managerial equity ownership is higher (Badertscher, Katz, and Rego 2013), and for family-run firms (Chen, Chen, Cheng, and Shevlin 2010), consistent with managers whose wealth is more closely tied to the firm being less willing to engage in risky TA activities. Hanlon, Maydew, and Saavedra (2017) find a positive association between TA and cash holdings, suggesting that firms with higher TA hold more cash in anticipation of higher cash flow risk. Given that managers are risk averse, to motivate managers to engage in more TA, executive compensation packages often provide risk-taking incentives (Phillips 2003; Armstrong et al. 2012; Rego and Wilson 2012; Gaertner 2014).

While the above analyses suggest that TA increases firm risk, Guenther et al. (2017) find no empirical evidence supporting this prediction. The authors offer two potential explanations. First, on average, firms are able to exploit benign tax-favored treatments that do not lead to to higher risk. Second, the risk caused by TA is, on average, too small to be detected.[9] However, it is possible that TA does increase firm risk, but firms that engage in TA are able to take actions that mitigate this effect of TA, and one such action is to make their financial statements less

[9] This second explanation suggests that, while, on average, TA might not be significantly related to firm risk, firms with extreme TA might indeed bear higher risk.

readable in order to reduce the likelihood of being audited and challenged by tax authorities. Inger et al. (2018) provide some evidence consistent with this explanation. Specifically, they find that investors value firms with aggressive TA more when their tax footnotes are less readable, suggesting that investors believe that opacity helps reduce risks caused by TA activities. Thus, our next hypothesis is:

H2: Among firms with high levels of tax avoidance, the association between tax avoidance and firm risk is less positive for firms with less readable financial statements.

2.3 Corporate Tax Avoidance, Financial Statement Readability, and Cost of Debt

Prior research suggests several reasons why TA can increase cost of debt. First, high TA increases the probability that the firm will experience negative cash flow shocks caused by its tax position being audited and overturned by tax authorities (Shevlin et al. 2013; Hasan et al. 2014). Second, some tax strategies increase the complexity of the firm's operations, which offers opportunities for managers to expropriate wealth from shareholders (Desai and Dharmapala 2006), thus negatively impacting the ability of the firm to pay debts. Consistent with these arguments, Hasan et al. (2014) find a positive association between TA and bank loan spreads. Similarly, Shevlin et al. (2013) find a positive association between TA and cost of public debt.

Having less readable financial statements can have two distinct effects on cost of debt. On the one hand, less readable financial statements increase information asymmetry between managers and lenders. This likely increases information risk and agency costs, which translates into higher cost of debt (Desai and Dharmapala 2006; Shevlin et al. 2013; Hasan et al. 2014). Consistent with this argument, Bonsall and Miller (2017) find a negative relation between FSR and cost of debt. On the other hand, less readable financial statements can reduce the likelihood that the firm's tax position is audited and overturned by tax authorities. This helps mitigate the

impact of TA on cost of debt by lowering the likelihood of negative cash flow shocks caused by TA. In other words, having less readable financial statements *moderates* the positive association between TA and cost of debt. Thus, our third hypothesis is:

H3: Among firms with high levels of tax avoidance, the association between tax avoidance and cost of debt is less positive for firms with less readable financial statements.

3. Data and Research Design

3.1 Data

Our initial sample includes all firm-years from Compustat for the period 1994-2015 for which the Bog Index is available. We follow Lo et al. (2017) and exclude utilities and financial firms (those with SIC codes in the ranges 4900-4999 and 6000-6999). We then exclude observations with negative pre-tax income ((pi-spi)<0) because ETRs are difficult to interpret for firms with negative pre-tax income (Dyreng, Hanlon, Maydew, and Thornock 2017). Next, we drop observations that have missing total tax expense (*txt*), which is necessary to calculate ETR. Finally, requiring data to calculate control variables used in our baseline model further reduces our sample by 11,170 observations. Table 1 describes the sample selection procedure. The final sample consists of 45,886 firm-years for 6,681 unique firms.

3.2 Measures of Financial Statement Readability and Tax Avoidance

Our main measure of FSR, the *Bog Index*, is calculated as follows: *Bog Index* = *Sentence Bog* + *Word Bog* – *Pep*, where higher *Bog Index* values indicate lower readability. In this formula, *Sentence Bog* captures the average sentence length across the whole document. *Word Bog* measures the complexity of words in the document along two dimensions: English style problems (e.g., passive verbs, hidden verbs, overwriting, legal terms, and so on) and word difficulty (e.g., heavy words, abbreviations, and specialist terms). The final term, *Pep,* captures

writing attributes that make texts more interesting, thus facilitating readability (e.g., interesting words, and sentence variety).[10]

We employ two common measures of TA: GAAP ETR (*ETR*) and cash ETR (*CETR*). *ETR* is calculated as total tax expense (*txt*) divided by pre-tax book income (*pi*) minus special items (*si*). *CETR* is computed as cash taxes paid (*txpd*) divided by pre-tax book income (*pi*) minus special items (*si*). These measures are commonly used to measure TA in prior research (e.g., Cen, Maydew, Zhang, and Zuo 2017; Hasan, Hoi, Wu, and Zhang 2017; Koester, Shevlin, and Wangerin 2017). We winsorize *ETR* and *CETR* to be in the range [0,1]. We then multiply *CETR* and *ETR* by negative one and denote the new variables as *TA_ETR* and *TA_CETR* with higher values indicating more TA. In addition to using one-year ETR as our primary proxies of TA, we also examine three-year and five-year *ETR* and *CETR* following Dyreng et al. (2008) to check the robustness of our results (results are available upon request).

3.3 Testing the Relation between Tax Avoidance and Financial Statement Readability

Hypothesis 1 predicts a negative (positive) association between TA and FSR for firms with relatively high (low) levels of TA. We test this hypothesis by estimating the following model separately for firms with high and low TA. We use the median TA of each industry-year as the threshold to identify high- and low-TA firms.

$$BOG = a_0 + a_1 TA + a_j Control_j \quad (1)$$

BOG is the *Bog Index,* our primary measure of FSR, with higher *BOG* indicating lower readability. Our independent variable of interest is *TA*, which is either *TA_ETR* or *TA_CETR*. Since higher *BOG* indicates less readable financial statements, Hypothesis 1 predicts that the

[10] Please refer to Bonsall et al. (2017) for more details on the Bog Index construction. The Bog Index data can be found at: https://kelley.iu.edu/bpm/activities/bogindex.html. We thank Brian P. Miller for making the data available.

coefficient on *TA* (i.e., a_1) is positive for firms with high levels of TA (above the industry-year median) and negative for firms with low levels of TA (below the industry-year median).

Following prior studies (e.g., Li 2008; Lo et al. 2017; Bonsall et al. 2018), control variables in model (1) include firm profitability (*EARN*), size (*SIZE*), growth opportunities (*MTB*), age (*AGE*), special items (*SI*), stock return and earnings volatility (*RET_VOL* and *EARN_VOL*), operational complexity (*NBSEG* and *NGSEG*), financial statement complexity (*NITEMS*), significant events (*MA* and *SEO*), and Delaware incorporation state (*DLW*). Including an extensive list of direct and indirect measures of complexity helps address concerns that certain TA activities are inherently complex, which result in low FSR (Desai and Dharmapala 2006). To mitigate the impacts of outliers, we winsorize all continuous variables at their 1st and 99th percentiles. Detailed definitions of all variables can be found in the Appendix. Model (1) also controls for industry (2-digit SIC codes) and year fixed effects. Table 2 reports summary statistics for variables used in the baseline model (1). The average Bog Index is 82.183, which is comparable to prior studies that also use this measure (e.g., Bonsall and Miller 2017; Bonsall et al. 2018). The average values for *ETR* and *CETR* are 0.288 and 0.254, which are consistent with those reported in other TA papers (e.g., Cen et al. 2017; Dyreng et al. 2017; Hasan et al. 2017; Koester et al. 2017).

3.4 Testing the Moderating Effect of Financial Statement Readability on the Relation Between Tax Avoidance and Firm Risk

Hypothesis 2 predicts that the relation between TA and firm risk is less positive for firms with less readable financial statements. We test this hypothesis using the following model, estimated on the subsample of firms with above-industry-median levels of TA:

$$RISK = b_0 + b_1 (TA*BOG) + b_2 TA + b_3 BOG + b_j Control_j \quad (2)$$

Following Guenther et al. (2017), we employ three proxies for firm risk including ETR persistence (*ETR_PERSIST* or *CETR_PERSIST*), future tax rate volatility (*FETR_VOL* or *FCETR_VOL*), and future stock return volatility (*FRET_VOL*). Our fourth measure of firm risk (*ETR_INCREASE* or C*ETR_INCREASE*) is intended to capture whether the firm's ETR will increase in the future. We include in model (2) the same control variables as in Guenther et al. (2017). Please refer to the Appendix (Panel E) for detailed definitions of these variables. In model (2), we expect the coefficient on the interaction term *TA*BOG*, b_1, to be positive when the dependent variable (*RISK*) is ETR persistence (*ETR_PERSIST* or *CETR_PERSIST*), and negative when the dependent variable is future tax rate volatility (*FETR_VOL* or *FCETR_VOL*), future stock return volatility (*FRET_VOL*), or a dummy variable capturing whether the firm experiences an increase in future ETR (*ETR_INCREASE* or C*ETR_INCREASE*).

3.5 Testing the Moderating Effect of Financial Statement Readability on the Relation Between Tax Avoidance and Bank Loan Spread

Hypothesis 3 predicts that the relation between TA and cost of debt is less positive for firms with less readable financial statements. We test this hypothesis by estimating the following regression model using the subsample of firms with above-industry-median levels of TA:

$$SPREAD = c_0 + c_1\,(TA*BOG) + c_2\,TA + c_3\,BOG + c_j\,Control_j \qquad (3)$$

We focus on cost of bank loans in this test. Following Hasan et al. (2014), we define cost of bank loans as the loan spread (*SPREAD*), measured as all-in spread drawn in the DealScan database. We merge loan spread data with financial statement data using the

DealScan-Compustat link file provided by Michael Roberts (see Chava and Roberts 2008).[11] Similar to Hasan et al. (2014) we treat each facility-year as a distinct observation and control for a comprehensive list of firm and loan attributes in model (3). Definitions of these variables can be found in the Appendix (Panel F). In model (3), we expect the coefficient on the interaction term *TA*BOG*, c_1, to be negative and significant, consistent with Hypothesis 3 that less readable financial statements mitigate the impact of TA on cost of debt.

4. Results

4.1 Tax Avoidance and Financial Statement Readability

Table 3 presents the estimation results for model (1) using the full sample and subsamples of firms with high and low TA. Columns (1) to (3) show the estimation results using GAAP ETR, while columns (4) to (6) report the results using cash ETR as the TA measure.

For the full sample, the coefficients on *TA_ETR* and *TA_CETR* in columns (1) and (4) are positive and statistically significant at less than one-percent level. These results suggest that, on average, higher TA is associated with lower FSR (i.e., higher *Bog Index*). However, subsample analysis indicates that the negative relation between TA and FSR only holds among firms with relatively high levels of TA (columns 2 and 5), while it is positive among firms with relatively low levels of TA (columns 3 and 6). The differences in the estimated coefficients on *TA_ETR* and *TA_CETR* between high- and low-TA subsamples are statistically significant at less than one-percent level (chi-square statistics are 92.23 for *TA_ETR* and 50.74 for *TA_CETR*). These results together are consistent with Hypothesis 1. Thus, firms with high TA tend to have less readable financial statements than firms with medium TA, consistent with

[11] The DealScan-Compustat link data can be found at: http://finance.wharton.upenn.edu/~mrrobert/styled-9/styled-12/index.html. We thank Michael R. Roberts for making the data available.

high-TA firms having incentives to obfuscate information in order to reduce the likelihood that their TA strategies are challenged and overturned by tax authorities. In contrast, firms with medium TA tend to have more readable financial statements than firms with low TA. For medium-TA firms, the benefits from providing more transparent information to investors likely outweigh the benefits from obfuscating information provided to tax authorities. These results are consistent with the findings in Inger et al. (2018) using tax footnotes.

4.1.1 Controlling for the Length of Financial Statement

In this section we examine whether the positive relation between FSR and TA still holds after controlling for the length of financial statements. In our untabulated test we observe that firms with more TA have longer financial statements (as measured by file size and number of words in the 10-K).[12] Although longer documents are commonly assumed to be less readable (Li 2008; Laughran and McDonald 2014), it is plausible that longer financial statements indicate managers' efforts to provide more disclosure so as to mitigate information asymmetry (Balakrishnan et al. 2019). We, therefore, do not use the length of financial statements as a measure of FSR. Nonetheless, we examine whether the association between TA and FSR is driven by the fact that firms with more TA provide more disclosure, and more disclosure is associated with higher *Bog Index*.

Table 4 reports the estimation results for model (1) with one additional control variable – financial statement length. We present the results using the number of words in 10-Ks, the results using the 10-K file size are similar. The coefficients on *TA_ETR* and *TA_CETR* remain positive and statistically significant for firms with above-industry-median TA. For firms with below-industry-median TA, the coefficients are negative, but only statistically significant at conventional levels for *TA_CETR*. Collectively, our findings suggest that firms with high TA

[12] The 10-K file size and number of words data can be found at: https://sraf.nd.edu/textual-analysis/resources/#LM_10X_Summaries. We thank Bill McDonald for making the data available.

not only produce longer financial statements but also make the content of their financial statements more difficult to read.

4.1.2 Addressing Endogeneity Concerns

To address the concern that the association between TA and FSR we documented earlier could be driven by some factors that are omitted from model (1), we conduct two additional tests. First, we estimate model (1) with firm fixed effects to control for any omitted time-invariant firm characteristics that affect both FSR and TA. Results for this test (presented in Table 5, Panel A) confirm our baseline findings that TA is positively related to the Bog Index for firms with relatively high levels of TA, while the relationship is negative for firms with relatively low levels of TA. These results further support Hypothesis 1.

Second, we follow Balakrishnan et al. (2019) and Dyreng, Jacob, Jiang, and Muller (2019) and exploit the "Check-the-Box" regulation (CTB) as a quasi-natural experiment to establish a causal impact of TA on FSR. The CTB regulation, effective January 1, 1997, provides companies that have foreign subsidiaries an opportunity to reduce taxes by obtaining a "disregarded entity status" for their foreign subsidiaries. Therefore, we utilize the CTB regulation as a clean setting to provide evidence on how exogenous increases in TA lead to changes in FSR.

Similar to Balakrishnan et al. (2019), we conduct our test using observations from the 1994-2000 period, which includes three years before and after the event year, 1997. Following Balakrishnan et al. (2019), treated firms are those with non-zero pre-tax foreign income (pifo#0) and control firms include those with zero or missing pre-tax foreign income.[13] To further control for fundamental differences between multinationals and domestic firms before

[13] We find consistent results if we define treated firms as those with positive foreign income (pifo>0) as in Dyreng et al. (2019).

the CTB event, we use the propensity score matching (PSM) technique to identify matched firms. First, we run a logit regression of the *TREATED* dummy on all control variables as per Table 3 (these variables are measured in year 1996) and obtain predicted values (propensity scores). We then match each treated firm to one control firm in the same industry (3-digit SIC codes) that has the nearest propensity score (within 0.01 caliper) with no replacement. This procedure yields a sample of 227 distinct pairs of firms with insignificant differences across all control characteristics used in the logit regression, confirming that control firms are similar to treated firms before the CTB event (*t*-test results are available upon request).

To provide evidence that firms that increase TA following the CTB regulation also reduce readability in financial statements, we estimate the following difference-in-differences (DiD) regression model using the PSM-matched sample:

$$BOG = d_0 + d_1 (TREATED*POST) + d_j Control_j + d_i + d_t \quad (4)$$

TREATED is a dummy variable equal to one for treated firms and zero for control firms; *POST* is a dummy variable equal to one for the post-CTB period (1997-2000) and 0 for the pre-CTB period (1994-1996). $Control_j$ is a list of control variables used in the first-stage logit regression. We do not include the variables *TREATED* and *POST* separately in model (4) since the model includes firm fixed effects (d_i) and year fixed effects (d_t) which absorb the effects of the firm-based *TREATED* dummy and the year-based *POST* dummy, respectively.

Table 5, Panel B presents our DiD regression results. The variable of interest is the interaction term, *TREATED*POST*, whose coefficient captures the differential change in FSR post-CTB between treated firms and control firms. The estimated coefficient on the interaction term *TREATED*POST* is positive and statistically significant for the subsample of firms with above-industry-median TA. This means treated firms significantly decrease their FSR compared to control firms after the introduction of the CTB regulation. For firms with below-

industry-median TA, the association is negative but not statistically significant at conventional levels.

Overall, our results presented in Tables 3 through 5 provide robust evidence of a positive association between TA and FSR for firms with relatively high levels of TA (above the industry median). For firms with relatively low levels of TA (below the industry median), the association between TA and FSR is negative in general but not as robust.

4.1.3 The Relation between Tax Avoidance and Readability Conditional on Tax Audit Risk

In this section, we conduct a cross-sectional analysis to provide further evidence that managers of high-TA firms reduce FSR in order to mitigate tax audit risk. Specifically, we condition the relation between TA and FSR on the ex ante probability of tax audit. We argue that if managers make their financial statements less readable in order to obfuscate information about their TA strategies, they are more likely to do so when their firm is faced with a higher probability of tax audit. To test our prediction, we modify model (1) by adding to it a measure of IRS audit probability (*IRSAUDIT*) and its interaction with the TA measure. We follow the approach in Hoopes et al. (2012) to construct the *IRSAUDIT* variable, where higher values of *IRSAUDIT* indicate higher likelihood of IRS audit.[14] Results presented in Table 6 show positive and statistically significant coefficients on the interaction terms *TA_ETR*IRSAUDIT* and *TA_CETR*IRSAUDIT* for firms with above-industry-median TA. These results further support our argument that firms engaging in high levels of TA make their financial statements harder to read in order to mitigate the risk of being audited and challenged by tax authorities. Although we have no directional predictions on the coefficients on the interaction terms *TA_ETR*IRSAUDIT* and *TA_CETR*IRSAUDIT* for firms with below-industry-median TA, we

[14] The data on IRS audit probability can be found at: http://www.jeffreyhoopes.com/data.html. We thank Jeffrey L. Hoopes for making the data available.

present the results for completeness. We find that the coefficients on the interaction terms *TA_ETR*IRSAUDIT* and *TA_CETR*IRSAUDIT* are negative for firms with low TA.

4.2 Tax Avoidance, Financial Statement Readability, and Firm Risk

Table 7 presents the results for model (2), which tests Hypothesis 2. We first show evidence on the general impact of TA on firm risk in Table 7 Panel A. Note that in this test we focus on the subsample of firms with relatively high levels of tax avoidance (above the industry-year median). The positive and statistically significant coefficients on *ETR_LOW* and *CETR_LOW* in columns (1) and (2), respectively, indicate that firms in the bottom quintile of ETR (extremely high tax avoidance) have higher ETR persistence than those in higher quintiles of ETR (lower tax avoidance). This result is consistent with Guenther et al. (2017) and contradicts the prediction that high TA firms have less persistent ETR. However, the positive and significant coefficients on *ETR_LOW* and *CETR_LOW* in columns (3) and (4) suggest that firms in the bottom quintile of ETR are more likely to experience an increase in ETR in the future (i.e., moving to a higher ETR quintile) than firms in higher quintiles of ETR. Similarly, the positive and statistically significant coefficients on *TA_ETR* and *TA_CETR* in columns (5), (7) and (8) suggest that higher TA is associated with higher future ETR volatility (*FETR_VOL* and *FCETR_VOL*) and higher future stock return volatility (*FRET_VOL)*. In sum, we find evidence that, among firms with above-industry-median tax avoidance, higher TA is associated with higher firm risk.[15]

Table 7 Panel B presents the estimation results for model (2), which tests Hypothesis 2. The positive coefficients on *ETR_LOW*BOG* and *CETR_LOW*BOG* in columns (1) and (2),

[15] Guenther et al. (2017) find no evidence that TA is related to firm risk. We find the same results when using the full sample. However, as presented in Table 7 Panel A, for the subsample of firms with relatively high TA, the association between TA and firm risk is positive and statistically significant (except for the measure of ETR persistence). One possible explanation is that firms tend to focus on low risk tax planning strategies first before moving to high risk tax planning strategies (Dyreng et al. 2019). Therefore, only firms with extremely high levels of tax avoidance exhibit higher risk.

respectively, suggest that among firms in the bottom quintile of TA (extremely high tax avoiders), those with less readable financial statements have more persistent ETR (i.e., lower tax-related risk). Likewise, the negative coefficients on *ETR_LOW*BOG* and *ETR_LOW*BOG* in columns (3) and (4) and on *TA_ETR*BOG* and *TA_CETR*BOG* in columns (5) through (8) indicate that among firms with high TA, those with less readable financial statements experience lower likelihood of future ETR increase, lower future tax rate volatility as well as lower future stock return volatility. Our results suggest that having less readable financial statements helps reduce the impact of TA on firm risk, consistent with Hypothesis 2.

4.3 Tax Avoidance, Financial Statement Readability, and Cost of Bank Loans

Table 8 presents estimation results for model (3), which tests Hypothesis 3. The positive and statistically significant coefficients on *TA_ETR* and *TA_CETR* in columns (1) and (3), respectively, suggest that on average firms with higher TA have higher cost of bank loans, consistent with Hasan et al. (2014). Consistent with Hypothesis 3, the negative and statistically significant coefficients of *TA_ETR*BOG* and *TA_CETR*BOG* in columns (2) and (4), respectively, indicate that the positive relation between TA and cost of bank loans is weaker for firms with less readable financial statements. Consistent with Bonsall and Miller (2017), we find that less readable financial statements are associated with higher cost of bank loans when we run regression of loan spread on *BOG* alone without including *TA* and its interaction with *BOG* (for brevity these results are not presented). Thus, our results highlight the multidimensional effects of FSR on cost of debt. On the one hand, lower FSR directly increases cost of debt by increasing information asymmetry between firms and their lenders. On the other hand, lower FSR indirectly decreases cost of debt by reducing the cash flow risk caused by TA.

Overall, our findings suggest that, if firms with high TA make their financial statements less readable in order to reduce potential negative impacts of TA, they appear to be successful,

since the impacts of TA on firm risk and cost of debt are weaker for firms with less readable financial statements (although less readable financial statements also have a direct impact that increases cost of debt).

5. Conclusion

We provide evidence that firms with more TA tend to produce financial statements that are more difficult to read. This relationship holds primarily among firms with above industry median TA. The relationship between TA and FSR is more pronounced for firms faced with higher ex ante probability of tax audit, consistent with those firms having stronger incentive to take action to reduce the likelihood that their TA strategies are challenged and overturned by tax authorities. We also provide evidence that having less readable financial statements mitigates the adverse impacts of high TA on both firm risk and cost of debt. Inger et al. (2018) find that investors discount valuation of firms with high TA, but the discount is smaller for firms that have less readable tax footnotes. One possible explanation is that investors perceive that firms with less readable disclosures are less likely to be challenged by tax authorities, thus having lower risk. Our results confirm that, among firms with high TA, those with less readable financial statements do, indeed, have lower risk as well as lower cost of bank loans.

References

Altman EI (1968) Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*. 23(4):589-609.

Armstrong CS, Blouin JL, Jagolinzer AD (2015) Corporate governance, incentives, and tax avoidance. *Journal of Accounting and Economics*. 60(1):1-17.

Armstrong CS, Blouin JL, Larcker DF (2012) The incentives for tax planning. *Journal of Accounting and Economics*. 53(1-2):391–411.

Balakrishnan K, Blouin JL, Guay WR (2019) Tax Aggressiveness and Corporate Transparency. *The Accounting Review*. 94(1):45-69.

Badertscher B, Katz S, Rego S (2013) The separation of ownership and control and its impact on corporate tax avoidance. *Journal of Accounting and Economics.* 56(2-3):228–250.

Bauselinck B, Blanco B, Dhole S, Lobo G (2018) Financial Statement Readability and Tax Aggressiveness. Working Paper. Available at SSRN: https://ssrn.com/abstract=3261115.

Biddle GC, Hilary G, Verdi RS (2009) How does financial reporting quality relate to investment efficiency?. *Journal of Accounting and Economics*. 48(2-3):112–131.

Bonsall SB, Holzman E, Miller BP (2018) The Role of Management Talent in the Production of Informative Regulatory Filings. Working Paper. Available at SSRN: https://ssrn.com/abstract=3029669.

Bonsall SB, Leone AJ, Miller BP, Rennekamp K (2017) A plain English measure of financial reporting readability. *Journal of Accounting and Economics*. 63(2-3):329–357.

Bonsall SB, Miller BP (2017) The impact of narrative disclosure readability on bond ratings and the cost of debt. *Review of Accounting Studies.* 22(2):608–643.

Bozanic Z, Hoopes JL, Thornock JR, Williams BM (2017) IRS Attention. *Journal of Accounting Research*. *55*(1):79–114.

Cen L, Maydew EL, Zhang L, Zuo L (2017) Customer-supplier relationships and corporate tax avoidance. *Journal of Financial Ecnomics*. 123(2):377-394.

Chava S, Roberts MR (2008) How does financing impact investment? The role of debt covenants. *The Journal of Finance*. *63*(5):2085-2121.

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

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

Davis AK, Tama-Sweet I (2012) Managers' use of language across alternative disclosure outlets: Earnings press releases versus MD&A. *Contemporary Accounting Research.* 29(3):804–837.

Desai MA, Dharmapala D (2006) Corporate tax avoidance and high-powered incentives. *Journal of Financial Economics*. 79(1):145-179.

Dyreng SD, Hanlon M, Maydew EL (2008) Long-run corporate tax avoidance. *The Accounting Review*. 83(1):61–82.

Dyreng SD, Hanlon M, Maydew EL (2019) When Does Tax Avoidance Result in Tax

Uncertainty?. *The Accounting Review*. 94(2):179-203.

Dyreng SD, Hanlon M, Maydew EL, Thornock JR (2017) Changes in corporate effective tax rates over the past 25 years. *Journal of Financial Economics*. 124(3):441-463.

Dyreng S, Jacob M, Jiang X, Muller MA (2019) Tax incidence and tax avoidance. Working Paper. Available at SSRN: https://ssrn.com/abstract=3070239.

Erickson M, Hanlon M, Maydew EL (2004) How Much Will Firms Pay for Earnings That Do Not Exist? Evidence of Taxes Paid on Allegedly Fraudulent Earnings. *The Accounting Review*. 79(2):387-408.

Ertugrul M, Lei J, Qiu J, Wan C (2017) Annual Report Readability, Tone Ambiguity, and the Cost of Borrowing. *Journal of Financial and Quantitative Analysis*. 52(2):811–836.

Ehinger A, Lee JA, Stomberg B, Towery E (2017) Let's talk about tax: The determinants and consequences of income tax mentions during conference calls. Working Paper. Available at SSRN: https://ssrn.com/abstract=3009390.

Gaertner FB (2014) CEO After-Tax Compensation Incentives and Corporate Tax Avoidance. *Contemporary Accounting Research*. 31(4):1077–1102.

Graham JR, Hanlon M, Shevlin T, Shroff N (2014) Incentives for Tax Planning and Avoidance: Evidence from the Field. *The Accounting Review*. 89(3):991–1023.

Graham JR, Li S, Qiu J (2008) Corporate misreporting and bank loan contracting. *Journal of Financial Economics.* 89(1):44-61.

Guenther DA, Matsunaga SR, Williams BM (2017) Is Tax Avoidance Related to Firm Risk?. *The Accounting Review*. 92(1):115–136.

Hanlon M, Maydew EL, Saavedra D (2017) The taxman cometh: Does tax uncertainty affect corporate cash holdings?. *Review of Accounting Studies*. 22(3):1198-1228.

Hanlon M, Slemrod J (2009) What does tax aggressiveness signal? Evidence from stock price reactions to news about tax shelter involvement. *Journal of Public Economics*. 93(1-2):126–141.

Hasan I, Hoi CKS, Wu Q, Zhang H (2014) Beauty is in the eye of the beholder: The effect of corporate tax avoidance on the cost of bank loans. *Journal of Financial Economics*. 113(1):109–130.

Hasan I, Hoi CK, Wu Q, Zhang H (2017) Does social capital matter in corporate decisions? Evidence from corporate tax avoidance. *Journal of Accounting Research*. 55(3):629-668.

Hope OK, Ma MS, Thomas WB (2013) Tax avoidance and geographic earnings disclosure. *Journal of Accounting and Economics*. 56(2-3):170–189.

Hoopes JL, Mescall D, Pittman JA (2012) Do IRS Audits Deter Corporate Tax Avoidance?. *The Accounting Review*. 87(5):1603–1639.

Hwang B.H, Kim HH (2017) It pays to write well. *Journal of Financial Economics*. 124(2):373–394.

Inger KK, Meckfessel MD, Zhou M, Fan W (2018) An Examination of the Impact of Tax Avoidance on the Readability of Tax Footnotes. *The Journal of the American Taxation Association*. 40(1):1–29.

Khan M, Srinivasan S, Tan L (2017) Institutional Ownership and Corporate Tax Avoidance:

New Evidence. *The Accounting Review*. 92(2):101–122.

Koester A, Shevlin T, Wangerin D (2017) The role of managerial ability in corporate tax avoidance. *Management Science*. 63(10):3285-3310.

Kubick TR, Lockhart GB, Robinson JR (2014) Does Inside Debt Moderate Corporate Tax Avoidance?. Working Paper. Available at SSRN: https://ssrn.com/abstract=2451228.

Law KK, Mills LF (2015) Taxes and Financial Constraints: Evidence from Linguistic Cues. *Journal of Accounting Research*. 53(4):777–819.

Li C, Ma M, Omer T, Sun K (2017) How Does Tax Avoidance Affect Corporate Transparency. Working Paper, Singapore Management University.

Li F (2008) Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics*. 45(2-3):221–247.

Lo K, Ramos F, Rogo R (2017) Earnings management and annual report readability. *Journal of Accounting and Economics*. 63(1):1–25.

Loughran T, McDonald B (2014) Measuring Readability in Financial Disclosures. *The Journal of Finance*. 69(4):1643–1671.

Mills LF, Sansing RC (2000) Strategic Tax and Financial Reporting Decisions: Theory and Evidence. *Contemporary Accounting Research.* 17:85–106.

Mills LF, Robinson LA, Sansing RC (2010) FIN 48 and Tax Compliance. *The Accounting Review.* 85(5):1721–1742.

Phillips JD (2003) Corporate Tax-Planning Effectiveness: The Role of Compensation-Based Incentives. *The Accounting Review*. 78(3):847–874.

Rego SO, Wilson R (2012) Equity Risk Incentives and Corporate Tax Aggressiveness. *Journal of Accounting Research*. 50(3):775–810.

Schwab CM (2009) The Determinants and Effects of Voluntary Book-Tax Difference Disclosures: Evidence from Earnings Press Releases. Working Paper, University of Texas at Austin.

Shevlin TJ, Urcan O, Vasvari FP (2013) Corporate Tax Avoidance and Public Debt Costs. Working Paper. Available at SSRN: https://ssrn.com/abstract=2228601

Appendix. Variable Definitions

Variable	Definition
	Panel A: Financial Statement Readability Variables
BOG	Bog Index as developed in Bonsall et al. (2017)
FILESIZE	Log of gross 10-K file size as developed in Laughran and McDonald (2014)
NWORDS	Log of total number of words in the 10-K as developed in Li (2008)
	Panel B: Tax Avoidance Variables
TA_CETR	Cash effective tax rate (CETR) is cash tax paid (TXPD) divided by pretax income, which is measured as the difference between pretax book income (PI) and special items (SPI). We winsorize CETR to the range [0,1]. TA_CETR is defined as (-1) times CETR
TA_ETR	Effective tax rate (ETR) is total tax expense (TXT) divided by pretax income, which is measured as the difference between pretax book income (PI) and special items (SPI). We winsorize ETR to the range [0,1]. TA_ETR is defined as (-1) times ETR
	Panel C: Control Variables for Readability Regressions
EARN	Operating income after depreciation (OIADP) divided by total assets (AT)
SIZE	Log of market value of equity (PRCC_F*CSHO)
MTB	Market value of assets (PRCC_F*CSHO + AT - CEQ) divided by book value of assets (AT)
AGE	Number of years since a firm shows up in CRSP monthly stock return files
SI	Special items (SPI) scaled by total assets (AT)
RET_VOL	Standard deviation of the monthly stock returns in the last year
EARN_VOL	Standard deviation of the operating earnings in the last five years (-4,0)
NBSEG	Log of 1 plus the number of business segments
NGSEG	Log of 1 plus the number of geographic segments
NITEMS	Number of non-missing items in Compustat
MA	Dummy that equals 1 if a firm appears as an acquirer in current year in SDC Platinum M&A database, and 0 otherwise
SEO	Dummy that equals 1 if a firm has seasoned equity offering in current year according to SDC Global New Issues database and 0 otherwise
DLW	Dummy that equals 1 if a company is incorporated in Delaware, and 0 otherwise
	Panel D: Variable used in Cross-sectional Regressions
IRSAUDIT	IRS audit probability as developed in Hoopes et al. (2012)
	Panel E: Variables used in Firm Risk Regressions
ETR_PERSIST	Dummy that equals 1 if a firm remains in the same current ETR quintile in the subsequent period, and 0 if the firm moves to a higher or lower ETR quintile in the subsequent period. Current ETR is 5-year ETR over (-4,0) window while subsequent ETR is 5-year ETR over (1,5) window. Similar to Guenther et al. (2017), we require available data for each year in the (-4,0) and (1,5) windows.
CETR_PERSIST	Dummy that equals 1 if a firm remains in the same current CETR quintile in the subsequent period, and 0 if the firm moves to a higher or lower CETR quintile in the subsequent period. Current CETR is 5-year CETR over (-4,0) window while subsequent ETR is 5-year CETR over (1,5) window. Similar to Guenther et al. (2017), we require available data for each year in the (-4,0) and (1,5) windows.
ETR_INCREASE	Dummy that equals 1 if a firm moves to a higher ETR quintile in the subsequent period, and 0 if the firm remains in the same current ETR quintile or moves to a lower ETR quintile in the subsequent period. Current ETR is 5-year ETR over (-4,0) window while subsequent ETR is 5-year ETR over (1,5) window. Similar to Guenther et al. (2017), we require available data for each year in the (-4,0) and (1,5) windows.

CETR_INCREASE	Dummy that equals 1 if a firm moves to a higher CETR quintile in the subsequent period, and 0 if the firm remains in the same current CETR quintile or moves to a lower CETR quintile in the subsequent period. Current CETR is 5-year CETR over (-4,0) window while subsequent ETR is 5-year CETR over (1,5) window. Similar to Guenther et al. (2017), we require available data for each year in the (-4,0) and (1,5) windows.
FETR_VOL	Log of standard deviation of ETR in the next five years (1,5). Similar to Guenther et al. (2017), we require a positive numerator and denominator for each year to calculate the volatility of the ETR
FCETR_VOL	Log of standard deviation of CETR in the next five years (1,5). Similar to Guenther et al. (2017), we require a positive numerator and denominator for each year to calculate the volatility of the CETR
FRET_VOL	Log of standard deviation of monthly stock returns in the next year (t+1)
ETR_LOW	Dummy that equals 1 if a firm is in the lowest current ETR quintile, and 0 for higher quintiles, where current ETR is 5-year ETR over (-4,0) window
CETR_LOW	Dummy that equals 1 if a firm is in the lowest current CETR quintile, and 0 for higher quintiles, where current CETR is 5-year ETR over (-4,0) window
LAT	Log of total assets (AT)
PTBI	Pretax book income (PI) scaled by lagged total assets (AT)
LEVERAGE	Long-term debt (DLTT) scaled by lagged total assets
VOL_PTBI	Log of standard deviation of the ratio of annual pretax book income (PI) to lagged total assets (AT) measured over (-4,0) window
BTM	Book value of equity (CEQ) scaled by market value of equity (PRCC_F*CSHO)
SQR_AA	Square of year-end discretionary accruals, estimated using the modified Jones model from Dechow, Sloan, and Sweeney (1996)
VOL_SPI	Log of standard deviation of special items (SPI) scaled by lagged total assets (AT) measured over the last five years (-4,0)
VOL_OCF	Log of standard deviation of operating cash flow (OANCF) scaled by lagged total assets (AT) measured over the last five years (-4,0)
VOL_ETBSO	Log of standard deviation of the excess tax benefit of stock options (TXBCOF + TXBCO) scaled by lagged total assets (AT) measured over the last five years (-4,0); VOL_ETBSO is set to 0 if missing
ETBSO	Excess tax benefit of stock options (TXBCOF + TXBCO) scaled by lagged total assets (AT); ETBSO is set to 0 if missing
CHG_NOLCF	Change in net operating loss carryforward (TLCF) scaled by lagged total assets (AT); NOLCF is set equal to 0 if missing (TLCF)
NOLCF	Net operating loss carryforward (TLCF) scaled by lagged total assets (AT); NOLCF is set equal to 0 if missing (TLCF)
BHR	Annual buy-and-hold stock return measured over the current fiscal year
SHARES_OUT	Log of the firm's common shares outstanding (CSHO)
INST_OWN	Average institutional ownership measured over the fiscal year (source: Thomson-Reuters Institutional Holdings (13F) Database)
LOSS	Dummy that equals 1 if a firm has negative pretax income, and 0 otherwise

Panel F: Variables used in Cost of Bank Loan Regressions	
SPREAD	Log of loan spread that is measured as all-in spread drawn in the DealScan database in year (t). All-in spread drawn is defined as the amount the borrower pays in basis points over London Interbank Borrowing Rate (LIBOR) or LIBOR equivalent for each dollar drawn down
LAT	Log of total assets (AT) in year (t-1)
LEV	Sum of long-term debt (DLTT) and debt in current liabilities (LCT) scaled by total assets (AT) in year (t-1)
PPE	Net property, plant, and equity (PPENT) scaled by total assets (AT) in year (t-1)
CASH	Cash and marketable securities (CHE) divided by total assets (AT) in year (t-1)
ROA	Operating income before depreciation (OIBDP) scaled by total assets (AT) in year (t-1)

MTBE	Market-to-book ratio is measured as market value of equity (PRCC_F*CSHO), scaled by book value of equity (CEQ) in year (t-1)
SALE_GR	The percentage growth rate of sales (SALE) from two years prior (i.e., year (t-2)) to the year immediately before the year of loan inception (i.e., year (t-1))
VOL_EPS	Log of standard deviation of quarterly earnings in the last five years (-5,-1)
ZSCORE	Modified Altman's (1968) Z-score (Graham, Li, and Qiu 2008) in year (t-1). Z-score = (1.2*WCAP + 1.4*RE + 3.3*PI + 0.999*SALE)/AT, where WCAP is working capital, RE is retained earnings, and PI is pretax income, SALE is total sales, and AT is total assets. We use this modified Z-score, which does not include the ratio of market value of equity to book value of total debt, because a similar term, market-to-book (MTBE), enters our regressions as a separate control variable
LOANSIZE	Log of total amount of a loan facility (in millions of US dollars) obtained by a firm in year t
LOANMAT	Log of number of months to maturity of a loan facility obtained by a firm in year t
SYNDICATION	Dummy that equals 1 if the loan obtained by a firm in year t is syndicated, and 0 otherwise

Table 1. Sample Selection

Total number of non-financial and non-regulated firm-years with available Bog Index (1994-2015)	102,938
Less: Observations with negative pre-tax income ((pi-spi)<0)	43,208
Less: Observations with insufficient data on COMPUSTAT to calculate tax avoidance proxy (*TA_ETR*)	2,674
Less: Observations with insufficient data on to calculate all control variables in model (1)	11,170
Final Sample	45,886

Notes. All variables are defined in Appendix

Table 2. Summary Statistics

Variable	N	Mean	SD	P25	Median	P75
Panel A: Main Dependent Variables						
BOG	45,886	82.183	7.468	77.000	82.000	87.000
Panel B: Main Independent Variables						
TA_ETR	45,886	-0.288	0.182	-0.380	-0.324	-0.183
TA_CETR	44,233	-0.254	0.214	-0.350	-0.234	-0.088
Panel C: Control Variables						
EARN	45,886	0.109	0.074	0.060	0.096	0.143
SIZE	45,886	6.182	2.015	4.761	6.202	7.530
MTB	45,886	2.029	1.819	1.172	1.558	2.260
AGE	45,886	17.330	15.886	5.000	13.000	25.000
SI	45,886	-0.012	0.052	-0.009	0.000	0.000
RET_VOL	45,886	0.128	0.074	0.077	0.110	0.158
EARN_VOL	45,886	0.060	0.175	0.019	0.034	0.063
NBSEG	45,886	1.554	0.742	1.099	1.386	2.303
NGSEG	45,886	1.637	0.809	1.099	1.386	2.303
NITEMS	45,886	337.907	45.505	295.000	342.000	378.000
MA	45,886	0.388	0.487	0.000	0.000	1.000
SEO	45,886	0.096	0.294	0.000	0.000	0.000
DLW	45,886	0.599	0.490	0.000	1.000	1.000

Notes. Sample period is 1994-2015. All continuous variables are winsorized at the 1st and 99th percentiles. All variables are defined in Appendix.

Table 3. The Relation between Tax Avoidance and 10-K Readability

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.		*BOG*			*BOG*	
Sample	All TA	High TA	Low TA	All TA	High TA	Low TA
TA_ETR	**1.451*****	**4.703*****	**-1.763*****			
	[6.47]	[9.27]	[-4.79]			
TA_CETR				**1.325*****	**4.613*****	**-0.516***
				[7.00]	[7.07]	[-1.94]
EARN	-11.802***	-7.441***	-11.437***	-12.551***	-8.275***	-12.314***
	[-13.68]	[-7.75]	[-9.66]	[-13.88]	[-8.18]	[-10.46]
SIZE	0.718***	0.753***	0.715***	0.708***	0.766***	0.683***
	[14.43]	[13.77]	[11.41]	[13.88]	[13.81]	[10.57]
MTB	-0.035	-0.027	-0.083*	-0.033	-0.043	-0.095*
	[-1.07]	[-0.82]	[-1.75]	[-0.91]	[-1.24]	[-1.65]
AGE	-0.054***	-0.048***	-0.061***	-0.053***	-0.046***	-0.059***
	[-8.87]	[-7.40]	[-8.12]	[-8.50]	[-7.05]	[-7.80]
SI	-4.835***	-2.731***	-8.662***	-6.210***	-6.571***	-4.654***
	[-7.19]	[-3.66]	[-6.51]	[-9.12]	[-7.15]	[-5.08]
RET_VOL	8.686***	6.292***	10.236***	9.020***	6.293***	10.976***
	[12.59]	[8.03]	[10.38]	[12.63]	[7.93]	[10.68]
EARN_VOL	0.450	0.125	0.563	0.460	0.388	0.114
	[1.60]	[0.52]	[1.35]	[1.55]	[1.29]	[0.34]
NBSEG	0.714***	0.658***	0.790***	0.704***	0.637***	0.783***
	[6.92]	[5.86]	[6.10]	[6.76]	[5.51]	[6.18]
NGSEG	-0.112	-0.130	-0.160	-0.088	-0.049	-0.152
	[-1.22]	[-1.24]	[-1.44]	[-0.94]	[-0.48]	[-1.32]
NITEMS	0.004	0.003	0.005	0.007*	0.008	0.008
	[0.93]	[0.62]	[1.03]	[1.65]	[1.62]	[1.49]
MA	0.636***	0.694***	0.620***	0.641***	0.651***	0.700***
	[7.22]	[6.65]	[5.31]	[7.18]	[6.06]	[5.87]
SEO	0.499***	0.643***	0.195	0.410***	0.194	0.717***
	[4.30]	[4.24]	[1.26]	[3.49]	[1.36]	[4.17]
DLW	0.860***	0.848***	0.812***	0.800***	0.773***	0.725***
	[5.33]	[4.88]	[4.03]	[4.90]	[4.53]	[3.57]
Industry FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
N	45,886	23,313	22,572	44,233	22,483	21,749
Adj. R^2	0.386	0.406	0.373	0.389	0.387	0.398
High – Low		6.466***			5.129***	
Chi-square Stat.		92.23			50.74	

Notes. This table presents estimation results for model (1). We run separate regressions on full sample and subsamples of high and low tax avoidance (relative to industry-year median). Industry and year fixed effects are included in all models. All continuous variables are winsorized at the 1st and 99th percentiles. All variables are defined in Appendix. T-statistics based on standard errors clustered at firm level are reported in brackets. The *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 4. Controlling for 10-K Length

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.		*BOG*			*BOG*	
Sample	All TA	High TA	Low TA	All TA	High TA	Low TA
TA_ETR	**1.341*** **	**4.314*** **	**-1.427*** **			
	[5.87]	[8.31]	[-3.75]			
TA_CETR				**1.238*** **	**4.085*** **	**-0.445***
				[6.47]	[6.14]	[-1.65]
NWORDS	**0.730*** **	**0.653*** **	**0.739*** **	**0.727*** **	**0.598*** **	**0.792*** **
	[17.42]	[12.74]	[13.35]	[17.15]	[11.57]	[13.72]
EARN	-11.022***	-6.881***	-11.053***	-11.723***	-7.734***	-11.569***
	[-12.68]	[-7.15]	[-9.14]	[-12.88]	[-7.63]	[-9.65]
SIZE	0.617***	0.669***	0.605***	0.610***	0.688***	0.567***
	[12.32]	[12.12]	[9.48]	[11.84]	[12.27]	[8.59]
MTB	-0.030	-0.027	-0.072	-0.030	-0.051	-0.075
	[-0.91]	[-0.83]	[-1.41]	[-0.81]	[-1.47]	[-1.21]
AGE	-0.050***	-0.045***	-0.056***	-0.049***	-0.043***	-0.055***
	[-7.96]	[-6.80]	[-7.24]	[-7.64]	[-6.30]	[-7.04]
SI	-4.440***	-2.575***	-7.971***	-5.707***	-6.000***	-4.345***
	[-6.42]	[-3.33]	[-6.00]	[-8.19]	[-6.30]	[-4.67]
RET_VOL	7.887***	5.882***	9.148***	8.219***	5.846***	9.990***
	[11.49]	[7.49]	[9.27]	[11.57]	[7.33]	[9.73]
EARN_VOL	0.537*	0.261	0.558	0.580*	0.624**	0.083
	[1.82]	[1.11]	[1.30]	[1.78]	[2.09]	[0.29]
NBSEG	0.693***	0.654***	0.749***	0.682***	0.612***	0.768***
	[6.62]	[5.74]	[5.69]	[6.46]	[5.21]	[5.93]
NGSEG	-0.087	-0.095	-0.136	-0.073	-0.042	-0.133
	[-0.94]	[-0.91]	[-1.22]	[-0.77]	[-0.40]	[-1.15]
NITEMS	0.002	0.001	0.004	0.005	0.006	0.006
	[0.52]	[0.26]	[0.76]	[1.06]	[1.16]	[1.01]
MA	0.611***	0.681***	0.577***	0.620***	0.635***	0.676***
	[6.85]	[6.35]	[4.88]	[6.85]	[5.75]	[5.63]
SEO	0.404***	0.558***	0.123	0.298**	0.128	0.543***
	[3.40]	[3.56]	[0.77]	[2.49]	[0.87]	[3.10]
DLW	0.858***	0.831***	0.823***	0.793***	0.798***	0.698***
	[5.27]	[4.72]	[4.04]	[4.80]	[4.60]	[3.40]
Industry FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
N	43,045	21,871	21,173	41,511	21,092	20,418
Adj. R^2	0.386	0.402	0.376	0.389	0.384	0.399

Notes. This table presents estimation results for model (1) including 10-K length as an additional control variable. We run separate regressions on full sample and subsamples of high and low tax avoidance (relative to industry-year median). Industry and year fixed effects are included in all models. All continuous variables are winsorized at the 1st and 99th percentiles. All variables are defined in Appendix. T-statistics based on standard errors clustered at firm level are reported in brackets. The *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 5. Addressing Endogeneity Concerns

Panel A: Firm Fixed Effect Tests						
	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.	*BOG*					
Sample	All TA		High TA		Low TA	
TA_ETR	**0.246***		**1.248*****		**-0.202**	
	[1.74]		[3.60]		[-0.74]	
TA_CETR		**0.236***		**2.059*****		**-0.294***
		[1.96]		[4.61]		[-1.65]
EARN	-5.373***	-5.380***	-4.605***	-4.831***	-4.982***	-4.957***
	[-8.96]	[-8.47]	[-5.44]	[-6.01]	[-5.92]	[-5.53]
SIZE	0.078	0.082	0.090	0.130*	0.110	0.138
	[1.24]	[1.26]	[1.12]	[1.71]	[1.27]	[1.47]
MTB	-0.054***	-0.060***	-0.051**	-0.068***	-0.053**	-0.083***
	[-3.26]	[-3.18]	[-2.25]	[-2.77]	[-2.38]	[-3.09]
AGE	-0.100*	-0.096*	-0.124**	-0.151**	-0.045	-0.052
	[-1.91]	[-1.88]	[-2.22]	[-2.54]	[-0.57]	[-0.77]
SI	-2.276***	-2.539***	-1.791***	-2.650***	-4.338***	-2.130***
	[-5.63]	[-6.31]	[-3.49]	[-4.76]	[-4.84]	[-3.57]
RET_VOL	2.930***	3.009***	2.311***	2.250***	2.807***	3.766***
	[6.89]	[6.85]	[4.19]	[3.92]	[4.49]	[5.79]
EARN_VOL	0.254	0.277	0.080	0.131	0.501	0.302
	[1.35]	[1.43]	[0.45]	[0.61]	[1.19]	[0.58]
NBSEG	0.517***	0.502***	0.471***	0.496***	0.533***	0.446***
	[6.38]	[6.11]	[4.79]	[4.56]	[4.83]	[4.62]
NGSEG	0.013	0.000	0.090	-0.051	-0.086	0.002
	[0.17]	[0.01]	[0.93]	[-0.53]	[-0.87]	[0.02]
NITEMS	0.004	0.005	0.010***	0.009**	0.002	0.001
	[1.47]	[1.60]	[2.76]	[2.42]	[0.63]	[0.36]
MA	0.411***	0.403***	0.405***	0.342***	0.383***	0.421***
	[8.57]	[8.25]	[6.08]	[5.00]	[5.65]	[6.06]
SEO	0.231***	0.247***	0.143	0.009	0.231**	0.533***
	[3.10]	[3.22]	[1.35]	[0.09]	[2.06]	[4.35]
Firm FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
N	44,754	43,168	21,920	21,178	21,125	20,324
Adj. R^2	0.787	0.788	0.792	0.789	0.799	0.800

Table 5. Addressing Endogeneity Concerns

Panel B: Difference-in-Differences Tests

	(1)	(2)	(3)
Dep. Var.	*BOG*		
Sample	All TA	High TA	Low TA
TREATED*POST	**0.754***	**1.872*****	**-0.097**
	[1.70]	[2.85]	[-0.14]
EARN	-1.862	-4.757	-0.043
	[-0.84]	[-1.42]	[-0.01]
SIZE	-0.460*	-0.130	-0.804*
	[-1.76]	[-0.37]	[-1.91]
MTB	-0.010	-0.090	0.057
	[-0.22]	[-1.48]	[0.73]
AGE	-0.489	0.642	-1.479***
	[-0.94]	[1.48]	[-5.12]
SI	-5.917***	-3.389	-14.784***
	[-3.24]	[-1.36]	[-2.61]
RET_VOL	2.469	2.269	3.768
	[1.24]	[0.70]	[1.43]
EARN_VOL	-1.569	8.812*	-1.483
	[-0.38]	[1.80]	[-0.23]
NBSEG	0.732**	0.989*	0.193
	[2.07]	[1.65]	[0.42]
NGSEG	0.211	-0.050	0.264
	[0.73]	[-0.12]	[0.59]
NITEMS	-0.007	0.024	-0.043*
	[-0.49]	[1.27]	[-1.80]
MA	0.331	0.213	0.113
	[1.42]	[0.63]	[0.30]
SEO	0.052	-0.823*	0.460
	[0.15]	[-1.75]	[0.80]
Firm FEs	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes
N	1,913	919	829
Adj. R^2	0.804	0.798	0.819

Notes. These tables present test results to address endogeneity concerns about the relation between tax avoidance and 10-K readability reported in Table 3. In Panel A, we control for firm fixed effects instead of industry fixed effects in the baseline models. In Panel B, we estimate a difference-in-differences model that exploits the 1997 "Check-the-Box" regulation which serves as a quasi-natural variation in tax planning opportunities. We run separate regressions on full sample and subsamples of high (higher than industry-year median) and low (lower than industry-year median) tax avoidance. Firm and year fixed effects are included in all regressions. All continuous variables are winsorized at the 1st and 99th percentiles. All variables are defined in Appendix. T-statistics based on standard errors clustered at firm level are reported in brackets. The *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 6. Cross-sectional Analysis

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.		*BOG*			*BOG*	
Sample	All TA	High TA	Low TA	All TA	High TA	Low TA
TA_ETR*IRSAUDIT	**3.607****	**11.667*****	**-7.813*****			
	[2.26]	[4.04]	[-3.03]			
TA_CETR*IRSAUDIT				**3.157****	**17.082*****	**-5.808*****
				[2.31]	[4.45]	[-3.23]
TA_ETR	0.531	1.503*	0.288			
	[1.22]	[1.65]	[0.38]			
TA_CETR				0.516	-0.297	1.011*
				[1.40]	[-0.25]	[1.95]
IRSAUDIT	1.178*	1.983**	-2.832**	1.125*	2.730***	-2.415**
	[1.69]	[2.42]	[-2.30]	[1.74]	[3.30]	[-2.53]
EARN	-11.789***	-7.397***	-11.437***	-12.502***	-8.204***	-12.394***
	[-13.70]	[-7.67]	[-9.66]	[-13.86]	[-8.07]	[-10.57]
SIZE	0.709***	0.754***	0.695***	0.688***	0.719***	0.692***
	[12.64]	[12.01]	[9.63]	[11.95]	[11.12]	[9.48]
MTB	-0.031	-0.022	-0.077	-0.027	-0.027	-0.096*
	[-0.93]	[-0.65]	[-1.57]	[-0.72]	[-0.74]	[-1.68]
AGE	-0.054***	-0.047***	-0.061***	-0.053***	-0.045***	-0.059***
	[-8.85]	[-7.23]	[-8.09]	[-8.47]	[-6.83]	[-7.80]
SI	-4.767***	-2.611***	-8.827***	-6.174***	-6.508***	-4.702***
	[-7.09]	[-3.51]	[-6.68]	[-9.07]	[-7.10]	[-5.13]
RET_VOL	8.743***	6.458***	10.239***	9.082***	6.466***	10.943***
	[12.68]	[8.28]	[10.38]	[12.71]	[8.15]	[10.65]
EARN_VOL	0.478*	0.175	0.545	0.481	0.439	0.098
	[1.66]	[0.71]	[1.26]	[1.60]	[1.43]	[0.29]
NBSEG	0.713***	0.656***	0.788***	0.703***	0.627***	0.784***
	[6.92]	[5.87]	[6.10]	[6.76]	[5.44]	[6.19]
NGSEG	-0.116	-0.152	-0.162	-0.089	-0.062	-0.153
	[-1.26]	[-1.46]	[-1.47]	[-0.95]	[-0.61]	[-1.33]
NITEMS	0.004	0.003	0.005	0.007*	0.008	0.008
	[0.94]	[0.62]	[1.05]	[1.69]	[1.64]	[1.42]
MA	0.634***	0.699***	0.624***	0.638***	0.653***	0.704***
	[7.20]	[6.72]	[5.35]	[7.15]	[6.09]	[5.91]
SEO	0.496***	0.619***	0.190	0.403***	0.176	0.717***
	[4.28]	[4.11]	[1.22]	[3.44]	[1.25]	[4.16]
DLW	0.863***	0.850***	0.815***	0.801***	0.771***	0.726***
	[5.35]	[4.90]	[4.05]	[4.91]	[4.52]	[3.58]
Industry FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
N	45,886	23,313	22,572	44,233	22,483	21,749
Adj. R^2	0.386	0.407	0.374	0.390	0.388	0.398

Notes. This table presents the relation between tax avoidance and 10-K readability, conditional on Internal Revenue Service audit probability. We run separate regressions on full sample and subsamples of high and low tax avoidance (relative to industry-year median). Industry and year fixed effects are included in all models. All continuous variables are winsorized at the 1st and 99th percentiles. All variables are defined in Appendix. T-statistics based on standard errors clustered at firm level are reported in brackets. The *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 7. Tax Avoidance, 10-K Readability and Firm Risk

Panel A: The Relation between Tax Avoidance and Firm Risk

Dep. Var.	(1) *ETR_ PERSIST*	(2) *CETR_ PERSIST*	(3) *ETR_ INCREASE*	(4) *CETR_ INCREASE*	(5) *FETR_ VOL*	(6) *FCETR_ VOL*	(7) *FRET_ VOL*	(8) *FRET_ VOL*
Sample	High TA	High TA	High TA	High TA	High TA	High TA	High TA	High TA
ETR_LOW	**0.048*** **		**0.325*** **					
	[3.18]		[23.12]					
CETR_LOW		**0.120*** **		**0.263*** **				
		[7.01]		[16.02]				
TA_ETR					**0.036** **		**0.009***	
					[2.38]		[1.92]	
TA_CETR						**-0.005**		**0.027*** **
						[-0.19]		[5.65]
LAT					0.001*	-0.008***	-0.011***	-0.010***
					[1.68]	[-5.85]	[-17.83]	[-16.58]
PTBI					-0.108***	-0.151***	-0.028***	-0.009
					[-8.12]	[-6.28]	[-4.83]	[-1.63]
LEVERAGE					0.007	0.013	0.025***	0.022***
					[1.15]	[1.08]	[11.46]	[10.05]
VOL_PTBI					-0.045	-0.073	0.151***	0.138***
					[-1.18]	[-1.00]	[10.01]	[8.98]
BTM					0.014***	0.037***	0.020***	0.019***
					[3.05]	[4.35]	[13.33]	[12.47]
SQR_AA					0.026	-0.067*	0.009	0.015
					[1.03]	[-1.89]	[0.70]	[1.32]
VOL_SPI					0.113**	0.132	-0.024	-0.015
					[1.97]	[1.35]	[-1.18]	[-0.76]
VOL_OCF					0.017	0.173***	0.118***	0.104***
					[0.52]	[2.60]	[8.40]	[7.66]
VOL_ETBSO					-1.504**	0.025	-0.454	-0.450
					[-2.06]	[0.02]	[-1.53]	[-1.56]
ETBSO					1.556	-2.321	-1.363***	-1.079**
					[1.04]	[-1.02]	[-2.66]	[-2.23]
CHG_NOLCF					0.092*	0.028	0.028*	0.051***
					[1.68]	[0.35]	[1.89]	[3.61]
NOLCF					0.042***	-0.006	0.001	0.004
					[2.79]	[-0.30]	[0.18]	[1.31]
BHR							0.007***	0.009***
							[6.25]	[8.75]
SHARES_OUT							0.009***	0.007***
							[11.88]	[10.21]
INST_OWN							0.010***	0.008***
							[6.18]	[4.75]
LOSS							0.017***	0.019***
							[8.69]	[8.21]
Industry FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	7,226	6,846	7,226	6,846	5,793	6,090	12,416	12,275
Adj. R^2	0.109	0.081	0.109	0.081	0.095	0.074	0.408	0.397

Table 7. Tax Avoidance, 10-K Readability and Firm Risk

Panel B: The Relation between Tax Avoidance and Firm Risk, conditional on 10-K Readability

Dep. Var.	(1) *ETR_ PERSIST*	(2) *CETR_ PERSIST*	(3) *ETR_ INCREASE*	(4) *CETR_ INCREASE*	(5) *FETR_ VOL*	(6) *FCETR_ VOL*	(7) *FRET_ VOL*	(8) *FRET_ VOL*
Samples	High TA	High TA	High TA	High TA	High TA	High TA	High TA	High TA
ETR_LOW*BOG	**0.008***** [4.31]		**-0.005**** [-2.53]					
CETR_LOW*BOG		**0.004** [1.61]		**-0.006**** [-2.48]				
TA_ETR*BOG					**-0.004*** [-1.91]		**-0.001***** [-3.11]	
TA_CETR*BOG						**-0.007***** [-2.92]		**-0.001**** [-2.53]
ETR_LOW	-0.650*** [-4.00]		0.722*** [4.57]					
CETR_LOW		-0.192 [-0.98]		0.728*** [3.86]				
TA_ETR					0.329** [2.11]		0.128*** [3.27]	
TA_CETR						0.591*** [2.87]		0.133*** [3.10]
BOG	-0.003*** [-2.94]	-0.002** [-2.14]	0.001 [0.96]	0.001 [0.64]	-0.001 [-0.94]	-0.000 [-0.75]	0.000 [0.86]	0.000** [2.07]
LAT					0.001 [1.33]	-0.008*** [-6.23]	-0.011*** [-17.84]	-0.010*** [-16.75]
PTBI					-0.106*** [-7.91]	-0.146*** [-6.12]	-0.025*** [-4.41]	-0.007 [-1.31]
LEVERAGE					0.007 [1.03]	0.010 [0.85]	0.024*** [10.94]	0.021*** [9.64]
VOL_PTBI					-0.046 [-1.20]	-0.069 [-0.94]	0.150*** [9.99]	0.137*** [8.93]
BTM					0.014*** [3.12]	0.037*** [4.34]	0.020*** [13.25]	0.019*** [12.39]
SQR_AA					0.026 [1.04]	-0.064* [-1.81]	0.010 [0.76]	0.017 [1.45]
VOL_SPI					0.103* [1.78]	0.106 [1.07]	-0.032 [-1.59]	-0.025 [-1.24]
VOL_OCF					0.009 [0.27]	0.157** [2.36]	0.115*** [8.20]	0.101*** [7.47]
VOL_ETBSO					-1.481** [-2.01]	-0.125 [-0.12]	-0.492* [-1.66]	-0.508* [-1.76]
ETBSO					1.397 [0.93]	-2.590 [-1.14]	-1.480*** [-2.89]	-1.108** [-2.29]
CHG_NOLCF					0.091* [1.66]	0.029 [0.36]	0.029* [1.91]	0.050*** [3.55]
NOLCF					0.042*** [2.84]	-0.005 [-0.22]	0.000 [0.13]	0.004 [1.22]
BHR							0.007*** [6.07]	0.009*** [8.64]
SHARES_OUT							0.008*** [11.49]	0.007*** [9.96]
INST_OWN							0.009*** [5.59]	0.007*** [4.21]
LOSS							0.017*** [8.82]	0.019*** [8.13]
Industry FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	7,226	6,846	7,226	6,846	5,793	6,090	12,416	12,275
Adj. R^2	0.110	0.081	0.110	0.081	0.096	0.076	0.410	0.400

Notes. Panel A presents estimation results for model (2) without controlling for the Bog Index and its interaction with TA, while Panel B presents estimation results for the full model (2). We run regressions on the subsample of high TA (above industry-year median). Note that while *BOG* in columns (1) through (4) in Panel B is measured as the average of *Bog Index* in the last five years (-4,0) to be consistent with the period over which *ETR_LOW*/*CETR_LOW and ETR_INCREASE/CETR_INCREASE* are calculated, *BOG* in columns (5) through (8) in Panel B is measured in current year *t* to be consistent with current *TA_ETR*/*TA_CETR*. Industry and year fixed effects are included in all models. All continuous variables are winsorized at the 1st and 99th percentiles. All variables are defined in Appendix. T-statistics based on standard errors clustered at firm level are reported in brackets. The *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 8. Tax Avoidance, 10-K Readability and Cost of Bank Loans

Samples	(1) High TA	(2) High TA	(3) High TA	(4) High TA
TA_ETR	**0.328*****	**1.969*****		
	[7.02]	[4.01]		
TA_ETR*BOG		**-0.020*****		
		[-3.39]		
TA_CETR			**0.489*****	**2.054*****
			[7.24]	[2.97]
TA_CETR*BOG				**-0.019****
				[-2.31]
BOG		0.003*		0.002
		[1.75]		[1.36]
LAT	-0.091***	-0.096***	-0.077***	-0.080***
	[-11.68]	[-12.31]	[-10.24]	[-10.67]
LEV	0.312***	0.316***	0.372***	0.373***
	[10.01]	[10.12]	[12.39]	[12.45]
PPE	-0.061	-0.030	-0.013	0.007
	[-1.59]	[-0.77]	[-0.35]	[0.20]
CASH	-0.190***	-0.204***	-0.119**	-0.132**
	[-3.36]	[-3.62]	[-2.22]	[-2.45]
ROA	-1.064***	-1.061***	-0.937***	-0.929***
	[-11.18]	[-11.20]	[-9.95]	[-9.86]
MTBE	-0.004***	-0.004***	-0.006***	-0.006***
	[-3.01]	[-2.96]	[-3.90]	[-3.88]
SALE_GR	0.038***	0.034***	0.027***	0.023**
	[4.50]	[4.05]	[2.89]	[2.46]
VOL_EPS	0.020***	0.020***	0.016***	0.016***
	[4.80]	[4.59]	[3.92]	[3.91]
ZSCORE	-0.048***	-0.045***	-0.038***	-0.037***
	[-8.98]	[-8.54]	[-7.09]	[-6.83]
LOANSIZE	-0.076***	-0.076***	-0.088***	-0.088***
	[-9.82]	[-9.84]	[-11.82]	[-11.84]
LOANMAT	-0.038**	-0.037**	-0.008	-0.008
	[-2.35]	[-2.31]	[-0.47]	[-0.46]
SYNDICATION	-0.060**	-0.061**	-0.111***	-0.111***
	[-2.30]	[-2.36]	[-4.12]	[-4.12]
Credit Rating, Loan Purpose & Type FEs	Yes	Yes	Yes	Yes
Industry & Year FEs	Yes	Yes	Yes	Yes
N	7,889	7,889	7,661	7,661
Adj. R^2	0.713	0.715	0.692	0.693

Notes. This table presents estimation results for model (3). We estimate the model on the subsample of high-TA firms (above industry-year median). Credit rating, loan purpose, loan type, industry, and year fixed effects are included in all models. All continuous variables are winsorized at the 1st and 99th percentiles. All variables are defined in Appendix. T-statistics based on standard errors clustered at firm level are reported in brackets. The *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.