

**Digital Disclosure Complexity and Tax Aggressiveness: Insights from
U.S. XBRL Filings**

Allam, A., Elmarzouky, M., Moussa, T., Hawas, A. and Karim, E.

This article is © Emerald Publishing Limited and permission has been granted for this version to appear here: <http://westminsterresearch.westminster.ac.uk/>

The final, published version in the Journal of Accounting Literature and is available at:

<https://doi.org/10.1108/JAL-05-2025-0227>

This manuscript version is made available under the CC BY-NC 4.0 licence

<https://creativecommons.org/licenses/by-nc/4.0/>

The WestminsterResearch online digital archive at the University of Westminster aims to make the research output of the University available to a wider audience. Copyright and Moral Rights remain with the authors and/or copyright owners.

Digital Disclosure Complexity and Tax Aggressiveness: Insights from U.S. XBRL Filings

Amir Allam, Ph.D.

Aston University
Aston Business School
295 Aston Express Way
Birmingham B4 7ER, UK
Email: a.allam@aston.ac.uk
ORCID: 0000-0001-5082-7936

Mahmoud Elmarzouky, Ph.D.¹

University of St Andrews
University of St Andrews Business School
St Andrews KY16 9AJ, UK
Email: mahmoud.elmarzouky@st-andrews.ac.uk
ORCID: 0000-0001-9867-6057

Tantawy Moussa, Ph.D.

University of Westminster & Cairo University, Egypt
Westminster Business School
35 Marylebone Road
London NW1 5LS, UK
Email: t.moussa@westminster.ac.uk
ORCID: 0000-0001-5349-2764

Amira Hawas, Ph.D.

University of Sussex & Cairo University, Egypt
University of Sussex Business School
Sussex House, Falmer
Brighton, BN1 9RH, UK
Email: a.hawas@sussex.ac.uk
ORCID: 0000-0002-3766-5685

Enayet Karim, Ph.D.

Queen's University Belfast, Queen's Business School,
Belfast, Northern Ireland, UK
Email: e.karim@qub.ac.uk
ORCID: 0000-0001-5468-3271

¹ Corresponding author.

Digital Disclosure Complexity and Tax Aggressiveness: Insights from U.S. XBRL Filings

Abstract

Purpose – This study examines the relationship between corporate tax aggressiveness and the complexity of financial reporting using a structured, XBRL-based measure derived from machine-readable 10-K filings. While prior research has linked tax avoidance to reduced transparency using narrative-based readability metrics, this paper contributes by evaluating complexity from the preparer’s perspective in the digital reporting environment, offering new insights into disclosure behavior among tax-aggressive firms.

Design/methodology/approach – The analysis is based on 18,235 firm-year observations from U.S. publicly listed firms between 2011 and 2020. We employ fixed effects regression to address unobserved heterogeneity and conduct a battery of robustness tests to mitigate endogeneity concerns, including two-stage least squares (2SLS) and difference-in-differences (DiD) analyses.

Findings – The results show a positive and statistically significant relationship between tax avoidance and the complexity of XBRL filings. Tax-aggressive firms report more total XBRL tags and rely more heavily on firm-specific extended tags, suggesting a role for discretionary reporting behaviour beyond operational complexity. The relationship is stronger under higher IRS audit risk and is amplified among firms with greater financial capacity, consistent with higher-capacity firms being better able to sustain sophisticated tax strategies and the associated reporting complexity.

Implications – The findings have implications for regulators such as the SEC and FASB, underscoring the importance of monitoring the customization of digital disclosures. The extensive use of extended tags in tax-sensitive contexts may challenge comparability and hinder transparency, potentially affecting the efficacy of financial reporting standards.

Originality/value – Although the XBRL-based complexity measure has been introduced in earlier research, this study applies it in a novel context by linking it to corporate tax avoidance. The approach expands the understanding of reporting complexity beyond narrative readability and contributes to a deeper understanding of how digital reporting structures relate to tax planning behavior.

Keywords: Tax avoidance, financial reporting complexity, XBRL, digital disclosures, extended tags, corporate transparency

1. Introduction

Corporate tax avoidance has attracted increasing attention from regulators, policymakers, and the public, leading to intensified scrutiny of firms' tax-related activities. Estimates suggest that global corporate tax avoidance costs governments approximately USD 301 billion in lost tax revenue annually (Tax Justice Network, 2023). A substantial body of research documents that many firms actively adopt aggressive tax strategies to reduce their tax liabilities (Allam et al., 2023; Dyreng et al., 2008, 2013; Jaafar & Thornton, 2015; Balakrishnan et al., 2019; Tang et al., 2022; Luo et al., 2023; Lin et al., 2025; Yang et al., 2025). While such strategies may generate short-term cash flow benefits, they also expose firms to non-trivial costs, including enforcement risk, litigation exposure, reputational damage, and adverse market reactions (Hanlon & Heitzman, 2010; Balakrishnan et al., 2019). As tax authorities and market participants rely heavily on financial statements to assess tax compliance and risk (Bozanic et al., 2017; Chen et al., 2021), firms engaging in tax avoidance may face incentives not only to manage their tax outcomes, but also to manage the transparency of their financial disclosures.

These incentives have motivated a growing literature examining whether tax-avoidant firms exhibit more complex financial reporting. One prominent explanation is the management obfuscation hypothesis, which posits that managers may increase disclosure complexity to reduce transparency and obscure economically sensitive activities while remaining formally compliant with reporting requirements (Bloomfield, 2002; Li, 2008; de Souza et al., 2019). Under this view, disclosure complexity serves as an indirect mechanism through which firms limit users' ability to detect or fully evaluate opportunistic behaviour. An alternative perspective suggests that greater complexity may instead reflect the inherently intricate operations and transactions associated with aggressive tax planning, rather than intentional obfuscation (Balakrishnan et al., 2019). These competing explanations raise a fundamental

empirical question: does reporting complexity among tax-avoidant firms primarily reflect structural necessity, discretionary reporting choices, or a combination of both?

Prior empirical research has largely examined disclosure complexity using linguistic readability measures, such as the Fog Index (Gunning, 1952, 1969) and the Bog Index (Bonsall et al., 2017), as well as simpler proxies such as file length or word count. Studies using these measures generally find that firms with lower readability are more likely to engage in earnings management and tax avoidance and are harder for users to evaluate (You & Zhang, 2009; Lehavy et al., 2011; Beuselinck et al., 2018; Inger et al., 2018; Nguyen, 2020; Dao & Xu, 2024). However, the effectiveness of these measures in financial reporting contexts has been widely challenged. Loughran and McDonald (2014, p. 1644) argue that traditional readability indices are outdated and poorly specified for evaluating financial disclosures, as they often misclassify commonly used accounting and business terms - such as “corporation” or “operations” - as complex, even though such language is well understood by financial users. More fundamentally, readability indices evaluate the experience of the reader rather than the complexity of the accounting information itself, limiting their ability to capture the preparer-side discretion involved in constructing financial reports (Hoitash & Hoitash, 2018). As a result, existing evidence on the relationship between tax avoidance and disclosure complexity may be sensitive to measurement choice.

This study addresses these limitations by adopting a preparer-side, non-linguistic measure of accounting reporting complexity based on XBRL (eXtensible Business Reporting Language) disclosures. Specifically, we employ the Accounting Reporting Complexity (ARC) measure developed by Hoitash and Hoitash (2018), which captures the structure of financial statements using XBRL metadata rather than narrative text. Since 2009, the U.S. Securities and Exchange

Commission has required public firms to file financial reports in XBRL format using a standard taxonomy maintained by the Financial Accounting Standards Board (SEC, 2009). While firms are required to use standardised tags when available, they retain discretion to create firm-specific extended tags when existing taxonomy elements do not fully capture their disclosures. Prior research suggests that extended tags can reduce comparability and increase information processing costs for users (Dhole et al., 2015; Huang et al., 2019), highlighting the potential of XBRL structure to reveal differences in reporting discretion.

By focusing on XBRL-based reporting complexity, this study contributes to the literature in several important ways. First, we provide novel empirical evidence linking corporate tax avoidance to accounting reporting complexity measured from the preparer's perspective, offering an alternative to reader-side linguistic proxies. In doing so, we extend prior XBRL research - including Hoitash and Hoitash (2018) - by shifting attention from the consequences of reporting complexity to its behavioural determinants, specifically tax avoidance incentives. Second, we distinguish between complexity arising from standardised taxonomy tags and complexity driven by discretionary extended tags, providing new insight into whether reporting complexity reflects structural necessity or managerial discretion. To the best of our knowledge, no prior study directly links extended XBRL tag usage to corporate tax avoidance behaviour.

Third, by examining the role of IRS audit probability, we shed light on how enforcement risk shapes the relationship between tax planning and reporting structure. Prior research suggests that audit exposure may lead firms to adopt more conservative tax strategies (Hoopes et al., 2012; Antinozzi & Stanley, 2024), but it is unclear how it affects disclosure complexity. Fourth, this study contributes to the tax avoidance literature by offering more precise evidence on the relationship between tax avoidance and disclosure complexity. Prior narrative-based studies

report inconsistent findings on this relationship, partly because linguistic measures cannot cleanly separate discretionary reporting choices from unavoidable technical complexity. By employing a preparer-side, XBRL-based measure, we provide more direct evidence that tax avoidance is associated with deliberate structuring of digital disclosures, extending the tax avoidance literature with a more robust and granular test of this relationship.

Using a large sample of U.S. firms, we document a robust positive relationship between corporate tax avoidance and accounting reporting complexity. Firms with higher levels of tax avoidance use more XBRL tags overall and rely more heavily on discretionary extended tags. Subsample analyses reveal that this relationship is more pronounced among firms with high levels of tax avoidance, and that the effect is strongest for the extensions component, indicating that discretionary structuring choices within XBRL play a central role. In contrast, among low tax-avoidance firms, the relationship is weaker or negative, particularly for standardised taxonomy tags. Moreover, the positive relationship between tax avoidance and reporting complexity is stronger among firms facing higher IRS audit risk, suggesting that complexity does not dissipate under heightened scrutiny. Finally, we find that the relationship is amplified for firms with greater financial capacity, consistent with higher-capacity firms being better able to sustain more sophisticated tax strategies and the associated reporting complexity. These findings remain robust across alternative tax avoidance measures, cross-sectional analyses, and a range of additional robustness tests, including 2SLS and difference-in-differences (DiD) analyses.

The remainder of the paper is organised as follows. Section 2 develops the theoretical framework and hypotheses. Section 3 describes the research design. Section 4 presents the empirical results and robustness analyses. Section 5 concludes and discusses implications for research, regulation, and practice.

2. Literature review, theory and hypothesis development

2.1. Corporate tax avoidance and its incentives

Corporate tax avoidance refers to a variety of legal and strategic actions that reduce a firm's tax burden. Dyreng et al. (2008) define it as any action taken by management to lower the firm's cash effective tax rate over time. Hanlon and Heitzman (2010) describe it more broadly as the reduction of explicit tax liability and emphasize that tax avoidance encompasses a spectrum of behaviors, ranging from benign tax planning to highly aggressive activities that risk regulatory challenge. Examples of such strategies include shifting profits to tax havens, manipulating transfer prices, or registering intangible assets in low-tax jurisdictions (Sikka & Willmott, 2010; Hasan et al., 2014; Sikka, 2017).

Despite variation in how tax avoidance is defined, it is well established in the accounting literature that managers have strong incentives to reduce corporate tax liabilities. These incentives are driven by the potential to enhance after-tax income and satisfy shareholder demands (Armstrong et al., 2012; Cao et al., 2012; Rego & Wilson, 2012; Khan et al., 2017; Ming-Chin et al., 2020; Baker et al., 2023; Shim, 2024; Chandrasena et al., 2025). Accordingly, firms that engage in higher levels of tax avoidance may also seek to reduce the visibility of these actions by limiting or complicating the information they disclose. As noted by Nguyen (2020), such firms may adopt strategies to obscure tax-related activities and reduce the likelihood of regulatory or public scrutiny.

Tax enforcement risk is a key consideration for managers contemplating aggressive tax strategies. Graham et al. (2014) find that firms explicitly factor the likelihood of detection by tax authorities into their tax planning decisions. Public disclosures, particularly 10-K filings, are known to be reviewed by the Internal Revenue Service (IRS) for clues about tax planning

and potential compliance risks (Bozanic et al., 2017; Chen et al., 2021). Faced with this scrutiny, firms may choose between two approaches. One is to enhance transparency and preempt tax authority concerns. The other is to increase reporting complexity, either as a deliberate obfuscation tactic or as a reflection of the inherent intricacy of their operations.

Several studies support the idea that firms reduce transparency when pursuing aggressive tax strategies. For example, Hope et al. (2013) find that firms discontinue geographic earnings disclosures to conceal profit shifting. Balakrishnan et al. (2019) and Beuselinck et al. (2018) report that tax-aggressive firms tend to issue more complex financial statements, which may make it harder for users to identify tax-related behavior. Nguyen (2020) similarly finds that lower readability of financial statements is associated with higher tax avoidance. However, another line of research argues that complexity in reporting may be unintentional. Studies by Bushee et al. (2018), Dyer et al. (2017), and Guay et al. (2016) suggest that disclosure complexity often reflects the nature of a firm's operations or the diversity of applicable accounting standards. In this view, complexity is informational rather than strategic. In fact, Guay et al. (2016) find that some managers respond to this issue by voluntarily increasing disclosure clarity to mitigate the negative effects of complexity. These competing interpretations raise an important consideration for this study: whether the complexity observed in the financial disclosures of tax-avoiding firms primarily reflects the underlying intricacies of their operations, or whether it may also be influenced by the discretion managers have in structuring their reporting.

Notwithstanding this evidence, the measures used in prior research have important limitations. Loughran and McDonald (2014) argue that traditional readability indices such as the Fog Index are poorly specified for financial texts, as they mechanically classify technical accounting

terminology as complex even when such language is well understood by sophisticated users. More fundamentally, linguistic measures reflect the experience of the reader rather than the reporting decisions of the preparer, potentially conflating strategic obfuscation with unavoidable technical complexity (Loughran and McDonald, 2014; Hoitash and Hoitash, 2018). As a result, existing evidence on the relationship between tax avoidance and disclosure complexity may be sensitive to measurement choice.

These limitations have motivated a shift toward preparer-side, structural measures of reporting complexity, particularly in the context of digital financial reporting. The mandatory adoption of XBRL was intended to enhance transparency and comparability through the standardisation of accounting information. However, while firms are required to use predefined taxonomy elements when available, the reporting framework allows discretion through the creation of firm-specific extensions. Prior research shows that extensive use of extensions can reduce cross-firm comparability, increase information processing costs, and impair the usefulness of financial reports for users (Dhole et al., 2015; Kim et al., 2019; Huang et al., 2019; Salem et al., 2025). This highlights the potential of XBRL structure to reveal differences in reporting discretion, and motivates our adoption of the ARC measure developed by Hoitash and Hoitash (2018) as a preparer-side, non-linguistic proxy for accounting reporting complexity.

2.2. Theory and hypothesis development

The management obfuscation hypothesis provides the primary theoretical framework for this study. Originally articulated by Bloomfield (2002), it posits that managers facing heightened scrutiny or adverse incentives may respond by altering the *form* rather than the *content* of disclosure. In environments with mandatory reporting, managers retain discretion over how information is structured and presented, enabling them to comply with disclosure requirements

while increasing information processing costs for external users (Bloomfield, 2002; Guay et al., 2016). Empirical research supports this view: Li (2008) shows that firms with weaker performance issue less readable annual reports, and You and Zhang (2009) and Lehavy et al. (2011) find that greater complexity impairs investors' ability to process information. However, as Bushee et al. (2018) note, complexity also captures an information component - firms may increase detail to convey economically complex activities rather than to obscure them.

These insights extend naturally to digital financial reporting. The mandatory adoption of XBRL was intended to enhance transparency and comparability, but the framework permits discretion through firm-specific extensions. Prior research shows that extensive extension use reduces comparability, increases information processing costs, and can impair report usefulness (Dhole et al., 2015; Huang et al., 2019). If managers engaged in aggressive tax planning face incentives to limit the visibility of sensitive activities, they may increase reporting complexity through discretionary XBRL structuring choices while remaining formally compliant.

An alternative view holds that complexity need not reflect obfuscation. Firms engaged in sophisticated activities such as advanced tax planning may provide more detailed and structured reporting to clarify economic substance and demonstrate compliance (Guay et al., 2016; Peterson, 2012; Bushee et al., 2018). Under this interpretation, reporting complexity is informational rather than strategic, reflecting the inherent intricacy of complex transactions and the application of detailed standards.

Taken together, although part of reporting complexity may reflect legitimate disclosure responses to complex activities, the management obfuscation hypothesis predicts that firms

engaging in activities associated with heightened scrutiny will, on average, exhibit greater reporting complexity. We therefore hypothesize:

H1: Corporate tax avoidance is positively associated with accounting reporting complexity.

3. Research design

3.1. Data and sample

Our initial sample comprises 40,737 U.S. firm-year observations with available ARC data over 2011–2020 (Table 1, Panel A). We exclude financial services and utilities because these industries face distinct regulatory and reporting environments that imply different disclosure requirements and tax incentives (Dyreng et al., 2017; Moussa et al., 2020). We then remove 785 observations lacking the data required to compute the GAAP effective tax rate (ETR) and the cash effective tax rate (CETR). Consistent with prior research (e.g., Dyreng et al., 2017; Inger et al., 2018; Nguyen, 2020), we exclude 17,509 observations with negative pre-tax income because ETR-based measures are not economically interpretable when the denominator is negative. Finally, we drop 4,208 observations with missing values for control variables. The final sample used in the empirical analyses comprises 18,235 firm-year observations. Table 1, Panel B reports the industry distribution of the final sample. Tax avoidance and financial data are obtained from Compustat, and ARC data are obtained from Hoitash and Hoitash (2018).²

[Insert Table 1 here]

3.2. Empirical model and variables measurement

To examine whether corporate tax avoidance influences accounting reporting complexity, we estimate the following empirical model (Eq. 1) (where i and t denote firm and fiscal-year indicators, respectively):

² An updated set of data is available at: www.xbrlresearch.com/accounting-reporting-complexity/

$$\text{ARC}_{it} = \alpha_0 + \beta_1 \text{TA}_{it} + \sum_{i=1}^n \beta_i \text{CONTROLS}_{it} + \text{Industry FE} + \text{Year FE} + \varepsilon_{it} \quad (1)$$

We use fixed-effects panel regressions. This approach effectively mitigates unobserved heterogeneity across industries and time periods, enhancing the robustness of our findings (Hsiao, 2022). Using the fixed effects regression model is advantageous because it increases the consistency and efficiency of estimations, allowing for more precise inferences by controlling for omitted variable bias and unobserved heterogeneity among the sampled firms over time (Wooldridge, 2010). We also winsorize all continuous variables at the 1st and 99th percentiles to mitigate the influence of extreme values. Moreover, we use heteroskedasticity-robust standard errors clustered by firm and year to account for within-firm and within-year correlation in the error terms (Petersen, 2009).

The dependent variable for this study is Accounting Reporting Complexity (ARC), which captures the complexity of financial reporting from the preparer's perspective using XBRL-tagged disclosures. We adopt the ARC measure developed by Hoitash and Hoitash (2018), who conceptualise accounting complexity as the difficulty faced by preparers in applying authoritative accounting standards and communicating the economic substance of a firm's activities through financial reports. This notion is consistent with the Advisory Committee on Improvements to Financial Reporting (SEC, 2008), which emphasises that accounting complexity arises from both the volume and diversity of accounting standards and the information required to apply them.

ARC is constructed using XBRL data from Item 8 of firms' Form 10-K filings, which includes the financial statements and accompanying notes (Hoitash and Hoitash, 2018). In the XBRL

U.S. GAAP taxonomy, each accounting concept (e.g., revenues, deferred tax assets, lease obligations) is represented by a unique tag that is linked to authoritative GAAP guidance. Preparing and disclosing a larger number of distinct accounting concepts requires preparers to collect, categorise, analyse, and apply a broader and more diverse set of accounting standards, thereby increasing preparation complexity. Consistent with task complexity theory, greater information load and diversity increase the difficulty of judgment and execution, even in the absence of intentional obfuscation (Campbell, 1988; Bonner, 1994).

ARC is measured as the natural logarithm of the total number of distinct, non-repeating, monetary, annual XBRL tags reported in Item 8 of the 10-K. Tags that repeat within a filing (for example, the same accounting concept disclosed for multiple fiscal years) are counted only once, as repetition does not materially increase preparation difficulty. By focusing exclusively on monetary accounting tags, ARC isolates accounting reporting complexity from narrative or non-accounting content. This construction yields an objective, replicable, and comprehensive measure of firm-level accounting reporting complexity that exhibits both cross-sectional and within-firm variation.

Hoitash and Hoitash (2018) distinguish between standardised and discretionary components of ARC, reflecting taxonomy-based tags and firm-specific extensions. In line with this distinction, we separately examine the use of standardised and extended XBRL tags in our empirical analysis. Specifically, TAXONOMY is measured as the natural logarithm of the number of distinct monetary tags drawn from the U.S. GAAP taxonomy available at the time of filing. These tags are constrained by regulatory definitions and reflect complexity arising from standardised reporting requirements. EXTENSIONS is measured as the natural logarithm of the number of distinct monetary extended tags created by the firm. Because extensions are permitted only when no suitable taxonomy element exists, they reflect greater preparer discretion in structuring disclosures. Examining these components alongside the overall ARC

measure allows us to assess whether observed reporting complexity is more closely related to standardised reporting requirements or discretionary reporting choices within the XBRL framework.

Our main independent variable is corporate tax avoidance (TA). Consistent with prior literature (e.g., Cook et al., 2017; Dyreng et al., 2017; Hanlon & Heitzman, 2010; Inger et al., 2018; Law & Mills, 2015; Minnick & Noga, 2010; Nguyen, 2020; Wang et al., 2021), we employ two common measures of tax avoidance: GAAP ETR (effective tax rate) and CETR (cash effective tax rate). GAAP ETR is computed as total tax expense divided by pre-tax book income after excluding special items. CETR is computed as cash taxes paid divided by pre-tax book income after excluding special items. We winsorize ETR and CETR to lie between 0 and 1 and then multiply them by negative one. The resulting variables are denoted as TA_ETR and TA_CETR, where higher values indicate higher levels of tax avoidance.

We also include several control variables to reduce omitted variables bias by following prior studies (see, for example, Balakrishnan et al., 2019; Benamraoui et al., 2023; Ezeani et al., 2022; Ezeani et al., 2025; Hoitash & Hoitash, 2018; Moussa et al., 2026; Moussa et al., 2022; Nguyen, 2020; Usman et al., 2026; Kwabi et al., 2024). These include firm size (FSIZE), firm profitability (ROE), liquidity (LIQU), leverage (LEV), growth opportunities (MTB), earnings volatility (EVOL), mergers and acquisitions (MA), business segments (B_SEG), foreign segments (F_SEG), and the number of employees (EMP). In addition to firm fundamentals and operating complexity controls, we incorporate variables related to management characteristics and corporate governance in additional analyses and robustness checks. Specifically, we include CEO_TENURE (the number of years the Chief Executive Officer has served in their current role), CEO_PQF (a binary dummy equal to 1 if the CEO possesses a professional qualification, such as CPA or CFA, and 0 otherwise), and BETA (a measure of the sensitivity of a firm's stock returns to a benchmark index, estimated using a least squares regression based

on up to 60 monthly returns, with a minimum of 24). We also include a governance factor, GOV, constructed using Principal Component Analysis (PCA) of audit committee independence, compensation committee independence, and board independence metrics, with the PCA yielding a single factor (eigenvalue > 1) that captures overall corporate governance quality. Table 2 provides a list of variables and their definitions.

[Insert Table 2 here]

4. Empirical results and discussion

4.1. Descriptive statistics

Table 3 reports summary statistics for all variables in the full sample, as well as for subsamples of high and low tax-avoiding firms. On average, firms report approximately 231 distinct XBRL tags in their financial statements and footnote disclosures, with a standard deviation of 151. This indicates a substantial variation in accounting reporting complexity (ARC) across firms. Regarding the sub-components of ARC, the average number of standard taxonomy tags (TAXONOMY) is approximately 195, while the mean number of extended tags (EXTENSIONS) is about 36. Table 3 also shows that the mean (median) values of TA_ETR and TA_CETR are 0.193 (0.200) and 0.189 (0.170), respectively, which are consistent with previous findings in the literature (e.g., Balakrishnan et al., 2019; Cen et al., 2017; Dyreng et al., 2017; Koester et al., 2017; Inger et al., 2018; Nguyen, 2020; Plečnik & Wang, 2021; Yang et al., 2025).

Panel B of Table 3 presents the results of univariate tests comparing ARC measures between high and low tax-avoiding firms (based on quintiles). Firms in the highest tax-avoidance quintile exhibit significantly higher values for ARC, TAXONOMY, EXTENSIONS, and ALL-FACTS than those in the lowest quintile. For example, the mean (median) ARC value is approximately 31 percent (45 percent) higher for high tax-avoiding firms relative to the full-

sample average (median), while it is approximately 31 percent (37 percent) lower among low tax-avoiding firms. This preliminary result is consistent with prior studies indicating that higher tax avoidance is related to greater disclosure complexity (Balakrishnan et al., 2019; Inger et al., 2018; Nguyen, 2020). This complexity may arise from either the underlying operational structure of these firms or from discretionary decisions in the construction of financial disclosures.

[Insert Table 3 here]

Table 4 presents the Pearson correlation matrix for the variables used in the regression analyses. Both tax avoidance proxies (TA_ETR and TA_CETR) show positive and statistically significant correlations with ARC, TAXONOMY, EXTENSIONS, and ALL-FACTS, with coefficients ranging from 0.059 to 0.367. These results are consistent with our hypothesis that greater tax avoidance is related to more complex financial reporting. Furthermore, the correlation analysis reveals generally low correlations among the independent and control variables, suggesting that multicollinearity is unlikely to be a material issue.

[Insert Table 4 here]

4.2. Main results

Table 5 presents the results of our main regressions examining whether corporate tax avoidance influences accounting reporting complexity. Models (1)-(3) report specifications that use TA_ETR as the tax avoidance proxy, with ARC, EXTENSIONS, and TAXONOMY as dependent variables, respectively. Models (4)-(6) repeat the analysis using TA_CETR. All models include the full set of control variables as well as year and industry fixed effects.

Models (1) and (4) report the baseline results for overall reporting complexity (ARC). The coefficients on both tax avoidance proxies are positive and statistically significant (TA_ETR = 1.056, $p < 0.05$; TA_CETR = 1.072, $p < 0.05$), indicating that firms with higher tax avoidance

exhibit higher accounting reporting complexity in XBRL disclosures. The effects are economically meaningful. Using TA_ETR, a one-standard-deviation increase in tax avoidance (0.169) implies a 0.178 increase in $\log(\text{ARC})$, corresponding to an approximate 19.5% increase in the number of distinct XBRL tags. Using TA_CETR, a one-standard-deviation increase (0.168) implies a 0.180 increase in $\log(\text{ARC})$, corresponding to an approximate 19.7% increase. This result is consistent with prior evidence that firms with higher tax avoidance exhibit more complex disclosure environments (e.g., Balakrishnan et al., 2019; Inger et al., 2018; Nguyen, 2020) and supports the view that activities subject to heightened scrutiny and potential regulatory or reputational costs are associated with greater reporting complexity.

The decomposition of ARC into EXTENSIONS and TAXONOMY provides clearer insight into the channels underlying the main effect. Models (2) and (5) focus on EXTENSIONS, the discretionary component of XBRL reporting. The coefficients on tax avoidance are positive and highly significant for both proxies (TA_ETR = 1.696, $p < 0.01$; TA_CETR = 0.989, $p < 0.01$), indicating that higher tax avoidance is linked to greater use of firm-created tags. This result is consistent with the management obfuscation hypothesis (Bloomfield, 2002), which posits that managers can comply with disclosure requirements while increasing information processing costs through presentation and structuring choices (Bloomfield, 2002; Guay et al., 2016). It also aligns with XBRL evidence that extensive extension use reduces comparability and increases processing costs (Dhole et al., 2015; Huang et al., 2019; Kim et al., 2019). The economic magnitude is substantial: based on Model (2), a one standard deviation increase in TA_ETR implies a 0.286 increase in $\log(\text{EXTENSIONS})$, corresponding to an approximate 33.1% increase in extended tags. The stronger and more consistent effects observed for EXTENSIONS indicate that discretionary digital reporting choices play a central role in how reporting complexity rises with tax avoidance.

Models (3) and (6) report results for TAXONOMY, the standardised component of XBRL tagging. Model (3) shows a positive and significant coefficient for TA_ETR (0.756, $p < 0.05$), whereas Model (6) shows an insignificant coefficient for TA_CETR (0.028, $p > 0.10$). The economic magnitude is smaller than for EXTENSIONS. Using TA_ETR, a one-standard-deviation increase in tax avoidance corresponds to a 0.128 increase in $\log(\text{TAXONOMY})$, implying an approximate 13.6% increase in taxonomy tags. This finding is consistent with the structural complexity and information provision perspective, which views part of reporting complexity as reflecting economically intricate activities and the application of detailed standards to complex transactions (Peterson, 2012; Guay et al., 2016; Bushee et al., 2018). However, the stronger and more consistent effects observed for EXTENSIONS indicate that the tax avoidance–complexity relationship is not explained solely by underlying economic structure, and that discretionary tagging choices play an important role in shaping digital reporting complexity.

Taken together, the results presented in Table 5 provide evidence that corporate tax avoidance is related to higher accounting reporting complexity, with the strongest effects concentrated in the discretionary EXTENSIONS component. This provides empirical support for H1. The results also align with the management obfuscation hypothesis, suggesting that managers can increase the complexity of digital disclosures through discretionary reporting choices while remaining compliant with mandatory reporting requirements (Bloomfield, 2002; Guay et al., 2016). These findings are also consistent with prior evidence on tax avoidance and disclosure complexity (Balakrishnan et al., 2019; Inger et al., 2018; Nguyen, 2020) by showing that the relationship manifests directly in the digital architecture of machine-readable reports. In terms of control variables, the results are generally consistent with the prior literature (e.g.,

Balakrishnan et al., 2019; Dyreng et al., 2017; Hoitash & Hoitash, 2018; Nguyen, 2020): firm size (FSIZE), leverage (LEV), merger and acquisition activity (MA), and business segments (B_SEG) are positively related to reporting complexity, while liquidity (LIQU), profitability (ROE), growth opportunities (MTB), and foreign segments (F_SEG) are negatively related to complexity across specifications.

[Insert Table 5 here]

To add more depth to the main results, Table 6 provides a subsample analysis by separating firms into high and low tax-avoidance subsamples. This analysis allows us to assess whether the tax avoidance–complexity relationship differs across firms with materially different tax avoidance profiles. Across panels, the results show that the positive relationship is stronger among high tax-avoiding firms and is strongest for the discretionary EXTENSIONS component. In Panel A (ARC), the coefficient on TA_ETR is positive and statistically significant in the high-tax-avoidance subsample (Model (1): $TA_ETR = 3.275$, $p < 0.01$), while it is negative and statistically significant in the low-tax-avoidance subsample (Model (2): $TA_ETR = -1.377$, $p < 0.01$). This result indicates that the positive effect between ARC and tax avoidance is more pronounced among firms already operating at higher levels of tax avoidance, where incentives to manage visibility and perceived costs are likely stronger.

Panel B (EXTENSIONS) provides the clearest evidence: tax avoidance is positively and statistically significantly related to EXTENSIONS in the high-tax-avoidance subsample (Model (1): $TA_ETR = 2.971$, $p < 0.01$; Model (3): $TA_CETR = 2.491$, $p < 0.01$), but negative and statistically significant in the low-tax-avoidance subsample (Model (2): $TA_ETR = -3.582$, $p < 0.01$; Model (4): $TA_CETR = -2.603$, $p < 0.01$). This differential effect across subsamples is consistent with the management obfuscation hypothesis (Bloomfield, 2002): where tax avoidance is already elevated and scrutiny is likely higher, managers appear to rely

more on discretionary XBRL structuring choices, which can raise information-processing costs and reduce comparability (Guay et al., 2016; Dhole et al., 2015; Huang et al., 2019; Kim et al., 2019). The fact that the strongest subsample effects emerge precisely in EXTENSIONS supports the interpretation that discretionary design choices within XBRL, rather than only underlying transaction complexity, are an important mechanism linking tax avoidance to higher reporting complexity.

Panel C (TAXONOMY) further supports this interpretation, showing negative and statistically significant coefficients in the high-tax-avoidance subsample (Model (1): $TA_ETR = -1.039$, $p < 0.01$; Model (3): $TA_CETR = -0.489$, $p < 0.01$), while the corresponding coefficients in the low-tax-avoidance subsample are statistically insignificant. This result suggests that, among high tax-avoiding firms, increases in tax avoidance are consistent with reduced reliance on standardised taxonomy tags and greater reliance on firm-specific extensions, reinforcing the view that complexity increases through discretionary structuring of disclosures rather than through greater reliance on standardised reporting elements. This result is also consistent with prior studies (e.g., Inger et al., 2018; Nguyen, 2020), indicating that firms with lower levels of tax avoidance tend to exhibit less complex disclosure environments.

Overall, the results reinforce that the relationship between tax avoidance and digital reporting complexity is most pronounced among high tax-avoiding firms and is driven primarily by discretionary EXTENSIONS rather than standardised taxonomy tags.

[Insert Table 6 here]

4.3. Endogeneity and robustness checks

In this section, we perform a battery of endogeneity and robustness checks to assess the stability of the main findings. First, we re-estimate the baseline specifications using alternative measures of accounting reporting complexity and tax avoidance to ensure that the results are not driven by variable construction. Second, we incorporate controls for CEO characteristics, corporate governance, and market volatility to address potential omitted variable concerns. Third, we examine a lagged ARC–tax avoidance structure to reduce concerns about reverse timing. Fourth, we implement a two-stage least squares (2SLS) approach to mitigate endogeneity in tax avoidance. Finally, we apply a propensity score matched difference-in-differences (PSM-DID) design to strengthen identification by comparing firms exposed to the relevant setting with matched control firms.

4.3.1. Alternative measures of ARC and tax avoidance

To assess the robustness of the main results, we re-estimate the baseline specifications using alternative proxies for both accounting reporting complexity and tax avoidance. Table 7 reports the results of these robustness tests. First, to test whether the findings are sensitive to the measurement of accounting reporting complexity, following Hoitash and Hoitash (2018), we replace ARC with ALL-FACTS, which captures the total number of annual monetary XBRL tags reported in the financial statements and notes, including repeating facts (see Table 2 for definitions). Models (1) and (2) of Table 7 report results using ALL-FACTS as the dependent variable and show positive and statistically significant coefficients on TA_ETR (Model (1): 6.651, $p < 0.01$) and TA_CETR (Model (2): 3.094, $p < 0.10$), indicating that the tax avoidance–complexity relationship is robust to an alternative complexity proxy that incorporates both the breadth and frequency of tagging. These results are generally consistent with those in Table 5, suggesting that our evidence is robust across alternative measures of ARC and tax avoidance.

Second, to examine whether the findings depend on effective tax rate–based measures, we use TAX_HAVEN, defined as the natural log of the number of tax havens (Dyreng and Lindsey, 2009). Model (3) shows a positive and statistically significant coefficient (TAX_HAVEN = 4.251, $p < 0.01$), which is consistent with the baseline evidence that higher tax avoidance is linked to greater reporting complexity. Third, we employ TAX_AVOID3, which captures tax avoidance as the difference between the statutory tax on pre-tax book income before special items and current taxes paid, scaled by pre-tax income over a three-year window (Atwood et al., 2012; Kanagaretnam et al., 2018). Model (4) reports a positive and statistically significant coefficient (TAX_AVOID3 = 3.535, $p < 0.05$). Finally, we use DTAX (Frank et al., 2009), measured as the residual from industry-year regressions of permanent book–tax differences on non-discretionary determinants such as intangible assets and state taxes. Model (5) shows a positive and statistically significant coefficient (DTAX = 0.076, $p < 0.01$). Taken together, the results in Table 7 indicate that the positive relationship between corporate tax avoidance and accounting reporting complexity remains stable across alternative measures of reporting complexity and a broader set of tax avoidance proxies.

[Insert Table 7 here]

4.3.2. Additional controls for CEO characteristics, governance, and market volatility

As an additional robustness test, Table 8 re-estimates the baseline specifications after incorporating controls for CEO characteristics, corporate governance, and market risk, which prior studies show can shape the readability and quality of corporate reporting (e.g., Bilal et al., 2023; Sun et al., 2022; Moussa et al., 2023; Duong et al., 2024; Oradi et al., 2024; Kalelkar et al., 2024; Rahman and Kabir, 2024; Tantawy and Moussa, 2023). In particular, we include CEO tenure (CEO_TENURE) and CEO professional qualification (CEO_PQF) to capture CEO experience and financial expertise, which have been associated with variation in

disclosure readability and reporting choices. We include a governance factor (GOV) to capture board-level monitoring that has been linked to reporting transparency and readability. Finally, we control for market risk (BETA) to account for differences in firms' information environments and risk profiles that may coincide with reporting outcomes.

The key inferences remain stable. In Models (1)–(3), the coefficients on TA_ETR are positive and statistically significant across ARC, EXTENSIONS, and TAXONOMY (TA_ETR = 0.153, $p < 0.01$; 0.546, $p < 0.01$; 0.109, $p < 0.01$), indicating that the main findings persist after accounting for CEO and governance characteristics. In Models (4)–(6), TA_CETR remains positive and statistically significant for ARC and EXTENSIONS (TA_CETR = 0.070, $p < 0.05$; 0.310, $p < 0.01$), while the coefficient for TAXONOMY is positive but not statistically significant (TA_CETR = 0.044). Moreover, CEO_TENURE is negatively related to EXTENSIONS in both TA specifications, and CEO_PQF is negatively related to ARC and TAXONOMY, consistent with prior research that CEO attributes can shape financial reporting quality and readability (e.g., Duong et al., 2024; Oradi et al., 2024; Kalelkar et al., 2024; Rahman and Kabir, 2024). Overall, the results indicate that the positive tax avoidance–complexity relationship remains stable after accounting for managerial characteristics, governance quality, and market risk.

[Insert Table 8 here]

4.3.3. Lagged ARC- TA structure

To address potential endogeneity concerns where accounting reporting complexity and tax avoidance may be simultaneously determined, we estimate a lagged ARC–tax avoidance structure (Larcker & Rusticus, 2010). Specifically, we employ a dynamic specification by incorporating a one-year lag of the dependent variable (i.e., L.ARC, L.EXTENSIONS, and L.TAXONOMY) to account for the stochastic persistence of reporting complexity, and to

reduce concerns that the results reflect lingering effects of prior firm conditions not fully captured by the control variables, thereby helping attenuate concerns about reverse causality.

Table 9 reports the results and shows that the coefficients on the tax avoidance proxies remain positive and statistically significant across the main specifications, with the strongest and most consistent effects observed for EXTENSIONS. The lagged dependent variables are positive and highly significant in all specifications (e.g., L.ARC = 0.265, $p < 0.01$; L.EXTENSIONS = 0.419, $p < 0.01$; L.TAXONOMY = 0.291, $p < 0.01$), indicating strong persistence in reporting complexity over time. Importantly, the continued significance of the tax avoidance coefficients after controlling for lagged complexity suggests that the main results are not driven solely by time-invariant firm characteristics or routine persistence in reporting structure. This result reinforces that the positive tax avoidance–complexity relationship remains stable under a lagged specification and continues to be most pronounced in the discretionary extension tags component.

[Insert Table 9 here]

4.3.4. Two-Stage Least Squares (2SLS) regression

To further address endogeneity concerns that may arise from omitted time-varying factors and potential simultaneity between corporate tax avoidance and accounting reporting complexity, we employ a Two-Stage Least Squares (2SLS) instrumental-variable (IV) approach. 2SLS is widely used to mitigate omitted-variable bias and reduce the risk that the regressor of interest is correlated with the error term (Angrist & Krueger, 2001), and it is particularly relevant when reverse causality is a plausible concern (Larcker & Rusticus, 2010).

In line with prior studies (e.g., Cai et al., 2011; Banerjee et al., 2022) that treat peer or industry-year averages of the endogenous variable as a valid source of exogenous variation, we

instrument firm-level tax avoidance using the leave-one-out industry-year average tax avoidance (excluding the focal firm). This IV is expected to satisfy (i) relevance, because firms in the same industry-year face similar tax planning opportunities, enforcement environments, and peer influences that shape tax behaviour, and (ii) the exclusion restriction, because other firms' industry-year tax avoidance should not directly affect the focal firm's XBRL reporting complexity once we condition on firm controls and year and industry fixed effects.

Table 10 reports the first-stage and second-stage 2SLS estimates. In Panel A, the excluded instruments are positively and statistically significantly associated with the endogenous tax avoidance measures, supporting instrument relevance. In Panel B, Models (1)–(3) use instrumented TA_ETR as the endogenous tax avoidance proxy, while Models (4)–(6) use instrumented TA_CETR. Across all six second-stage models, the coefficients on the instrumented tax avoidance measures remain positive and statistically significant, consistent with the baseline results. This evidence supports the conclusion that the positive association between tax avoidance and XBRL reporting complexity is unlikely to be driven by endogenous variation in the tax avoidance proxies.

The IV diagnostics reported in Table 10 strongly support instrument relevance and strength. First, the underidentification tests (Kleibergen–Paap rk LM statistics) are large (100.026 for the TA_ETR specifications and 78.421 for the TA_CETR specifications), rejecting the null of underidentification and indicating that the model is identified. Second, the weak-instrument diagnostics indicate strong instrument relevance. The Cragg–Donald Wald F-statistics (106.485 for TA_ETR and 86.074 for TA_CETR) and the corresponding Kleibergen–Paap rk Wald F-statistics (101.061 and 84.057) comfortably exceed the Stock–Yogo critical value of 16.38 (10% maximal IV size), alleviating concerns about weak instruments (Stock & Yogo,

2005). Taken together, these diagnostics indicate that the instrument is sufficiently relevant and strong for the 2SLS estimation. Collectively, the 2SLS evidence reinforces the interpretation that corporate tax avoidance influences accounting reporting complexity, including complexity reflected in firms' XBRL tagging structure, and that the main results are not an artefact of reverse causality or omitted time-varying determinants of reporting outcomes.

[Insert Table 10 here]

4.3.5. Propensity Score Matched Difference-in-Differences (PSM-DID) model

To further address endogeneity concerns, we follow prior tax settings that exploit major regulatory shocks as quasi-natural experiments and implement a propensity score matched difference in differences design. Prior studies (e.g., Boubaker et al., 2016; Elmarzouky et al., 2025; Nguyen, 2020) show that combining propensity score matching with a DiD framework can sharpen identification by comparing treated and control firms that are similar immediately before a tax shock, and then examining their differential change after the reform.

In our setting, the 2017 Tax Cuts and Jobs Act (TCJA) provides a large, exogenous shock to U.S. corporate taxation, including a reduction in the statutory corporate rate and a major redesign of international tax rules. These changes are expected to alter the net benefits of sophisticated tax planning for firms with higher foreign exposure, and therefore offer a setting to test whether the link between tax avoidance incentives and digital reporting complexity reflects a causal response to tax incentives rather than a purely correlational association. The treatment group comprises firms with high foreign exposure (more directly affected by the international tax provisions), while the control group comprises firms with low foreign exposure.

To strengthen comparability between the two groups, we first apply propensity score matching using pre reform covariates (measured in 2016), excluding the treatment defining foreign segment measure (F_SEG) and the closely related segment count (B_SEG). The post match balance diagnostics in Panel A of Table 11 indicate that most matched covariates exhibit small and statistically insignificant mean differences between treatment and control firms, suggesting improved comparability in observable characteristics before the TCJA. Some differences remain (for example, in MA and segment measures), so we retain the full set of controls and fixed effects in the DiD estimation to further absorb residual observable heterogeneity.

Panel B of Table 11 reports the DiD estimates. The coefficient on the interaction term (POST*TREATMENT), which captures the differential post-TCJA change in reporting complexity for treated firms relative to control firms, is negative and statistically significant for ARC (POST*TREATMENT = -0.035 , $p < 0.10$) and TAXONOMY (POST*TREATMENT = -0.036 , $p < 0.05$), while it is negative but statistically insignificant for EXTENSIONS. These results indicate that, following the TCJA, firms with high foreign exposure experience a relative decline in digital reporting complexity compared to otherwise similar low-exposure firms. Taken together, the PSM-DiD evidence supports the interpretation that the tax avoidance–complexity relation is responsive to tax incentives, consistent with the view that at least part of XBRL-based complexity reflects discretionary reporting choices that adjust when the value of tax planning changes.

[Insert Table 11 here]

4.4. Additional analyses and underlying channels

4.4.1. The impact of IRS audit probability

To sharpen the mechanism behind the main results, we conduct a cross-sectional analysis motivated by the argument of tax audit risk mitigation. Prior research (e.g., Hoopes et al., 2012; Nguyen, 2020) suggests that tax enforcement risk may shape both tax planning and disclosure incentives. When audit likelihood is higher, managers face greater expected costs if tax positions are detected and challenged (including potential penalties, reputational costs, and managerial consequences). Importantly, tax authorities do not rely solely on private tax-return information: they can also use firms' public financial statements and 10-K disclosures to corroborate private claims and identify risk indicators (see, for example, Hoopes et al., 2012; Nguyen, 2020). For example, Nguyen (2020) finds that the negative relationship between tax avoidance and financial statement readability is stronger when ex ante audit probability is higher and in industries exposed to greater tax risk. Building on this enforcement-based argument, we test whether the tax avoidance–complexity relation varies with IRS audit probability in our XBRL setting.

Table 12 reports subsample regressions split by IRS audit probability (above versus below the median) and re-estimates the tax avoidance–ARC relation within each group. The results indicate that the positive relationship between tax avoidance and accounting reporting complexity is more pronounced when audit probability is higher. Using TA_ETR, the coefficient is 0.637 ($p < 0.01$) in the high-audit subsample versus 0.407 ($p < 0.05$) in the low-audit subsample. Using TA_CETR, the coefficient is 0.355 ($p < 0.05$) in the high-audit subsample versus 0.221 ($p < 0.10$) in the low-audit subsample. This result indicates that the tax avoidance–complexity relation persists under both high and low scrutiny but is more pronounced when expected enforcement is higher.

This result is consistent with the audit-risk mitigation mechanism: when the likelihood of IRS scrutiny increases, managers' expected costs of detection and challenge rise, strengthening incentives to manage visibility through the structure of disclosures while remaining compliant. In particular, higher enforcement risk can amplify incentives to increase information-processing costs for external users through the digital structure of reporting, consistent with the broader view that complexity can be strategically employed in disclosure settings. The finding also aligns with Nguyen (2020), which shows that the tax avoidance–disclosure relation strengthens with higher audit likelihood, supporting the view that disclosure properties respond to enforcement risk rather than reflecting only passive operational complexity. At the same time, the relation remains positive even among low-audit firms, which is compatible with the structural complexity view that some portion of ARC captures underlying transaction and reporting environment complexity linked to tax planning.

[Insert Table 12 here]

4.4.2. The impact of the firm's financial capacity

To explore firm financial capacity as a potential channel, we conduct a cross-sectional analysis that partitions firms by pre-tax income. Prior research (e.g., Rajabalizadeh, 2023; Nguyen, 2020) suggests that firms' financial conditions can shape managerial incentives and disclosure choices, and may therefore condition how strongly tax avoidance is reflected in accounting reporting complexity. Firms with stronger financial capacity (proxied by higher pre-tax income) may have greater incentives and resources to implement, sustain, and defend more complex tax positions, which can increase the need or willingness to structure digital disclosures in ways that raise information-processing costs. Conversely, financially constrained firms may face tighter limits on the scale and sophistication of tax planning, potentially attenuating the tax avoidance–complexity relation, particularly for cash-tax-based

measures that capture realised tax payments. For example, Rajabalizadeh (2023) finds that CEO overconfidence is associated with more complex financial reporting (lower readability), and that the relationship is stronger among firms with lower profitability, underscoring that financial conditions can affect incentives to increase complexity. In our setting, financial capacity may shape the intensity and sophistication of tax planning, and therefore the extent to which tax avoidance is reflected in more complex XBRL tagging.

Table 13 reports subsample regressions split by firm financial capacity (below versus above the median of pre-tax income) and re-estimates the tax avoidance–complexity relation within each group. The results indicate that the positive relationship between tax avoidance and reporting complexity is more pronounced among firms with higher financial capacity. In Panel A (low financial capacity), TA_ETR is positively and statistically significantly associated with ARC (0.153, $p < 0.01$), EXTENSIONS (0.443, $p < 0.01$), and TAXONOMY (0.127, $p < 0.01$). In contrast, the cash-based proxy is weaker in this subsample: TA_CETR is statistically insignificant for ARC (0.063) and TAXONOMY (0.039), although it remains positive and significant for EXTENSIONS (0.276, $p < 0.01$). In Panel B (high financial capacity), the relationship is stronger and more consistent across measures. Using TA_ETR, the coefficients are positive and significant for ARC (0.373, $p < 0.01$), EXTENSIONS (0.478, $p < 0.01$), and TAXONOMY (0.369, $p < 0.01$). Using TA_CETR, the coefficients remain positive for ARC (0.121, $p < 0.05$), EXTENSIONS (0.340, $p < 0.01$), and TAXONOMY (0.099, $p < 0.10$).

These findings suggest that financial capacity strengthens the tax avoidance–complexity relation. When firms have higher pre-tax income, tax avoidance is associated with higher reporting complexity under both effective and cash tax measures, consistent with the view that higher-capacity firms can pursue more sophisticated strategies and absorb the associated

reporting and compliance costs. This is consistent with a capacity-based channel in which stronger financial resources facilitate more intensive tax planning and expand the scope for complex digital disclosure structuring. In particular, extension tags remain positive and significant across both panels and both proxies, reinforcing the broader implication from the main results that the tax avoidance–complexity relation is particularly tied to discretionary extension tags rather than reflecting only mechanical increases in standardised tagging. At the same time, the persistence of positive coefficients even among low-capacity firms (particularly under TA_ETR) is compatible with the idea that some component of reporting complexity reflects underlying transaction and reporting-environment complexity that accompanies tax planning, even where discretionary capacity is more constrained.

[Insert Table 13 here]

4.4.3. The impact of CEO professional qualifications

To examine whether CEO financial expertise conditions the tax avoidance–complexity relation, Table 14 augments the baseline models by adding CEO professional qualifications (CEO_PQF) and the interaction terms TA*CEO_PQF. Prior research suggests that CEO attributes can influence financial reporting quality and readability through differences in expertise, monitoring, and disclosure choices (e.g., Allam et al., 2024; Duong et al., 2024; Ezeani et al., 2025; Ezeani et al., 2023; Oradi et al., 2024; Kalelkar et al., 2024; Rahman & Kabir, 2024). Conceptually, CEO professional qualifications could attenuate complexity if financially skilled CEOs favour clearer, more comparable reporting, or they could amplify complexity if expertise enables more sophisticated structuring of disclosures when incentives to manage visibility are strong.

Table 14 shows that the core tax avoidance effects remain positive and statistically significant across the main specifications, consistent with earlier results. However, the moderation results

do not indicate that CEO professional qualifications systematically change the strength of the tax avoidance–complexity relation. The interaction terms (TA*CEO_PQF) are statistically insignificant across outcomes for both tax avoidance proxies, implying that the positive relationship between tax avoidance and XBRL-based complexity is broadly similar whether or not the CEO holds a professional qualification.

At the same time, CEO_PQF itself is negative and statistically significant in some specifications, consistent with the argument that financially qualified CEOs are associated with higher financial reporting quality/readability (Duong et al., 2024; Oradi et al., 2024; Kalelkar et al., 2024; Rahman & Kabir, 2024). This result suggests that while CEO professional qualifications may relate to lower baseline complexity in certain dimensions, they do not materially weaken the positive tax avoidance–complexity relationship.

[Insert Table 14 here]

5. Conclusion

This study examines the relationship between corporate tax avoidance and accounting reporting complexity using a structured, XBRL-based measure derived from firms' 10-K filings. Moving beyond traditional narrative readability metrics, we adopt a preparer-side perspective and show that tax avoidance is systematically related to the digital architecture of financial reports. Across a large sample of U.S. firms, we document a robust positive relationship between tax avoidance and accounting reporting complexity, measured by the total number of distinct XBRL tags and, more importantly, by the discretionary use of firm-specific extension tags.

Our empirical results reveal several important patterns. First, the tax avoidance–complexity relationship is not uniform across firms. Subsample analyses show that the positive relationship is strongest among firms with high levels of tax avoidance, whereas it is weaker or even negative among low tax-avoidance firms. This heterogeneity is most pronounced in the extensions component of reporting complexity, indicating that discretionary structuring choices within XBRL play a central role in shaping complexity where incentives to manage visibility are likely strongest. In contrast, standardised taxonomy tags exhibit weaker and less consistent associations, suggesting that the documented effects cannot be explained solely by underlying operational or transactional complexity.

Second, we find that the relationship between tax avoidance and reporting complexity persists - and intensifies - under heightened enforcement risk. Firms facing a higher probability of IRS audit exhibit a stronger positive relationship between tax avoidance and reporting complexity, particularly through discretionary extensions. This finding is notable because increased scrutiny might otherwise be expected to encourage transparency. Instead, the evidence suggests that reporting complexity does not dissipate under audit risk, consistent with the view that managers may respond to higher expected enforcement costs by managing the structure of disclosures while remaining formally compliant.

Third, our analysis highlights the role of firm financial capacity as an important conditioning factor. The positive relationship between tax avoidance and reporting complexity is more pronounced among firms with higher pre-tax income, consistent with a capacity-based channel in which financially stronger firms are better able to sustain sophisticated tax strategies and absorb the associated reporting and compliance costs. While some evidence of increased complexity also appears among lower-capacity firms, particularly for effective tax rate-based

measures, the results suggest that financial resources amplify the extent to which tax avoidance is reflected in complex digital disclosure structures.

Collectively, these findings contribute to the literature in several ways. The study advances research on tax avoidance and disclosure by shifting attention from narrative readability to the structure of machine-readable financial reports, demonstrating that tax planning behaviour is reflected directly in the digital design of disclosures. By distinguishing between standardised taxonomy tags and discretionary extension tags, the analysis provides new insight into how reporting complexity arises from a combination of structural firm characteristics and managerial discretion. Moreover, by documenting heterogeneity across tax avoidance levels, audit risk, and financial capacity, the study refines existing debates on whether disclosure complexity reflects obfuscation, information provision, or both.

The results also carry important practical and policy implications. For regulators such as the SEC and standard-setters like the FASB, the systematic association between tax avoidance and the use of XBRL extensions raises questions about the current balance between flexibility and comparability in digital reporting. While extensions play a legitimate role in capturing firm-specific information, their disproportionate use in tax-sensitive contexts suggests a need for more targeted guidance, monitoring, or post-filing review of extension practices. Enhanced transparency around the justification for extensions, clearer constraints on their use, or greater standardisation of commonly extended items could improve comparability without undermining the benefits of XBRL.

For tax authorities, the findings suggest that XBRL-based complexity measures - particularly those capturing discretionary reporting choices - may serve as a useful supplementary

screening tool for risk assessment. Although accounting reporting complexity is not a direct measure of tax avoidance, its strong relationship with aggressive tax behaviour, persistence under audit risk, and amplification among high-capacity firms indicate that digital disclosure structure may convey incremental information relevant for enforcement prioritisation in resource-constrained environments.

As with all empirical research, this study has limitations that point to avenues for future work. The analysis focuses on U.S. firms operating under a specific tax and reporting regime, and the results may not generalise to jurisdictions with different disclosure mandates or enforcement environments. Future research could examine how international variation in XBRL implementation, taxonomy design, or regulatory oversight shapes the relationship between tax planning and reporting complexity. In addition, combining structural XBRL measures with advanced textual or machine-learning approaches may provide a more granular understanding of how complexity is constructed across both narrative and digital dimensions of financial reporting.

Acknowledgements

The authors gratefully acknowledge the Editor-in-Chief, Professor Tom Smith, and the anonymous reviewers for their thoughtful and constructive feedback.

For the purposes of open access, the authors have applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript (AAM) version arising from this submission.

Statements and Declarations

The authors declare that they have no conflict of interest.

The authors did not receive support from any organization for the submitted work.

The authors have no relevant financial or non-financial interests to disclose.

References

- Adams, M. T., Inger, K. K., Meckfessel, M. D., & Maher, J. J. (2022). Tax-Related Restatements and Tax Avoidance Behavior. *Journal of Accounting, Auditing & Finance*, 0148558X221115482. <https://doi.org/10.1177/0148558X221115482>
- Allam, A., Moussa, T. and Elmarzouky, M. (2024). Examining the relationship between CEO power and modern slavery disclosures: The moderating role of board gender diversity in UK companies. *Business Strategy and The Environment*. 33 (8), pp. 8067-8085. <https://doi.org/10.1002/bse.3910>
- Allam, A., Moussa, T., Abdelhady, M. and Yamen, A. (2023). National culture and tax evasion: The role of the institutional environment quality. *Journal of International Accounting, Auditing and Taxation*. 52 100559. <https://doi.org/10.1016/j.intaccaudtax.2023.100559>
- Angrist, J. D., & Krueger, A. B. (2001). Instrumental variables and the search for identification: From supply and demand to natural experiments. *Journal of Economic Perspectives*, 15(4), 69–85.
- Armstrong, C. S., Blouin, J. L., & Larcker, D. F. (2012). The incentives for tax planning. *Journal of Accounting and Economics*, 53(1–2), 391–411.
- Atwood, T. J., Drake, M. S., Myers, J. N., & Myers, L. A. (2012). Home country tax system characteristics and corporate tax avoidance: International evidence. *The Accounting Review*, 87(6), 1831–1860. <https://doi.org/10.2308/accr-50222>
- Baker, P. L., Lyu, P., & Perotti, P. (2023). Does tax avoidance impair accounting comparability? *Journal of Accounting Literature*. <https://www.emerald.com/insight/content/doi/10.1108/JAL-07-2022-0078/full/html>
- Balakrishnan, K., Blouin, J. L., & Guay, W. R. (2019). Tax aggressiveness and corporate transparency. *The Accounting Review*, 94(1), 45–69. <https://doi.org/10.2308/accr-52130>
- Banerjee, R., Gupta, K., & Krishnamurti, C. (2022). Does corrupt practice increase the implied cost of equity? *Journal of Corporate Finance*, 73, 102191.
- Benamraoui, A., Moussa, T. and Hussien, M. (2023). Islamic Banks' Sharia Compliance Disclosure: An International Evidence. *Accounting Research Journal*. 36 (4/5), pp. 327-348. <https://doi.org/10.1108/ARJ-07-2022-0157>
- Bilal, Komal, B., Ezeani, E., Usman, M., Kwabi, F., & Ye, C. (2023). Do the educational profile, gender, and professional experience of audit committee financial experts improve financial reporting quality? *Journal of International Accounting, Auditing and Taxation*, 53, 100580. <https://doi.org/10.1016/j.intaccaudtax.2023.100580>
- Boubaker, S., Nguyen, P. and Rouatbi, W., 2016. Multiple large shareholders and corporate risk-taking: Evidence from French family firms. *European Financial Management*, 22(4), pp.697-745.
- Beuselinck, C., Blanco, B., Dhole, S., & Lobo, G. J. (2018). *Financial statement readability and tax aggressiveness* (SSRN Scholarly Paper ID 3261115). Social Science Research Network. <https://doi.org/10.2139/ssrn.3261115>
- Bloomfield, R. J. (2002). The “Incomplete Revelation Hypothesis” and financial reporting. *Accounting Horizons*, 16(3), 233–243. <https://doi.org/10.2308/acch.2002.16.3.233>
- Bonner, S. E. (1994). A model of the effects of audit task complexity. *Accounting, Organizations and Society*, 19(3), 213–234.
- Bonsall, S. B., Leone, A. J., Miller, B. P., & Rennekamp, K. (2017). A plain English measure of financial reporting readability. *Journal of Accounting and Economics*, 63(2), 329–357. <https://doi.org/10.1016/j.jacceco.2017.03.002>
- Bozanic, Z., Hoopes, J. L., Thornock, J. R., & Williams, B. M. (2017). IRS Attention. *Journal of Accounting Research*, 55(1), 79–114. <https://doi.org/10.1111/1475-679X.12154>

- Bushee, B. J., Gow, I. D., & Taylor, D. J. (2018). Linguistic complexity in firm disclosures: Obfuscation or information? *Journal of Accounting Research*, 56(1), 85–121. <https://doi.org/10.1111/1475-679X.12179>
- Cai, Y., Jo, H., & Pan, C. (2011). Vice or Virtue? The Impact of Corporate Social Responsibility on Executive Compensation. *Journal of Business Ethics*, 104(2), 159–173. <https://doi.org/10.1007/s10551-011-0909-7>
- Campbell, D. J. (1988). Task complexity: A review and analysis. *The Academy of Management Review*, 13(1), 40. <https://doi.org/10.2307/258353>
- Cao, Y., Dong, Y., Guo, T., & Ma, D. (2022). Short-sale deregulation and corporate tax aggressiveness: Evidence from the Chinese market. *The European Journal of Finance*, 28(4–5), 326–355. <https://doi.org/10.1080/1351847X.2021.1958890>
- Cen, L., Maydew, E. L., Zhang, L., & Zuo, L. (2017). Customer–supplier relationships and corporate tax avoidance. *Journal of Financial Economics*, 123(2), 377–394.
- Chandrasena, S., Matthews, L., & Gerged, A. M. (2025). Does the presence of a sustainability committee strengthen the impact of ESG disclosure on tax aggressiveness? Insights from North America. *Review of Quantitative Finance and Accounting*, 65(3), 1039–1065. <https://doi.org/10.1007/s11156-024-01368-z>
- Chang, H. S., Donohoe, M., & Sougiannis, T. (2016). Do analysts understand the economic and reporting complexities of derivatives? *Journal of Accounting and Economics*, 61(2), 584–604. <https://doi.org/10.1016/j.jacceco.2015.07.005>
- Chen, J. Z., Hong, H. A., Kim, J.-B., & Ryou, J. W. (2021). Information processing costs and corporate tax avoidance: Evidence from the SEC’s XBRL mandate. *Journal of Accounting and Public Policy*, 40(2), 106822. <https://doi.org/10.1016/j.jaccpubpol.2021.106822>
- Cook, K. A., Moser, W. J., & Omer, T. C. (2017). Tax avoidance and ex ante cost of capital. *Journal of Business Finance & Accounting*, 44(7/8), 1109–1136. <https://doi.org/10.1111/jbfa.12258>
- Dao, M., & Xu, H. (2024). Shareholder activism and accounting reporting complexity. *Journal of Accounting Literature*, 46(1), 121–149.
- de Souza, J. A. S., Rissatti, J. C., Rover, S., & Borba, J. A. (2019). The linguistic complexities of narrative accounting disclosure on financial statements: An analysis based on readability characteristics. *Research in International Business and Finance*, 48, 59–74. <https://doi.org/10.1016/j.ribaf.2018.12.008>
- Dhole, S., Lobo, G. J., Mishra, S., & Pal, A. M. (2015). Effects of the SEC’s XBRL mandate on financial reporting comparability. *International Journal of Accounting Information Systems*, 19, 29–44. <https://doi.org/10.1016/j.accinf.2015.11.002>
- Duong, K. T., Elmahgoub, M., Gaia, S., & Malikov, K. T. (2024). CEO social capital and the readability of 10-K reports. *Accounting and Business Research*, 56(1), 99–139. <https://doi.org/10.1080/00014788.2024.2424532>
- Dyer, T., Lang, M., & Stice-Lawrence, L. (2017). The evolution of 10-K textual disclosure: Evidence from Latent Dirichlet Allocation. *Journal of Accounting and Economics*, 64(2), 221–245. <https://doi.org/10.1016/j.jacceco.2017.07.002>
- Dyreng, S. D., Hanlon, M., & Maydew, E. L. (2008). Long-run corporate tax avoidance. *The Accounting Review*, 83(1), 61–82. <https://doi.org/10.2308/accr.2008.83.1.61>
- Dyreng, S. D., Hanlon, M., & Maydew, E. L. (2019). When does tax avoidance result in tax uncertainty? *Accounting Review*, 94(2), 179–203. <https://doi.org/10.2308/accr-52198>
- Dyreng, S. D., Hanlon, M., Maydew, E. L., & Thornock, J. R. (2017). Changes in corporate effective tax rates over the past 25 years. *Journal of Financial Economics*, 124(3), 441–463. <https://doi.org/10.1016/j.jfineco.2017.04.001>

- Dyreng, S. D., & Lindsey, B. P. (2009). Using financial accounting data to examine the effect of foreign operations located in tax havens and other countries on U. S. multinational firms' tax rates. *Journal of Accounting Research*, 47(5), 1283–1316.
- Dyreng, S. D., Lindsey, B. P., & Thornock, J. R. (2013). Exploring the role Delaware plays as a domestic tax haven. *Journal of Financial Economics*, 108(3), 751–772. <https://doi.org/10.1016/j.jfineco.2013.01.004>
- Elmarzouky, M., Moussa, T., & Allam, A. (2025). Cybersecurity disclosure: Board commitment and regulatory impact in the United Kingdom. *The International Journal of Accounting*, 2542005. <https://doi.org/10.1142/S1094406025420053>
- Ezeani, E., Bilal, B., & Samuel, F. (2025). CEO age and capital structure dynamics: The moderating effect of overconfidence and tenure. *International Journal of Finance & Economics*, 30, 3092–3118. <https://doi.org/10.1002/ijfe.3059>
- Ezeani, E., Salem, R. I. A., Usman, M., Kwabi, F., & Bilal. (2023). Board characteristics and corporate cash holding: evidence from the UK, France and Germany. *International Journal of Accounting & Information Management*, 31, 413–439. <https://doi.org/10.1108/IJAIM-09-2022-0184>
- Ezeani, E., Salem, R., Kwabi, F., Boutaine, K., Bilal, & Komal, B. (2022). Board monitoring and capital structure dynamics: evidence from bank-based economies. *Review of Quantitative Finance and Accounting*, 58, 473–498. <https://doi.org/10.1007/s11156-021-01000-4>
- Ezeani, E., Fulgence, S., Hu, W., Kwabi, F., & Wonu, C. (2025). Corporate board reform and capital structure dynamics: evidence from UK. *Review of Quantitative Finance and Accounting*, 65, 973–1003. <https://doi.org/10.1007/s11156-024-01365-2>
- Frank, M. M., Lynch, L. J., & Rego, S. O. (2009). Tax reporting aggressiveness and its relation to aggressive financial reporting. *The Accounting Review*, 84(2), 467–496.
- Graham, J. R., Hanlon, M., Shevlin, T., & Shroff, N. (2014). Incentives for tax planning and avoidance: Evidence from the field. *Accounting Review*, 89(3), 991–1023. <https://doi.org/10.2308/accr-50678>
- Guay, W., Samuels, D., & Taylor, D. (2016). Guiding through the Fog: Financial statement complexity and voluntary disclosure. *Journal of Accounting and Economics*, 62(2), 234–269. <https://doi.org/10.1016/j.jacceco.2016.09.001>
- Gunning, R. (1952). *Technique of clear writing*. McGraw-Hill.
- Gunning, R. (1969). The fog index after twenty years. *Journal of Business Communication*, 6(2), 3–13.
- Hanlon, M., & Heitzman, S. (2010). A review of tax research. *Journal of Accounting and Economics*, 50(2–3), 127–178.
- Hasan, I., Hoi, C. K. (Stan), 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. <https://doi.org/10.1016/j.jfineco.2014.03.004>
- He, G., Ren, H. M., & Taffler, R. (2020). The impact of corporate tax avoidance on analyst coverage and forecasts. *Review of Quantitative Finance & Accounting*, 54(2), 447–477. <https://doi.org/10.1007/s11156-019-00795-7>
- Hoitash, R., & Hoitash, U. (2018). Measuring accounting reporting complexity with XBRL. *The Accounting Review*, 93(1), 259–287. <https://doi.org/10.2308/accr-51762>
- Hoopes, J. L., Mescall, D., & Pittman, J. A. (2012). Do IRS audits deter corporate tax avoidance? *The Accounting Review*, 87(5), 1603–1639. <https://doi.org/10.2308/accr-50187>
- Hope, O.-K., Ma, M. S., & Thomas, W. B. (2013). Tax avoidance and geographic earnings disclosure. *Journal of Accounting and Economics*, 56(2–3), 170–189.
- Hsiao, C. (2022). *Analysis of panel data*. Cambridge university press.

- Huang, F., No, W. G., & Vasarhelyi, M. A. (2019). Do managers use extension elements strategically in the SEC's tagged data for financial statements? Evidence from XBRL complexity. *Journal of Information Systems*, 33(3), 61–74. <https://doi.org/10.2308/isys-52162>
- Inger, K. K., Meckfessel, M. D., Zhou, M. (Jamie), & Fan, W. (Patrick). (2018). An examination of the impact of tax avoidance on the readability of tax footnotes. *Journal of the American Taxation Association*, 40(1), 1–29. <https://doi.org/10.2308/atax-51812>
- Internal Revenue Service (n.d.) *Tax Cuts and Jobs Act: A comparison for businesses*. Available at: <https://www.irs.gov/newsroom/tax-cuts-and-jobs-act-a-comparison-for-businesses> (Accessed: 21 November 2025).
- Jaafar, A., & Thornton, J. (2015). Tax havens and effective tax rates: An analysis of private versus public European firms. *The International Journal of Accounting*, 50(4), 435–457. <https://doi.org/10.1016/j.intacc.2015.10.005>
- Kalelkar, R., Xu, H., Nguyen, D., & Chen, Z. (2024). Generalist CEOs and the readability of the 10-K report. *Advances in Accounting*, 65, 100680. <https://doi.org/10.1016/j.adiac.2023.100680>
- Kanagaretnam, K., Lee, J., Lim, C. Y., & Lobo, G. (2018). Societal trust and corporate tax avoidance. *Review of Accounting Studies*, 23(4), 1588–1628. <https://doi.org/10.1007/s11142-018-9466-y>
- Khan, M., Srinivasan, S., & Liang Tan. (2017). Institutional ownership and corporate tax avoidance: New evidence. *The Accounting Review*, 92(2), 101–122. <https://doi.org/10.2308/accr-51529>
- Kim, J.-B., Kim, J. W., & Lim, J.-H. (2019). Does XBRL adoption constrain earnings management? Early evidence from mandated U.S. filers. *Contemporary Accounting Research*, 36(4), 2610–2634. <https://doi.org/10.1111/1911-3846.12493>
- Koester, A., Shevlin, T., & Wangerin, D. (2017). The role of managerial ability in corporate tax avoidance. *Management Science*, 63(10), 3285–3310. <https://doi.org/10.1287/mnsc.2016.2510>
- Kwabi, F. O., Adegbite, E., Ezeani, E., Wonu, C., & Mumbi, H. (2024). Political uncertainty and stock market liquidity, size, and transaction cost: The role of institutional quality. *International Journal of Finance & Economics*, 29, 2030–2048. <https://doi.org/10.1002/ijfe.2760>
- Larcker, D. F., & Rusticus, T. O. (2010). On the use of instrumental variables in accounting research. *Journal of Accounting and Economics*, 49(3), 186–205. <https://doi.org/10.1016/j.jacceco.2009.11.004>
- Law, K. K. F., & Mills, L. F. (2015). Taxes and financial constraints: Evidence from linguistic cues. *Journal of Accounting Research*, 53(4), 777–819. <https://doi.org/10.1111/1475-679X.12081>
- Lehavy, R., Feng Li, & Merkley, K. (2011). The effect of annual report readability on analyst following and the properties of their earnings forecasts. *The Accounting Review*, 86(3), 1087–1115. <https://doi.org/10.2308/accr.00000043>
- Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics*, 45(2), 221–247. <https://doi.org/10.1016/j.jacceco.2008.02.003>
- Lin, H., He, Y., Wang, M., & Zhang, J. (2025). How does state-controlling shareholders affect corporate tax avoidance? *Journal of International Financial Management & Accounting*, 36(2), 346–367. <https://doi.org/10.1111/jifm.12228>
- Loughran, T., & McDonald, B. (2014). Measuring readability in financial disclosures. *The Journal of Finance*, 69(4), 1643–1671. <https://doi.org/10.1111/jofi.12162>

- Luo, L., Ma, M. S., Omer, T. C., & Xie, H. (2023). Tax avoidance and firm value: Does qualitative disclosure in the tax footnote matter? *Review of Accounting Studies*. <https://doi.org/10.1007/s11142-023-09773-w>
- Ming-Chin, C., Chia-Wen, C., & Mei-Chueh, L. (2020). The effect of chief financial officers' accounting expertise on corporate tax avoidance: The role of compensation design. *Review of Quantitative Finance and Accounting*, 54(1), 273–296. <https://doi.org/10.1007/s11156-019-00789-5>
- Minnick, K., & Noga, T. (2010). Do corporate governance characteristics influence tax management? *Journal of Corporate Finance*, 16(5), 703–718.
- Moussa, T., Krasnikova, N., Postoyeva, M., Hawas, A. and Helfaya, A. (2026). Sustainability Performance, Cost of Debt, and Institutional Environment Quality: Global Evidence From the Hospitality and Tourism Industry. *Business Strategy and The Environment*. 35 (2), pp. 1712-1734. <https://doi.org/10.1002/bse.70230>
- Moussa, T., Allam, A. and Elmarzouky, M. (2023). An examination of UK companies' modern slavery disclosure practices: Does board gender diversity matter? *Business Strategy and The Environment*. 32 (8), pp. 5382-5402. <https://doi.org/10.1002/bse.3426>
- Moussa, T., Allam, A., Elbanna, S. and Bani-Mustafa, A. (2020). Can board environmental orientation improve U.S. firms' carbon performance? The mediating role of carbon strategy. *Business Strategy and The Environment*. 29 (1), pp. 72-86. <https://doi.org/10.1002/bse.2351>
- Moussa, T., Kotb, A. and Helfaya, A. (2022). An Empirical Investigation of U.K. Environmental Targets Disclosure: The Role of Environmental Governance and Performance. *European Accounting Review*. 31 (4), pp. 937-971. <https://doi.org/10.1080/09638180.2021.1890173>
- Nguyen, J. H. (2020). Tax avoidance and financial statement readability. *European Accounting Review*, 1–24. <https://doi.org/10.1080/09638180.2020.1811745>
- Oradi, J., Hesarzadeh, R., E-Vahdati, S., & Nadeem, M. (2024). CEO succession origin and annual reports readability. *The British Accounting Review*, 56(6), 101384. <https://doi.org/10.1016/j.bar.2024.101384>
- Petersen, M. A. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *The Review of Financial Studies*, 22(1), 435–480. <https://doi.org/10.1093/rfs/hhn053>
- Peterson, K. (2012). Accounting complexity, misreporting, and the consequences of misreporting. *Review of Accounting Studies*, 17(1), 72–95. <https://doi.org/10.1007/s11142-011-9164-5>
- Plečnik, J. M., & Wang, S. (2021). Top management team intrapersonal functional diversity and tax avoidance. *Journal of Management Accounting Research*, 33(1), 103–128. <https://doi.org/10.2308/JMAR-19-058>
- Plumlee, M. A. (2003). The effect of information complexity on analysts' use of that information. *Accounting Review*, 78(1), 275. <https://doi.org/10.2308/accr.2003.78.1.275>
- Rahman, D., & Kabir, M. (2024). Does board independence influence annual report readability? *European Accounting Review*, 33(5), 1923–1950. <https://doi.org/10.1080/09638180.2023.2223590>
- Rego, S. O., & Wilson, R. (2012). Equity risk incentives and corporate tax aggressiveness. *Journal of Accounting Research*, 50(3), 775–810. <https://doi.org/10.1111/j.1475-679X.2012.00438.x>
- Salem, R., Ezeani, E., Gerged, A. M., & Bilal, B. (2025). Credit confidence: the impact of disclosure quality and audit quality on bank ratings in emerging economies. *Journal of*

- Accounting in Emerging Economies*, 15, 571-622. <https://doi.org/10.1108/JAEE-04-2024-0181>
- SEC. (2008). *Final Report of the Advisory Committee on Improvements to Financial Reporting*. <https://www.sec.gov/about/offices/oca/acifr/acifr-finalreport.pdf>
- SEC. (2009). *Interactive data to improve financial reporting*. Securities and Exchange Commission. <http://www.sec.gov/rules/final/2009/33-9002.pdf>
- Sianipar, N. K., Yahya, I., & Sadalia, I. (2020). The determinants of tax avoidance with firm size as moderating variable at multinational companies. *International Journal of Research and Review*, 7(7), 237–242.
- Shim, H. (2024). The role of financial affiliates in tax avoidance by business groups: Evidence from Korea. *Journal of International Financial Management & Accounting*, 35(1), 290–326. <https://doi.org/10.1111/jifm.12196>
- Sikka, P. (2017). Accounting and taxation: Conjoined twins or separate siblings? *Accounting Forum*, 41(4), 390–405. <https://doi.org/10.1016/j.accfor.2016.12.003>
- Sikka, P., & Willmott, H. (2010). The dark side of transfer pricing: Its role in tax avoidance and wealth retentiveness. *Critical Perspectives on Accounting*, 21(4), 342–356. <https://doi.org/10.1016/j.cpa.2010.02.004>
- Sun, L., Johnson, G., & Bradley, W. (2022). CEO power and annual report reading difficulty. *Journal of Contemporary Accounting & Economics*, 18(2), 100315. <https://doi.org/10.1016/j.jcae.2022.100315>
- Tang, T., Xu, L., Yan, X., & Yang, H. (2022). Simultaneous debt–equity holdings and corporate tax avoidance. *Journal of Corporate Finance*, 72, 102154. <https://doi.org/10.1016/j.jcorpfin.2021.102154>
- Tantawy, S. and Moussa, T. (2023). The effect of political connections on firms' auditor choice decisions and audit opinions: evidence from Egypt. *Asian Review of Accounting*, 31 (3), pp. 414-436. <https://doi.org/10.1108/ara-07-2022-0161>
- Tax Justice Network. (2023). *The State of Tax Justice 2023—Tax Justice Network*. Tax Justice Network. <https://taxjustice.net/reports/the-state-of-tax-justice-2023/>
- Ullah, S., Zaefarian, G., & Ullah, F. (2021). How to use instrumental variables in addressing endogeneity? A step-by-step procedure for non-specialists. *Industrial Marketing Management*, 96, A1–A6. <https://doi.org/10.1016/j.indmarman.2020.03.006>
- Usman, M., Hussain, M. J., Akuffo, I. N., Ezedebego, I. R., & Ezeani, E. (2026). Independent female directors and green innovation in China: The moderating roles of ownership structure and market competition. *Business Strategy and the Environment*, 1-20. <https://doi.org/10.1002/bse.70572>
- Wang, C., Wilson, R. J., Zhang, S., & Zou, H. (2021). Political costs and corporate tax avoidance: Evidence from sin firms. *Journal of Accounting and Public Policy*, 106861. <https://doi.org/10.1016/j.jaccpubpol.2021.106861>
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.
- Yang, D., Zhou, G., & Zhu, X. (2025). Cross-border cooperation and tax avoidance of U.S. cross-listed firms. *Journal of International Financial Management & Accounting*, 36(1), 144–183. <https://doi.org/10.1111/jifm.12221>
- You, H., & Zhang, X. (2009). Financial reporting complexity and investor underreaction to 10-K information. *Review of Accounting Studies*, 14(4), 559–586. <https://doi.org/10.1007/s11142-008-9083-2>

Tables

Table 1. Sample selection and industry composition

Panel (A) Sample selection		
	firm-year observations	
Total firm-year observations with available ARC of non-financial firms for the period 2011 – 2020	40737	
Less:		
(1) Firm-year observations with insufficient data to compute tax avoidance	(785)	
(2) Firm-year observations with negative pre-tax income	(17509)	
(3) Observations with insufficient data to calculate control variables	(4208)	
Final sample	18235	
Panel (B) Industry composition		
Industry	Firm-year observations	Percent
Agriculture, Forestry and Fishing	306	0.017
Mining	1299	0.071
Construction	287	0.016
Manufacturing	9222	0.506
Transportation, Communications, Electric, Gas and Sanitary service	1050	0.058
Wholesale Trade	994	0.055
Retail Trade	1537	0.084
Services	3429	0.188
Other	111	0.006
Total	18235	1.000

Table 2. Variable definitions

Variables	Definition
Panel A: Accounting Reporting Complexity Variables	
<i>Accounting Reporting Complexity (ARC)</i>	Natural log of the total number of distinct non-repeating and monetary annual tags reported in the financial statements and notes, where each XBRL tag refers to accounting standards and regulations.
<i>TAXONOMY</i>	Natural log of the total number of non-repeating monetary taxonomy facts within a disclosure, which are tags approved by FASB and appear on the XBRL taxonomy at the time of the filing.
<i>EXTENSIONS</i>	Natural log of the total number of non-repeating monetary extended tags. Extensions are tags created by companies when firms determine that the available taxonomy tags cannot accurately capture their company-specific economic transactions.
<i>ALL-FACTS</i>	Natural log of the total number of annual monetary tags reported in the financial statements and notes, including repeating facts. This complexity measure is a combination of the number of tags used and their frequency of use.
Panel B: Corporate Tax Avoidance Variables	
<i>TA_ETR</i>	Effective tax rate (ETR) is defined as the total tax expense (txt) divided by pre-tax book income (pi) before special items (spi). ETR with negative denominators are removed, and then the remaining are winsorized to the range [0,1]. TA_ETR is defined as (−1) times ETR
<i>TA_CETR</i>	Cash effective tax rate (CETR) is defined as cash tax paid (txpd) divided by pre-tax book income (pi) before special items (spi). CETR with negative denominators are removed, and then the remaining are winsorized to the range [0,1]. TA_CETR is defined as (−1) times CETR
<i>TAX_HAVEN</i>	Natural log of the number of tax havens.
<i>TAX_AVOID3</i>	The difference between the tax on pre-tax book income before special items at the statutory tax rate and the current tax paid is expressed as a percentage of pre-tax income over a three-year window.
<i>DTAX</i>	The residuals (ε) from regressions (estimated by industry and year) of permanent differences on non-discretionary items such as intangible assets and state taxes.
Panel C: Control Variables	
<i>Firm size (FSIZE)</i>	Natural log of total assets of the firm.
<i>Liquidity (LIQU)</i>	Percentage of current assets to current liabilities.
<i>Profitability (ROE)</i>	Percentage of net income to total equity
<i>Growth opportunities (MTB)</i>	Percentage of the market to book value of equity.

<i>Leverage (LEV)</i>	Percentage of total debt to total assets
<i>Earnings volatility (EVOL)</i>	Standard deviation of the operating earnings in the last five years.
<i>Merger and acquisition (MA)</i>	Dummy equals one if a firm had a merger or acquisition during the year and zero otherwise.
<i>Business segments (B_SEG)</i>	Natural log of the number of business segments.
<i>Foreign segments (F_SEG)</i>	Natural log of the number of foreign segments.
<i>Employees (EMP)</i>	Natural log of the number of employees of the firm.
<i>Beta (BETA)</i>	A measurement of a firm's stock price sensitivity (or volatility) to the overall fluctuation of a given benchmark index (e.g., S&P 500 for U.S. stocks). It is calculated using a least squares regression analysis based on 60 monthly returns (or a minimum of 24).
<i>CEO tenure (CEO_TENURE)</i>	The number of years the Chief Executive Officer (CEO) has served in their current role.
<i>CEO professional qualification (CEO_PQF)</i>	A binary dummy variable equal to 1 if the CEO possesses a professional qualification (e.g., CPA, CFA, ACCA, or ICAEW), and 0 otherwise.
<i>Governance (GOV)</i>	A factor score derived from Principal Component Analysis (PCA) capturing corporate governance quality. It is constructed from metrics of audit committee independence, compensation committee independence, and board independence.

Table 3. Descriptive statistics

Panel A: Descriptive statistics for all variables in the full sample								
Variables	Mean	Std. Dev.	P25	Median	P75			
ARC	230.680	151.020	104.000	200.000	332.000			
TAXONOMY	194.820	127.930	87.000	166.000	284.000			
EXTENSIONS	35.860	39.040	12.000	27.000	48.000			
ALL-FACTS	678.828	533.417	258.000	501.000	973.000			
TA_ETR	-0.193	0.169	-0.340	-0.200	0.000			
TA_CETR	-0.189	0.168	-0.300	-0.170	-0.020			
FSIZE	5.444	3.193	3.600	5.850	7.630			
LIQU	2.728	2.584	1.110	1.880	3.310			
ROE	-0.055	0.229	-0.060	0.030	0.060			
MTB	2.782	4.437	1.050	2.130	4.000			
LEV	0.456	0.907	0.000	0.180	0.710			
EVOL	95.121	177.729	4.930	20.020	79.540			
MA	0.392	0.488	0.000	0.000	1.000			
B_SEG	2.140	0.880	1.610	2.200	2.830			
F_SEG	1.730	0.860	1.100	1.790	2.300			
EMP	8.466	16.288	0.080	1.100	7.350			
Panel B: Descriptive statistics for ARC proxies for the bottom and top quintiles of tax avoidance								
Variables of interest	High tax-avoiding firms			Low tax-avoiding firms			High vs low tax-avoiding firms	
	Mean	Median	SD	Mean	Median	SD	t-test	z-test
ARC	298.88	290.00	135.49	156.60	126.00	114.46	30.35***	21.23***
TAXONOMY	258.87	251.00	112.24	127.05	101.00	90.96	32.4***	19.92***
EXTENSIONS	40.02	32.00	34.00	29.55	19.00	38.94	-8.60***	23.99***
ALL-FACTS	920.97	793.00	576.22	448.89	303.00	520.05	22.90***	14.64***

Notes: This table reports univariate comparisons of mean and median values of ARC measures (number of tags). The last two columns report the t-statistics for differences in means and z-statistics of the Wilcoxon test for differences in medians between the bottom and top GAAP ETR quintiles. ***, **, and * indicate significance for a two-tailed test at the 1%, 5% and 10% levels, respectively.

Table 4. Correlation matrix

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
(1) ARC	1.000															
(2) TAXONOMY	0.978	1.000														
(3) EXTENSIONS	0.771	0.652	1.000													
(4) ALL-FACTS	0.934	0.927	0.706	1.000												
(5) TA_ETR	0.329	0.367	0.090	0.329	1.000											
(6) TA_CETR	0.237	0.269	0.059	0.241	0.587	1.000										
(7) FSIZE	0.668	0.696	0.381	0.684	-0.480	-0.359	1.000									
(8) LIQU	-0.185	-0.175	-0.158	-0.186	0.089	-0.026	-0.056	1.000								
(9) ROE	0.256	0.295	0.033	0.286	-0.428	-0.300	0.454	0.066	1.000							
(10) MTB	0.095	0.103	0.037	0.099	-0.043	-0.059	0.174	0.165	0.165	1.000						
(11) LEV	0.372	0.375	0.248	0.370	-0.170	-0.092	0.363	-0.150	0.172	0.396	1.000					
(12) EVOL	0.383	0.373	0.327	0.444	-0.141	-0.078	0.402	-0.173	0.152	0.077	0.163	1.000				
(13) MA	0.380	0.399	0.207	0.380	-0.272	-0.190	0.408	-0.140	0.240	0.060	0.175	0.180	1.000			
(14) B_SEG	0.458	0.477	0.257	0.500	-0.330	-0.249	0.505	-0.078	0.302	0.048	0.156	0.312	0.265	1.000		
(15) F_SEG	0.343	0.360	0.186	0.374	-0.202	-0.161	0.352	-0.028	0.207	0.042	0.088	0.214	0.165	0.455	1.000	
(16) EMP	0.400	0.411	0.259	0.461	-0.225	-0.188	0.486	-0.220	0.213	0.069	0.194	0.434	0.285	0.351	0.239	1.000

Notes: All variable definitions are presented in Table 2. Significant correlation coefficients at 1% level are presented in bold.

Table 5. Tax avoidance and accounting reporting complexity

Model	(1)	(2)	(3)	(4)	(5)	(6)
(Variables)	ARC	EXTENSIONS	TAXONOMY	ARC	EXTENSIONS	TAXONOMY
TA_ETR	1.056** (0.355)	1.696*** (0.285)	0.756** (0.312)			
TA_CETR				1.072** (0.463)	0.989*** (0.258)	0.028 (0.153)
FSIZE	2.455*** (0.364)	3.059*** (0.104)	2.167*** (0.295)	2.577*** (0.412)	2.946*** (0.117)	2.304*** (0.331)
LIQU	-0.888*** (0.224)	-1.369*** (0.069)	-0.730*** (0.179)	-0.914*** (0.295)	-1.562*** (0.088)	-0.739*** (0.236)
ROE	-2.911*** (0.879)	-2.219** (0.951)	-1.559*** (0.455)	-4.947*** (0.205)	-1.732*** (0.276)	-3.197*** (0.357)
MTB	-0.472*** (0.149)	-0.460*** (0.048)	-0.217* (0.118)	-0.301*** (0.105)	-0.645*** (0.055)	-0.230** (0.107)
LEV	1.563*** (0.517)	3.170*** (0.239)	1.184* (0.663)	1.710** (0.819)	4.003*** (0.259)	1.256*** (0.445)
EVOL	0.030*** (0.005)	0.019*** (0.001)	0.005 (0.004)	0.017*** (0.006)	0.020*** (0.002)	-0.005 (0.005)
MA	2.592*** (0.429)	3.190*** (0.419)	2.327*** (0.161)	2.371*** (0.502)	2.929*** (0.417)	2.103*** (0.217)
B_SEG	2.580*** (0.151)	0.359*** (0.044)	2.260*** (0.122)	2.440*** (0.158)	0.410*** (0.045)	2.070*** (0.128)
F_SEG	-0.848*** (0.196)	-0.159*** (0.056)	-0.730*** (0.160)	-0.762*** (0.202)	-0.150*** (0.056)	-0.642*** (0.164)
EMP	-0.101* (0.053)	0.004 (0.015)	-0.083* (0.043)	-0.149*** (0.055)	0.002 (0.016)	-0.136*** (0.044)
Constant	-14.520*** (1.890)	-15.39*** (1.946)	-14.270*** (1.752)	-14.440*** (1.360)	-18.920*** (1.049)	-14.260*** (1.156)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	18,235	18,235	18,235	15,511	15,511	15,511
Adjusted R ²	0.649	0.297	0.687	0.631	0.311	0.667

Notes: All variable definitions are presented in Table 2. Models (1)–(3) estimate the association between tax avoidance and ARC measures using TA_ETR, while Models (4)–(6) use TA_CETR. ***, **, and * indicate significance for a two-tailed test at the 1%, 5% and 10% levels, respectively. Robust standard errors are reported in parentheses.

Table 6. Tax avoidance and accounting reporting complexity: High- and low-tax-avoidance subsamples

Sample	Model (1) High TA	Model (2) Low TA	Model (3) High TA	Model (4) Low TA
Panel A: Dependent Variable/ ARC				
TA_ETR	3.275*** (1.252)	-1.377*** (0.473)		
TA_CETR			1.246 (2.745)	-0.762*** (0.272)
FSIZE	2.419*** (0.360)	3.488** (1.414)	2.469*** (0.432)	3.600*** (1.168)
LIQU	-0.887*** (0.278)	-0.723*** (0.277)	-0.906** (0.386)	-0.742*** (0.167)
ROE	-2.342*** (0.326)	-2.243* (1.303)	-4.487*** (0.428)	-2.350 (1.492)
MTB	-0.270*** (0.046)	-0.180*** (0.059)	-0.261*** (0.023)	-0.353*** (0.044)
LEV	1.578*** (0.082)	1.072*** (0.232)	1.562*** (0.099)	1.891*** (0.180)
EVOL	0.028*** (0.005)	-0.047*** (0.016)	0.016** (0.007)	-0.015 (0.011)
MA	2.745*** (0.143)	1.427*** (0.389)	2.570*** (0.169)	1.625*** (0.275)
B_SEG	2.698*** (0.149)	2.152*** (0.406)	2.609*** (0.177)	1.949*** (0.282)
F_SEG	-0.993*** (0.190)	-0.512 (0.505)	-0.956*** (0.222)	-0.299 (0.356)
EMP	-0.107** (0.050)	-0.236 (0.170)	-0.046 (0.059)	-0.571*** (0.102)
Constant	-13.900*** (1.009)	-28.430*** (0.528)	-13.840*** (1.151)	-27.510*** (0.277)
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Observations	16,318	1,917	11,782	3,729
Adjusted R ²	0.655	0.636	0.639	0.621
Panel B: Dependent Variable/ EXTENSIONS				
TA_ETR	2.971*** (1.060)	-3.582*** (1.202)		
TA_CETR			2.491*** (0.951)	-2.603*** (0.813)
FSIZE	3.086*** (0.105)	5.616*** (0.398)	2.838*** (0.124)	5.671*** (0.326)
LIQU	-1.377*** (0.081)	-1.150*** (0.275)	-1.504*** (0.111)	-1.249*** (0.214)

(Continued).

Table 6. Continued.

Sample	Model (1) High TA	Model (2) Low TA	Model (3) High TA	Model (4) Low TA
ROE	-1.219*** (0.395)	-0.413 (0.367)	-1.544*** (0.423)	-0.063 (0.416)
MTB	-0.460*** (0.048)	-0.464*** (0.168)	-0.544*** (0.061)	-0.577*** (0.123)
LEV	3.170*** (0.239)	2.055*** (0.651)	3.567*** (0.285)	3.861*** (0.502)
EVOL	0.019*** (0.002)	0.006 (0.004)	0.019*** (0.002)	0.011*** (0.003)
MA	3.190*** (0.419)	1.881* (1.095)	3.460*** (0.488)	1.641** (0.766)
B_SEG	0.359*** (0.044)	0.370*** (0.114)	0.451*** (0.051)	0.364*** (0.079)
F_SEG	-0.159*** (0.056)	-0.132 (0.142)	-0.199*** (0.064)	-0.080 (0.109)
EMP	0.004 (0.016)	-0.096** (0.048)	0.047*** (0.017)	-0.136*** (0.028)
Constant	-15.390*** (2.946)	-64.160*** (1.485)	-16.940*** (3.308)	-58.420*** (7.734)
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Observations	16,318	1,917	11,782	3,729
Adjusted R ²	0.299	0.370	0.309	0.361
Panel C: Dependent Variable/ TAXONOMY				
TA_ETR	-1.039*** (0.349)	0.208 (1.567)		
TA_CETR			-0.489*** (0.182)	0.182 (2.017)
FSIZE	2.126*** (0.290)	2.968*** (1.055)	2.206*** (0.346)	3.056*** (0.953)
LIQU	-0.728*** (0.223)	-0.603*** (0.179)	-0.735** (0.310)	-0.599*** (0.163)
ROE	-1.184*** (0.262)	-2.080* (1.064)	-2.860*** (0.344)	-2.877** (1.217)
MTB	-0.318** (0.132)	-0.127*** (0.049)	-0.401** (0.171)	-0.281*** (0.036)
LEV	1.199*** (0.066)	0.854*** (0.189)	1.148*** (0.079)	1.450*** (0.147)
EVOL	0.018*** (0.004)	-0.049*** (0.013)	-0.004 (0.005)	-0.027*** (0.009)
MA	2.457*** (0.115)	1.229*** (0.318)	2.265*** (0.136)	1.448*** (0.224)

(Continued).

Table 6. Continued.

	Model (1)	Model (2)	Model (3)	Model (4)
Sample	High TA	Low TA	High TA	Low TA
B_SEG	2.346*** (0.120)	1.827*** (0.332)	2.212*** (0.142)	1.626*** (0.230)
F_SEG	-0.837*** (0.153)	-0.349 (0.413)	-0.799*** (0.178)	-0.227 (0.290)
EMP	-0.089** (0.040)	-0.211 (0.139)	-0.073 (0.047)	-0.441*** (0.083)
Constant	-13.950*** (5.107)	-22.430*** (4.314)	-14.110*** (5.231)	-21.520*** (2.264)
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Observations	16,318	1,917	11,782	3,729
Adjusted R ²	0.694	0.655	0.679	0.638

*Notes: We use three proxies for accounting reporting complexity: the ARC Index (Panel A), the ARC Extensions (Panel B), and the ARC Taxonomy (Panel C). We estimate separate regressions for subsamples of high and low tax avoidance using TA_ETR (Models 1–2) and TA_CETR (Models 3–4). All variable definitions are presented in Table 2. ***, **, and * indicate significance for a two-tailed test at the 1%, 5%, and 10% levels, respectively. Robust standard errors are reported in parentheses.*

Table 7. Robustness check: alternative measures of accounting reporting complexity and tax avoidance

Model (Variables)	(1) ALL-FACTS	(2) ALL-FACTS	(3) ARC	(4) ARC	(5) ARC
TA_ETR	6.651*** (1.770)				
TA_CETR		3.094* (1.782)			
TAX_HAVEN			4.251*** (0.433)		
TAX_AVOID3				3.535** (1.442)	
DTAX					0.076*** (0.029)
FSIZE	9.071*** (1.556)	9.647*** (1.786)	2.393*** (0.662)	2.432*** (0.335)	2.278*** (0.976)
LIQU	-0.313*** (0.087)	-0.335*** (0.116)	-0.762*** (0.051)	-0.887*** (0.067)	-0.727*** (0.076)
ROE	-7.840*** (0.125)	-1.710*** (0.177)	-3.751*** (0.633)	-3.045*** (0.310)	-4.039*** (0.876)
MTB	-0.883*** (0.161)	-1.015*** (0.107)	-1.492*** (0.131)	-2.696*** (0.160)	-2.331*** (0.148)
LEV	5.866*** (0.352)	6.579*** (0.397)	1.216*** (0.150)	1.564*** (0.077)	1.391*** (0.226)
EVOL	0.222*** (0.023)	0.195*** (0.024)	0.078*** (0.009)	0.026*** (0.005)	0.046*** (0.016)
MA	7.710*** (0.585)	6.945*** (0.617)	2.013*** (0.239)	2.601*** (0.135)	3.190*** (0.406)
B_SEG	1.352** (0.615)	1.304** (0.646)	1.911*** (0.231)	2.593*** (0.140)	3.143*** (0.451)
F_SEG	-0.605*** (0.083)	-0.575*** (0.085)	-0.649** (0.029)	-0.868*** (0.018)	-1.029* (0.058)
EMP	1.368*** (0.226)	1.010*** (0.231)	-0.154* (0.089)	-0.085* (0.048)	0.108 (0.149)
Constant	-7.384** (3.210)	-7.515* (4.014)	-4.550** (1.583)	-3.488*** (1.005)	-2.719** (1.187)
Year FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Observations	18,235	15,511	7,165	18,222	1,879
Adjusted R ²	0.643	0.626	0.606	0.650	0.669

Notes: All variable definitions are presented in Table 2. Models (1)–(2) use ALL-FACTS as an alternative proxy for accounting reporting complexity. Models (3)–(5) use the baseline ARC measure. Each model employs a different tax avoidance proxy: TA_ETR (Model 1), TA_CETR (Model 2), TAX_HAVEN (Model 3), TAX_AVOID3 (Model 4), and DTAX (Model 5). All regressions include year and industry fixed effects. ***, **, and * indicate significance for a two-tailed test at the 1%, 5%, and 10% levels, respectively. Robust standard errors are reported in parentheses.

Table 8. Robustness check: additional controls for CEO characteristics, governance, and market volatility

Model (Variables)	(1)	(2)	(3)	(4)	(5)	(6)
	ARC	EXTENSIO NS	TAXONO MY	ARC	EXTENSIO NS	TAXONO MY
TA_ETR	0.153*** (0.035)	0.546*** (0.070)	0.109*** (0.033)			
TA_CETR				0.070** (0.034)	0.310*** (0.067)	0.044 (0.032)
CEO_TENURE	-0.002 (0.001)	-0.005*** (0.001)	-0.003 (0.001)	-0.001 (0.001)	-0.004*** (0.002)	-0.001 (0.001)
CEO_PQF	-0.041*** (0.013)	-0.042 (0.025)	-0.044*** (0.012)	-0.042*** (0.014)	-0.033 (0.027)	-0.048*** (0.013)
GOV	0.010** (0.004)	0.001 (0.008)	0.011*** (0.004)	0.012*** (0.005)	-0.003 (0.009)	0.014*** (0.005)
BETA	0.004 (0.006)	0.002 (0.011)	0.004 (0.005)	0.001 (0.006)	-0.009 (0.013)	0.003 (0.006)
FSIZE	0.122*** (0.005)	0.228*** (0.011)	0.113*** (0.005)	0.124*** (0.006)	0.223*** (0.011)	0.116*** (0.006)
LIQU	-0.021*** (0.003)	-0.045*** (0.005)	-0.018*** (0.002)	-0.020*** (0.003)	-0.055*** (0.006)	-0.016*** (0.003)
ROE	-0.098** (0.039)	-0.609*** (0.078)	-0.052 (0.037)	-0.214*** (0.051)	-0.702*** (0.101)	-0.181*** (0.049)
MTB	-0.010*** (0.001)	-0.009*** (0.003)	-0.010*** (0.001)	-0.010*** (0.002)	-0.011*** (0.003)	-0.009*** (0.001)
LEV	0.081*** (0.006)	0.114*** (0.013)	0.076*** (0.006)	0.083*** (0.007)	0.124*** (0.014)	0.077*** (0.007)
EVOL	-0.004*** (0.000)	-0.002*** (0.000)	-0.001*** (0.000)	-0.003*** (0.000)	-0.003*** (0.000)	-0.002*** (0.000)
MA	0.071*** (0.010)	0.071*** (0.020)	0.072*** (0.009)	0.067*** (0.010)	0.074*** (0.020)	0.067*** (0.010)
B_SEG	0.001 (0.001)	0.001 (0.002)	0.001 (0.001)	0.001 (0.001)	0.003 (0.002)	0.001 (0.001)
F_SEG	0.004*** (0.001)	0.002 (0.002)	0.005*** (0.001)	0.005*** (0.001)	0.002 (0.003)	0.005*** (0.001)
EMP	-0.014*** (0.004)	-0.072*** (0.008)	-0.007* (0.004)	-0.016*** (0.005)	-0.062*** (0.009)	-0.010** (0.004)
Constant	4.843*** (0.035)	2.072*** (0.071)	4.751*** (0.034)	4.818*** (0.038)	2.023*** (0.076)	4.726*** (0.037)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8,170	8,170	8,170	7,487	7,487	7,487
Adjusted R ²	0.300	0.201	0.303	0.267	0.209	0.261

Notes: All variable definitions are presented in Table 2. ***, **, and * indicate significance for a two-tailed test at the 1%, 5% and 10% levels, respectively. Robust standard errors are reported in parentheses.

Table 9. Robustness Check: lagged ARC- tax avoidance structure

Model	(1)	(2)	(3)	(4)	(5)	(6)
(Variables)	ARC	EXTENSIO NS	TAXONO MY	ARC	EXTENSIO NS	TAXONO MY
TA_ETR	0.073*** (0.021)	0.265*** (0.054)	0.037* (0.019)			
TA_CETR				0.063*** (0.020)	0.181*** (0.051)	0.049*** (0.019)
L.ARC	0.265*** (0.006)			0.259*** (0.006)		
L.EXTENSIONS		0.419*** (0.008)			0.411*** (0.008)	
L.TAXONOMY			0.291*** (0.006)			0.284*** (0.006)
CEO_TENURE	-0.002*** (0.000)	-0.004*** (0.001)	-0.001*** (0.000)	-0.002*** (0.000)	-0.004*** (0.001)	-0.002*** (0.000)
CEO_PQF	-0.039*** (0.007)	-0.064*** (0.018)	-0.039*** (0.007)	-0.035*** (0.008)	-0.048** (0.019)	-0.038*** (0.007)
GOV	0.004* (0.002)	-0.005 (0.006)	0.005** (0.002)	0.003 (0.003)	-0.012* (0.007)	0.004* (0.002)
BETA	0.001 (0.003)	-0.004 (0.008)	0.002 (0.003)	0.000 (0.004)	-0.007 (0.010)	0.002 (0.003)
FSIZE	0.066*** (0.003)	0.078*** (0.008)	0.058*** (0.003)	0.066*** (0.003)	0.069*** (0.009)	0.060*** (0.003)
LIQU	-0.016*** (0.001)	-0.023*** (0.004)	-0.013*** (0.001)	-0.017*** (0.002)	-0.033*** (0.004)	-0.014*** (0.002)
ROE	0.058** (0.024)	-0.154** (0.061)	0.093*** (0.022)	-0.015 (0.033)	-0.159* (0.085)	0.009 (0.031)
MTB	-0.012*** (0.001)	-0.016*** (0.002)	-0.011*** (0.001)	-0.013*** (0.001)	-0.019*** (0.002)	-0.011*** (0.001)
LEV	0.064*** (0.004)	0.083*** (0.009)	0.058*** (0.003)	0.067*** (0.004)	0.094*** (0.010)	0.059*** (0.004)
EVOL	-0.000*** (0.000)	0.000 (0.000)	-0.000*** (0.000)	-0.000*** (0.000)	0.000 (0.000)	-0.000*** (0.000)
MA	0.073*** (0.006)	0.073*** (0.014)	0.074*** (0.005)	0.073*** (0.006)	0.082*** (0.015)	0.072*** (0.005)
B_SEG	0.006*** (0.001)	0.009*** (0.001)	0.006*** (0.000)	0.006*** (0.001)	0.010*** (0.001)	0.006*** (0.001)
F_SEG	-0.003*** (0.001)	-0.006*** (0.002)	-0.002*** (0.001)	-0.002*** (0.001)	-0.007*** (0.002)	-0.002*** (0.001)
EMP	-0.008*** (0.002)	-0.032*** (0.006)	-0.004* (0.002)	-0.010*** (0.003)	-0.026*** (0.007)	-0.006*** (0.002)
Constant	3.782*** (0.035)	1.648*** (0.055)	3.575*** (0.034)	3.824*** (0.037)	1.697*** (0.059)	3.622*** (0.035)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,012	6,012	6,012	5,508	5,508	5,508
Adjusted R ²	0.671	0.504	0.692	0.635	0.510	0.647

Notes: All variable definitions are presented in Table 2. ***, **, and * indicate significance for a two-tailed test at the 1%, 5% and 10% levels, respectively. Robust standard errors are reported in parentheses.

Table 10. IV 2SLS regression

Panel A. First-stage regressions		
Model (Variables)	(1) TA_ETR	(2) TA_CETR
IND_AVG_TA_ETR	0.419*** (0.042)	
IND_AVG_TA_CETR		0.552*** (0.060)
CEO_TENURE	-0.001*** (0.000)	-0.000 (0.000)
CEO_PQF	-0.006 (0.004)	0.004 (0.004)
GOV	0.002 (0.001)	0.006*** (0.002)
BETA	0.003* (0.002)	0.004* (0.002)
FSIZE	0.002 (0.002)	0.015*** (0.002)
LIQU	0.006*** (0.001)	0.000 (0.001)
ROE	-0.271*** (0.014)	-0.162*** (0.019)
MTB	0.003*** (0.000)	0.001 (0.001)
LEV	0.002 (0.002)	0.007*** (0.003)
EVOL	0.000*** (0.000)	0.000 (0.000)
MA	0.003 (0.003)	-0.005 (0.003)

B_SEG	-0.001**	-0.002***
	(0.000)	(0.000)
F_SEG	0.001**	0.001***
	(0.000)	(0.000)
EMP	-0.008***	-0.019***
	(0.002)	(0.002)
Constant	-0.176***	-0.160***
	(0.021)	(0.025)
Year FE	Yes	Yes
Industry FE	Yes	Yes
Observations	8,170	7,487

Panel B. Second-stage regressions

Model (Variables)	Instrumented TA_ETR			Instrumented TA_CETR		
	(1)	(2)	(3)	(4)	(5)	(6)
	ARC	EXTENSIONS	TAXONOMY	ARC	EXTENSIONS	TAXONOMY
TA_ETR	3.138*** (0.380)	1.024* (0.573)	3.313*** (0.389)			
TA_CETR				1.716*** (0.296)	2.009*** (0.600)	1.630*** (0.280)
CEO_TENURE	0.001 (0.001)	-0.006*** (0.001)	0.002* (0.001)	-0.002*** (0.001)	-0.005*** (0.001)	-0.002*** (0.001)
CEO_PQF	-0.024 (0.016)	-0.041* (0.024)	-0.025 (0.016)	-0.051*** (0.013)	-0.042 (0.026)	-0.056*** (0.012)
GOV	0.000 (0.005)	-0.004 (0.008)	0.000 (0.005)	-0.002 (0.005)	-0.018* (0.009)	-0.000 (0.004)
BETA	-0.016** (0.007)	-0.002 (0.011)	-0.017** (0.007)	-0.015** (0.006)	-0.018 (0.013)	-0.014** (0.006)
FSIZE	0.108*** (0.007)	0.232*** (0.010)	0.096*** (0.007)	0.094*** (0.008)	0.198*** (0.016)	0.085*** (0.007)
LIQU	-0.041*** (0.004)	-0.048*** (0.006)	-0.040*** (0.004)	-0.019*** (0.003)	-0.054*** (0.006)	-0.015*** (0.003)
ROE	0.803*** (0.115)	-0.428** (0.173)	0.917*** (0.118)	0.160** (0.069)	-0.352** (0.140)	0.191*** (0.065)

MTB	-0.017*** (0.002)	-0.009*** (0.003)	-0.018*** (0.002)	-0.010*** (0.001)	-0.011*** (0.003)	-0.009*** (0.001)
LEV	0.051*** (0.008)	0.086*** (0.012)	0.045*** (0.008)	0.044*** (0.007)	0.079*** (0.014)	0.037*** (0.006)
EVOL	-0.003*** (0.000)	-0.001 (0.000)	-0.003*** (0.000)	-0.004*** (0.000)	-0.001 (0.000)	-0.002*** (0.000)
MA	0.063*** (0.012)	0.074*** (0.018)	0.063*** (0.012)	0.080*** (0.010)	0.088*** (0.019)	0.080*** (0.009)
B_SEG	0.010*** (0.001)	0.009*** (0.002)	0.011*** (0.001)	0.012*** (0.001)	0.014*** (0.002)	0.012*** (0.001)
F_SEG	-0.007*** (0.002)	-0.006*** (0.002)	-0.007*** (0.002)	-0.006*** (0.001)	-0.009*** (0.003)	-0.006*** (0.001)
EMP	0.024*** (0.007)	-0.064*** (0.010)	0.034*** (0.007)	0.025*** (0.008)	-0.020 (0.017)	0.029*** (0.008)
Constant	4.496*** (0.134)	0.483** (0.202)	4.524*** (0.137)	4.030*** (0.105)	0.769*** (0.212)	3.973*** (0.099)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8,170	8,170	8,170	7,487	7,487	7,487
Wald chi ²	4896	4168	4484	7117	3730	7475
F test	178.25	144.98	163.67	267.67	133.03	277.90
Prob > F	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***
Underidentification test						
Kleibergen–Paap rk LM statistic	100.026	100.026	100.026	78.421	78.421	78.421
Weak identification tests						
Cragg–Donald Wald <i>F</i> -statistic	106.485	106.485	106.485	86.074	86.074	86.074
Kleibergen–Paap rk Wald <i>F</i> -statistic	101.061	101.061	101.061	84.057	84.057	84.057
Stock–Yogo weak ID test critical values						
10% maximal IV size	16.38	16.38	16.38	16.38	16.38	16.38
15% maximal IV size	8.96	8.96	8.96	8.96	8.96	8.96
20% maximal IV size	6.66	6.66	6.66	6.66	6.66	6.66
25% maximal IV size	5.53	5.53	5.53	5.53	5.53	5.53

Notes: All variable definitions are presented in Table 2. ***, **, and * indicate significance for a two-tailed test at the 1%, 5% and 10% levels, respectively. Robust standard errors are reported in parentheses.

Table 11. Propensity Score Matched Difference-in-Differences (PSM-DID) model

Panel A: PSM- covariate differences between control and treatment group						
	Control	Treatment	Control Mean	Treatment Mean	t value	p value
CEO TENURE	378	378	6.351	5.861	1.050	0.293
CEO_PQF	378	378	0.167	0.167	0.000	1.000
GOV	378	378	0.081	0.003	1.050	0.289
BETA	378	378	0.935	0.985	-1.050	0.286
FSIZE	378	378	7.348	7.190	1.150	0.252
LIQU	378	378	2.636	2.793	-1.050	0.286
ROE	378	378	0.032	0.018	1.900	0.056
MTB	378	378	3.588	3.413	0.600	0.562
LEV	378	378	0.651	0.530	1.800	0.071
EVOL	378	378	95.934	115.746	-1.550	0.117
MA	378	378	0.582	0.476	2.950	0.004
B_SEG	378	378	12.944	14.95	-3.15	0.002
F_SEG	378	378	3.728	6.455	-6.05	0.000
EMP	378	378	1.455	1.484	-0.200	0.850
Panel B: DiD Effect on accounting reporting complexity						
Model			(1)	(2)	(3)	
Variables			ARC	EXTENSIONS	TAXONOMY	
POST			0.001 (0.010)	0.048 (0.036)	-0.001 (0.009)	
TREATMENT			0.010 (0.021)	0.067 (0.063)	0.003 (0.020)	
POST*TREATMENT			-0.035* (0.019)	-0.030 (0.057)	-0.036** (0.018)	
CONTROLS			Included	Included	Included	
Constant			5.480*** (0.079)	3.154*** (0.254)	5.372*** (0.074)	
Year FE			Yes	Yes	Yes	
Industry FE			Yes	Yes	Yes	
Observations			756	756	756	
Adjusted R ²			0.583	0.247	0.593	

Notes: All variable definitions are presented in Table 2. ***, **, and * indicate significance for a two-tailed test at the 1%, 5% and 10% levels, respectively. Robust standard errors are reported in parentheses.

Table 12. Channel analysis: IRS audit probability

Model	(1)	(2)	(3)	(4)
(Variables)	High IRS audit	Low IRS audit	High IRS audit	Low IRS audit
TA_ETR	0.637*** (0.211)	0.407** (0.185)		
TA_CETR			0.355** (0.162)	0.221* (0.129)
FSIZE	2.756*** (0.393)	2.320*** (0.407)	2.717*** (0.398)	2.408*** (0.474)
LIQU	-0.553** (0.264)	-0.718** (0.303)	-0.662** (0.267)	-1.705*** (0.396)
ROE	-0.066 (3.318)	-3.939*** (0.360)	0.637 (3.347)	-6.472*** (0.475)
MTB	-0.402*** (0.082)	-0.263*** (0.019)	-0.415*** (0.084)	-0.299*** (0.023)
LEV	3.182*** (0.352)	1.799*** (0.093)	3.293*** (0.355)	2.057*** (0.107)
EVOL	-0.001 (0.014)	0.066*** (0.008)	0.002 (0.015)	0.052*** (0.009)
MA	2.194*** (0.559)	2.797*** (0.163)	2.307*** (0.564)	2.634*** (0.178)
B_SEG	2.054*** (0.439)	1.383*** (0.167)	1.983*** (0.444)	1.199*** (0.182)
F_SEG	-0.458 (0.589)	0.616*** (0.222)	-0.348 (0.593)	0.771*** (0.240)
EMP	-0.677*** (0.128)	0.214*** (0.069)	-0.698*** (0.130)	0.152** (0.073)
Constant	10.080*** (3.523)	13.830*** (2.288)	9.523*** (3.516)	13.180*** (2.727)
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Observations	2,300	15,918	2,247	13,248
Adjusted R ²	0.123	0.489	0.124	0.454

Notes: All variable definitions are provided in Table 2. The dependent variable is ARC in all models. Models (1)–(2) use TA_ETR as the tax avoidance proxy, while Models (3)–(4) use TA_CETR. High (low) IRS audit probability denotes firms above (below) the sample median of the IRS audit probability measure. ***, **, and * indicate significance for a two-tailed test at the 1%, 5% and 10% levels, respectively. Robust standard errors are reported in parentheses.

Table 13. Channel analysis: firm financial capacity

Panel A: Low financial capacity (lower pre-tax income)						
Model	(1)	(2)	(3)	(4)	(5)	(6)
(Variables)	ARC	EXTENSIO NS	TAXONO MY	ARC	EXTENSIO NS	TAXONO MY
TA_ETR	0.153*** (0.049)	0.443*** (0.103)	0.127*** (0.046)			
TA_CETR				0.063 (0.042)	0.276*** (0.089)	0.039 (0.040)
CEO_TENURE	-0.000 (0.001)	-0.003 (0.002)	0.000 (0.001)	-0.000 (0.001)	-0.003 (0.002)	-0.000 (0.001)
CEO_PQF	-0.068*** (0.019)	-0.020 (0.041)	-0.086*** (0.018)	-0.070*** (0.019)	-0.028 (0.041)	-0.088*** (0.018)
GOV	0.012 (0.007)	0.009 (0.015)	0.011* (0.007)	0.012 (0.007)	0.009 (0.015)	0.011 (0.007)
BETA	-0.015* (0.009)	-0.039** (0.018)	-0.012 (0.008)	-0.014* (0.009)	-0.038** (0.018)	-0.012 (0.008)
FSIZE	0.132*** (0.008)	0.272*** (0.018)	0.122*** (0.008)	0.132*** (0.008)	0.270*** (0.018)	0.122*** (0.008)
LIQU	-0.021*** (0.004)	-0.060*** (0.008)	-0.017*** (0.003)	-0.020*** (0.004)	-0.058*** (0.008)	-0.016*** (0.003)
ROE	-0.261*** (0.057)	-0.883*** (0.120)	-0.210*** (0.054)	-0.283*** (0.056)	-0.931*** (0.119)	-0.231*** (0.054)
MTB	-0.006** (0.002)	0.006 (0.005)	-0.006** (0.002)	-0.006** (0.002)	0.006 (0.005)	-0.006** (0.002)
LEV	0.055*** (0.010)	0.046** (0.022)	0.053*** (0.010)	0.056*** (0.010)	0.048** (0.022)	0.054*** (0.010)
EVOL	-0.002*** (0.000)	-0.002** (0.000)	-0.001*** (0.000)	-0.003*** (0.000)	-0.001** (0.000)	-0.001*** (0.000)
MA	0.061*** (0.014)	0.081*** (0.030)	0.059*** (0.014)	0.062*** (0.014)	0.086*** (0.030)	0.060*** (0.014)
B_SEG	0.001 (0.002)	0.001 (0.003)	0.002 (0.001)	0.001 (0.002)	0.000 (0.003)	0.001 (0.001)
F_SEG	0.004** (0.002)	0.001 (0.004)	0.005** (0.002)	0.005** (0.002)	0.002 (0.004)	0.005*** (0.002)
EMP	0.000 (0.006)	-0.061*** (0.013)	0.005 (0.006)	0.001 (0.006)	-0.057*** (0.014)	0.005 (0.006)
Constant	4.813*** (0.052)	1.835*** (0.110)	4.729*** (0.050)	4.785*** (0.052)	1.779*** (0.109)	4.702*** (0.049)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,743	3,743	3,743	3,743	3,743	3,743
Adjusted R ²	0.255	0.179	0.256	0.253	0.177	0.255

Panel B: High financial capacity (higher pre-tax income)						
Model	(1)	(2)	(3)	(4)	(5)	(6)
(Variables)	ARC	EXTENSIO NS	TAXONO MY	ARC	EXTENSIO NS	TAXONO MY
TA_ETR	0.373*** (0.064)	0.478*** (0.119)	0.369*** (0.062)			
TA_CETR				0.121** (0.059)	0.340*** (0.109)	0.099* (0.057)

CEO_TENURE	-0.003** (0.001)	-0.007*** (0.002)	-0.002** (0.001)	-0.003** (0.001)	-0.007*** (0.002)	-0.002* (0.001)
CEO_PQF	-0.011 (0.019)	-0.027 (0.035)	-0.005 (0.018)	-0.013 (0.019)	-0.030 (0.035)	-0.007 (0.018)
GOV	0.012* (0.006)	-0.013 (0.011)	0.015** (0.006)	0.013** (0.006)	-0.013 (0.011)	0.016*** (0.006)
BETA	0.013 (0.010)	0.002 (0.018)	0.015 (0.009)	0.013 (0.010)	0.001 (0.018)	0.016* (0.009)
FSIZE	0.092*** (0.011)	0.158*** (0.020)	0.085*** (0.010)	0.100*** (0.011)	0.163*** (0.020)	0.093*** (0.010)
LIQU	-0.021*** (0.005)	-0.050*** (0.010)	-0.017*** (0.005)	-0.019*** (0.005)	-0.047*** (0.010)	-0.015*** (0.005)
ROE	-0.711*** (0.239)	0.374 (0.442)	-0.855*** (0.230)	-0.596** (0.240)	0.523 (0.441)	-0.742*** (0.231)
MTB	-0.015*** (0.002)	-0.023*** (0.004)	-0.014*** (0.002)	-0.015*** (0.002)	-0.022*** (0.004)	-0.014*** (0.002)
LEV	0.114*** (0.010)	0.191*** (0.018)	0.104*** (0.010)	0.113*** (0.010)	0.186*** (0.019)	0.103*** (0.010)
EVOL	-0.001* (0.000)	0.004 (0.000)	-0.003*** (0.000)	-0.004** (0.000)	0.002 (0.000)	-0.003*** (0.000)
MA	0.058*** (0.014)	0.051* (0.027)	0.060*** (0.014)	0.065*** (0.014)	0.061** (0.027)	0.067*** (0.014)
B_SEG	0.002 (0.001)	0.005* (0.002)	0.001 (0.001)	0.001 (0.001)	0.004* (0.002)	0.001 (0.001)
F_SEG	0.004** (0.002)	0.002 (0.003)	0.004** (0.002)	0.005*** (0.002)	0.003 (0.003)	0.005*** (0.002)
EMP	-0.034*** (0.006)	-0.069*** (0.012)	-0.027*** (0.006)	-0.033*** (0.007)	-0.065*** (0.012)	-0.027*** (0.006)
Constant	5.209*** (0.090)	2.500*** (0.166)	5.129*** (0.087)	5.061*** (0.087)	2.380*** (0.160)	4.975*** (0.084)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,744	3,744	3,744	3,744	3,744	3,744
Adjusted R ²	0.155	0.148	0.143	0.148	0.147	0.136

Notes: All variable definitions are presented in Table 2. ***, **, and * indicate significance for a two-tailed test at the 1%, 5% and 10% levels, respectively. Robust standard errors are reported in parentheses.

Table 14. Channel analysis: CEO professional qualifications

Model (Variables)	(1)	(2)	(3)	(4)	(5)	(6)
	ARC	EXTENSIO NS	TAXONO MY	ARC	EXTENSIO NS	TAXONO MY
TA_ETR	0.142*** (0.037)	0.539*** (0.075)	0.093*** (0.036)			
TA_CETR				0.072** (0.036)	0.270*** (0.072)	0.048 (0.035)
CEO_PQF	-0.022 (0.026)	-0.030 (0.052)	-0.016 (0.025)	-0.046* (0.025)	0.036 (0.050)	-0.054** (0.024)
TA_ETR * CEO_PQF	0.076 (0.089)	0.047 (0.179)	0.108 (0.086)			
TA_CETR * CEO_PQF				-0.018 (0.097)	0.315 (0.192)	-0.030 (0.093)
CEO_TENURE	-0.001 (0.001)	-0.005*** (0.001)	-0.000 (0.001)	-0.001 (0.001)	-0.004*** (0.002)	-0.001 (0.001)
GOV	0.010** (0.004)	0.001 (0.008)	0.011*** (0.004)	0.012*** (0.005)	-0.003 (0.009)	0.014*** (0.005)
BETA	0.004 (0.006)	0.002 (0.011)	0.005 (0.005)	0.001 (0.006)	-0.009 (0.013)	0.003 (0.006)
FSIZE	0.122*** (0.005)	0.228*** (0.011)	0.113*** (0.005)	0.124*** (0.006)	0.222*** (0.011)	0.116*** (0.006)
LIQU	-0.021*** (0.003)	-0.045*** (0.005)	-0.018*** (0.002)	-0.020*** (0.003)	-0.055*** (0.006)	-0.016*** (0.003)
ROE	-0.097** (0.039)	-0.608*** (0.078)	-0.050 (0.037)	-0.214*** (0.051)	-0.699*** (0.101)	-0.181*** (0.049)
MTB	-0.010*** (0.001)	-0.009*** (0.003)	-0.010*** (0.001)	-0.010*** (0.002)	-0.011*** (0.003)	-0.009*** (0.001)
LEV	0.081*** (0.006)	0.114*** (0.013)	0.076*** (0.006)	0.084*** (0.007)	0.123*** (0.014)	0.077*** (0.007)
EVOL	-0.003*** (0.000)	-0.002*** (0.000)	-0.001*** (0.000)	-0.002*** (0.000)	-0.002** (0.000)	-0.001*** (0.000)
MA	0.071*** (0.010)	0.071*** (0.020)	0.072*** (0.009)	0.067*** (0.010)	0.074*** (0.020)	0.067*** (0.010)
B_SEG	0.001 (0.001)	0.001 (0.002)	0.001 (0.001)	0.001 (0.001)	0.002 (0.002)	0.001 (0.001)
F_SEG	0.004*** (0.001)	0.002 (0.002)	0.005*** (0.001)	0.005*** (0.001)	0.002 (0.003)	0.005*** (0.001)
EMP	-0.014*** (0.004)	-0.072*** (0.008)	-0.007* (0.004)	-0.016*** (0.005)	-0.061*** (0.009)	-0.010** (0.004)
Constant	4.840*** (0.035)	2.070*** (0.071)	4.747*** (0.034)	4.818*** (0.038)	2.018*** (0.076)	4.726*** (0.037)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8,170	8,170	8,170	7,487	7,487	7,487
Adjusted R ²	0.317	0.247	0.300	0.304	0.257	0.281

Notes: All variable definitions are presented in Table 2. ***, **, and * indicate significance for a two-tailed test at the 1%, 5% and 10% levels, respectively. Robust standard errors are reported in parentheses.

Appendix Table A1. Tax Avoidance and XBRL Reporting Complexity After Excluding Manufacturing Firms

Model (Variables)	(1) ARC	(2) EXTENSIONS	(3) TAXONOMY	(4) ARC	(5) EXTENSIONS	(6) TAXONOMY
TA_ETR	0.202*** (0.033)	0.500*** (0.069)	0.139*** (0.031)			
TA_CETR				0.120*** (0.032)	0.332*** (0.065)	0.079*** (0.030)
CONTROLS	Included	Included	Included	Included	Included	Included
Constant	4.874*** (0.021)	2.387*** (0.044)	4.726*** (0.020)	4.883*** (0.023)	2.424*** (0.047)	4.728*** (0.022)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8,033	8,033	8,033	7,331	7,331	7,331
Adjusted R ²	0.392	0.209	0.427	0.352	0.197	0.381

*Notes: All variable definitions are presented in Table 2. Models (1)–(3) estimate the association between tax avoidance and ARC measures using TA_ETR, while Models (4)–(6) use TA_CETR. ***, **, and * indicate significance for a two-tailed test at the 1%, 5% and 10% levels, respectively. Robust standard errors are reported in parentheses.*