Employees in Tax Havens: A Human Capital Measure of Economic Substance

Michael Mayberry
mayberry@ufl.edu
University of Florida

Mark Zakota
mzakota@umd.edu
University of Maryland, College Park

We are thankful for the helpful comments from Nathan Goldman and Jaron Wilde. Michael Mayberry gratefully acknowledges funding from the Jack Kramer Term Professorship and the Fisher School of Accounting. Mark Zakota is grateful for financial support from the Robert H. Smith School of Business.

Employees in Tax Havens: A Human Capital Measure of Economic Substance

Abstract

Under international tax regulations, multinational corporations must increasingly demonstrate that the economic substance of their tax haven operations justifies their reported profits in these locations. However, due to data limitations, prior research focuses on the mere legal presence rather than the substance of haven operations. We address this gap by introducing *Haven Substance*, a novel proxy for economic substance within tax haven locations. Unlike tax haven indicators based on subsidiaries, *Haven Substance* is based on the proportion of employees located in tax havens using data from employees' online professional profiles. Consistent with *Haven Substance* capturing a distinct dimension of haven operations, we find that it has significantly stronger associations with both tax avoidance and outbound income shifting than traditional tax haven indicators. We also find that firms achieving low tax rates enjoy lower risks of tax settlements when they have greater economic substance, further validating our measure. Overall, we develop and validate a novel proxy for economic substance in tax havens, offering insights into how multinationals leverage haven jurisdictions to reduce their tax burdens.

Keywords: Tax haven, human capital, tax avoidance, income shifting, tax settlements.

JEL Classification: G30; H25; H26; M41.

Data Availability: All data used in this study are obtained from publicly available sources.

1. Introduction

Multinational corporations manage their tax burdens by strategically structuring activities across foreign jurisdictions. Tax havens, in particular, enable firms to achieve significant tax savings, leading to billions in lost tax revenues worldwide (Dyreng and Lindsey, 2009; Dyreng et al., 2019; Law and Mills, 2022; Torslov et al., 2023; Olbert et al., 2024). To counteract the erosion of tax bases, tax authorities increasingly monitor the economic substance behind profit allocations. Most notably, the OECD's Base Erosion and Profit Shifting (BEPS) Action 13 requires multinational corporations to disclose to tax authorities economic indicators, such as revenues, profits, and number of employees, on a disaggregated country-by-country basis. Despite the growing importance of having economic substance in tax-advantaged jurisdictions for justifying profit allocations, the SEC requires firms to disclose only the existence of "significant subsidiaries," resulting in a notable lack of jurisdiction-specific public information on the scale of offshore operations. Due to this lack of data, empirical proxies for economic substance are limited and little research examines the *scale* of tax haven operations as a mechanism for tax avoidance.

We address this gap in the literature by developing a novel proxy for the economic substance of tax haven operations based on labor data. Our methodology overcomes the lack of publicly available country-specific information by leveraging location data from employees' LinkedIn profiles, obtained from Revelio Labs (hereafter Revelio). Our new measure (*Haven Substance*) is based on the proportion of employees located in tax havens relative to the total number of employees of the firm. As such, *Haven Substance* reflects economic substance through the concentration of a firm's human capital stock—a key driver of economic productivity and growth, especially for multinational corporations (Honigsberg and Rajgopal, 2022).

Haven Substance offers several advantages over existing tax-haven-related measures. First, by focusing on a firm's human capital stock in tax haven countries, *Haven Substance* reflects not only the presence but also the scale and scope of tax haven operations. In contrast, current proxies in the literature, which are typically indicators for the existence of a haven subsidiary, do not capture the construct of economic substance, as they only reflect legal presence rather than the scale of haven operations. Second, unlike existing measures, *Haven Substance* is derived independently from corporate disclosures. This approach overcomes a key limitation of traditional measures that rely on annual reports or Exhibit 21 to identify haven operations, as these sources may introduce systematic measurement error due to managers' strategic disclosure of haven entities (Dyreng et al., 2020). Finally, *Haven Substance* can be disaggregated along occupational dimensions to capture the intensity of specific offshore activities, such as manufacturing, administrative, financial, and scientific activities. This granularity allows *Haven Substance* to provide insights into not only the scale and scope but also the functions within haven operations.

Our primary motivation is to extend the literature on tax havens by operationalizing a new measure that reflects the economic substance of tax haven operations. We build on insights from prior research to validate our proposed measure. Although studies on jurisdiction-specific employment are scarce due to data limitations, with most research focusing on simple foreign versus domestic classifications (Drake et al., 2022), Williams (2018) provides small sample evidence on the geographic distribution and determinants of offshore workforce allocations. Therefore, to validate the coverage of our online profile data, we begin by confirming that the geographic distribution of employees in our sample is similar to the distribution reported by Williams (2018). We further assess the coverage of our LinkedIn data by examining the determinants of foreign employment in a *firm-country-year* level analysis and find results

consistent with those of Williams (2018). To validate our novel proxy, *Haven Substance*, we first compare it to other firm characteristics that are associated with haven operations. We provide evidence that *Haven Substance* is positively correlated with the number of geographic segments, number of countries of operation, foreign income, profitability, and R&D intensity. Moreover, we find significantly negative correlations between *Haven Substance* and both cash and GAAP effective tax rates (ETRs), consistent with tax motivated haven operations.

We further assess the relation between *Haven Substance* and tax avoidance in multivariate analyses. Conceptually, tax avoidance through tax havens can involve making real investments to establish economic substance that justifies allocating profits to these locations (De Simone and Olbert, 2022). However, not all real investments are tax motivated and tax considerations are often not a first order determinant of offshore operations (MacCarthy and Atthirawang, 2003; De Simone et al., 2022). Accordingly, we begin by investigating the relation between foreign employees in high-tax, low-tax, and tax haven jurisdictions and tax avoidance, measured by cash and GAAP ETRs. We find that firms with a higher proportion of employees in high-tax jurisdictions face higher cash ETRs. Conversely, firms with a higher proportion of employees located in low-tax non-haven jurisdictions enjoy lower GAAP ETRs. However, only *Haven Substance* has a significantly negative association with both GAAP and cash ETRs. Specifically, a one standard deviation increase in *Haven Substance* lowers cash (GAAP) ETRs by 0.7 (1.1) percentage points. These results are consistent with the notion that *Haven Substance* reflects economic activities that distinctly influence a firm's tax burdens. In contrast, locating employees in non-haven foreign jurisdictions does not seem to be motivated by tax considerations.

Next, to corroborate that *Haven Substance* reflects economic substance rather than merely indicating legal presence or the existence of these operations, we assess if it explains tax avoidance

beyond prior tax-haven-related measures. First, we compare *Haven Substance* to Dyreng and Lindsay's (2009) tax haven indicator, which captures the existence of a significant subsidiary in a tax haven country based on Exhibit 21 disclosures. Consistent with the notion that *Haven Substance* captures a distinct dimension of haven operations, we find that it is significantly associated with tax avoidance even after controlling for the existence of an Exhibit 21 tax haven subsidiary. Second, we investigate whether *Haven Substance* explains tax avoidance beyond Law and Mills' (2022) *Active Haven* measure, which identifies the presence of a significant haven subsidiary that is likely more than a dormant shell company by supplementing Exhibit 21 data with textual analysis of financial statements. Notably, we find that *Haven Substance* is associated with tax avoidance even when controlling for Law and Mills' (2022) active haven indicator. Moreover, *Haven Substance* has a significantly stronger association with tax avoidance than either of the two tax haven subsidiary indicators. Collectively, these results corroborate that *Haven Substance* captures unique information about the economic substance within tax havens, which is not reflected by prior measures that only indicate the presence of tax haven operations.

We complement our ETR analyses by examining a particular tax avoidance strategy, income shifting. To the extent that *Haven Substance* captures the economic substance of tax haven operations, we expect it to be associated with greater outbound income shifting. Following Dyreng and Markle (2016), we estimate a system of equations mapping changes in foreign and domestic sales into changes in foreign and domestic income. We find that outbound income shifting significantly increases with the proportion of a firm's employees located in tax havens. Moreover, we find that *Haven Substance* explains income shifting even after controlling for Dyreng and Linsday's (2009) Exhibit 21-based indicator for havens and Law and Mills' (2022) measure for the presence of active havens. These results corroborate the role of economic substance, as

reflected by *Haven Substance*, as a mechanism behind firms' income shifting strategies.

Having validated *Haven Substance* in the context of lower ETRs and greater income shifting, we next consider the tax settlements associated with these tax benefits. In particular, if locating more employees in tax havens provides economic substance in tax planning strategies, we expect firms with lower tax rates and greater *Haven Substance* to face fewer settlements with tax authorities. We find that firms with lower ETRs are more likely to face tax settlements and incur larger settlement amounts, consistent with prior research (Dyreng et al., 2019). However, both the likelihood and magnitude of these settlements decrease with greater *Haven Substance*. These results are consistent with economic substance in tax havens mitigating the risk of facing challenges by tax authorities, further validating our measure.

Last, in exploratory analysis, we exploit the granularity of the employment data by considering whether particular occupations in tax havens facilitate greater tax savings. We separate tax haven employees into those engaged in core operations (e.g., operating, engineering, and scientific) and those in support functions (e.g., marketing, admin, and finance). We primarily find significant tax benefits for firms whose haven employees support core operations, suggesting that different types of employees augment the economic substance behind tax positions in unique ways.

We acknowledge several limitations of our measure and analyses. First, our analyses are associational and aim to validate *Haven Substance* in the contexts of tax avoidance, income shifting, and tax settlements. Therefore, we do not draw causal inferences from our results. Second, *Haven Substance* relies on employees' disclosures on LinkedIn, which could introduce noise into our analyses. Such noise is unlikely to systematically bias our analyses, as LinkedIn disclosures are by employees rather than management, and several tests corroborate the representative coverage of our LinkedIn data. Nonetheless, the voluntary nature of these disclosures may limit

the measure's effectiveness in capturing economic substance within certain contexts. Finally, *Haven Substance* captures economic substance through employee locations alone, without considering economic substance from physical assets due to data limitations. Although this focus narrows the measure's scope, we believe it is justified for several reasons: (1) employees are among the most important value-producing resources of multinational corporations; (2) the placement of physical assets is often tied to the placement of employees; and (3) the OECD's profit attribution guidelines emphasize the central role of employees in establishing economic substance and directly specify that physical assets without personnel, such as automated equipment, are insufficient to justify profit allocations (OECD, 2010).

We make several contributions to the literature. First, we extend prior research on tax haven operations by developing and validating a new human-capital-based measure of economic substance within these operations. Prior research finds that the presence of significant subsidiaries in tax havens yields tax benefits, especially when these subsidiaries have active intercompany transactions (Dyreng and Lindsay, 2009; Law and Mills, 2022). However, due to the limitations of Exhibit 21 data, prior measures of haven operations only capture the existence of a subsidiary in a tax haven, which does not reflect the true economic substance of haven operations. For example, Exhibit 21 data provide only simple indicators, and the vast majority of our sample firms report at least one subsidiary in a tax haven. Unlike existing measures, *Haven Substance* captures the scale of tax haven operations by focusing on the concentration of a firm's human capital stock in these locations. Validating the measure, we provide evidence that firms derive additional tax benefits from more extensive haven operations, beyond those obtained by simply incorporating or transacting with a haven subsidiary. These findings can guide regulators in understanding the benefits of public country-by-country reporting and aid external stakeholders in assessing the tax

uncertainties of foreign operations. In addition, our findings highlight the usefulness of public, employee-initiated disclosures in identifying tax-motivated offshore operations and evaluating the mechanisms behind firms' tax strategies. Accordingly, we answer Olbert et al.'s (2024, p.4) call to investigate "the magnitudes and exact mechanisms of [multinational firms'] tax haven activities."

Second, we contribute to the literature on income shifting. Klassen and Laplante (2012) find that firms shift income out of the U.S. into jurisdictions with relatively lower tax rates, while other studies identify specific transfer pricing channels such as intangible assets (Grubert, 2003) and intercompany debt (Mintz and Smart, 2004). Most related to our study, Drake et al. (2022) establish the conceptual link between foreign employment and income shifting, finding that firms with higher-than-expected foreign employees engage in greater income shifting without a concomitant increase in tax uncertainty. We contribute beyond Drake et al. (2022) by developing a human-capital-based measure of economic substance that draws directly on granular, jurisdiction-specific data.[1] While Drake et al. (2022) treat all foreign employees as a homogenous group, we link the concentration of employees specifically in tax havens—but not other offshore locations—to both income shifting and tax avoidance. Our findings underscore the importance of distinguishing between economic substance in tax havens and in other offshore locations not driven by tax considerations. As such, our study not only introduces the first measure of U.S. multinationals' economic substance in tax havens but also answers Lester and Olbert's (2024, p.35) call to "shed light on the actual activity in havens and quantify tax and non-tax motives" behind global operations, extending the insights from Drake et al. (2022).

[1] Empirically, Drake et al. (2022) construct an indicator variable to "capture firm-years that have more or less foreign employment than expected based on economic fundamentals" (p.187). Drake et al.'s (2022) proxy differs from *Haven Substance* in several key regards: (1) it is a binary variable capturing higher-than-expected foreign employment, not a continuous measure of economic substance; (2) it treats all foreign employees as a homogeneous group; (3) it relies on segment disclosures of foreign employees, which are voluntary disclosures that follow the "management approach" and are therefore susceptible to bias from managerial discretion; and (4) it is derived from the error term of a first-stage regression, making the resulting proxy a black box, indirect, and potentially sensitive to model specification.

Finally, we contribute to the literature on tax settlements and tax uncertainty. Prior research indicates that foreign operations are associated with uncertain tax benefits (Cazier et al., 2010) and that tax avoidance increases such uncertainty, especially for firms with tax haven operations (Dyreng et al., 2019). However, we find that economic substance in tax havens, as captured by *Haven Substance*, mitigates both the risk and severity of challenges by tax authorities for certain tax strategies. Thus, we provide important nuance that not all tax haven operations are equally uncertain and more substantial tax haven operations may shield the firm from challenges by tax authorities. Moreover, our results suggest that firms might strategically locate productive resources in tax havens to substantiate the economic fundamentals of their tax positions—an important consideration given the growing regulatory pressure on multinationals to justify their tax strategies based on economic substance. Our measure can aid future research in evaluating the real consequences of these regulatory pressures and new anti-tax-avoidance policies.

2. Literature Review and Motivation for *Haven Substance*

Multinational firms regularly structure their operations, investment, and financing to pass through tax-advantaged jurisdictions known as tax havens. In addition to directly generating profits through haven operations, multinationals can utilize various strategies to realize income in these locations. Common strategies include intracompany debt, cost sharing arrangements, transfer pricing, and management consulting fees, among others (e.g., Dyreng and Markle, 2016; Drake et al., 2022; Kim, 2024; Hepfer et al., 2023). However, IRC Section 7701(o) requires all intercompany transactions, even common ones, to possess economic substance. This requirement is implemented through a two-pronged test, namely that (1) a transaction must meaningfully change the taxpayer's economic position beyond tax considerations, and (2) the taxpayer must have a non-tax-related purpose for the transaction. These provisions aim to ensure that value

creation aligns with reported profits across jurisdictions, emphasizing the economic impact over the legal form of transactions (known as "substance over form"). If an intercompany transaction lacks economic substance, tax authorities may challenge the arrangement, increasing the threat that the firm must pay back taxes, penalties, and interest.

Prior research indicates that firms with tax haven subsidiaries reap substantial tax benefits through intercompany transactions (Olbert et al., 2024). Specifically, Dyreng and Lindsay (2009) find that firms disclosing a tax haven in their Exhibit 21 filings have approximately 1.5 percentage points lower ETRs than firms without such subsidiaries. Supplementing the Exhibit 21 data with disclosures in annual reports, Law and Mills (2022) find that tax savings are concentrated among firms that engage in intercompany transactions with their haven subsidiaries (i.e., firms with an "active" tax haven subsidiary).

Despite the importance of tax havens in firms' tax strategies, disclosures about the scale of tax haven operations are notably sparse. The SEC requires firms to publicly disclose the name and incorporation location of "significant subsidiaries" in Exhibit 21 filings. While useful, such Exhibit 21 disclosures fail to provide information to external stakeholders across several dimensions. First, firms have discretion in how they define significant subsidiaries (Gramlich and Whiteaker-Poe, 2013; Dyreng et al., 2020) and therefore can selectively withhold relevant information in an attempt to avoid scrutiny.[2] Second, disclosures on the locations of incorporation, rather than the countries of actual operations, do not reflect the increasing importance of economic substance over legal form in tax considerations. Third, Exhibit 21 contains no information about the scale of operations in a given jurisdiction, as it merely indicates the legal incorporation of significant subsidiaries without any operational details.

[2] Per Dyreng et al. (2020, p.648), the SEC defines a significant subsidiary as "any subsidiary whose assets are greater than 10% of consolidated assets or whose income is greater than 10% of consolidated income."

In contrast to the sparse public disclosures in Exhibit 21, many tax authorities possess more granular information regarding firms' operations in foreign jurisdictions, including tax havens. Since 1983, Congress has required firms to file Form 5471 for every controlled foreign corporation within their organizational structure, providing the IRS with extensive details about the financial position and performance of these subsidiaries, and thus the scale of firms' tax haven operations. More recently, the OECD implemented BEPS Action 13, requiring certain large multinational enterprises to disclose to tax authorities detailed information about their sales, taxes paid and accrued, employment, and capital, disaggregated by jurisdiction.[3] Reflecting the usefulness of these detailed disclosures to tax authorities, Joshi (2020) finds that the implementation of country-by-country reporting under BEPS Action 13 reduced both tax avoidance and income shifting.

Thus, while tax authorities increasingly utilize detailed jurisdiction-specific information, researchers face significant limitations when strictly relying on Exhibit 21 data to assess the economic substance of a firm's tax haven operations. To address this gap, we develop and validate a novel measure of economic substance in tax havens—*Haven Substance*—by leveraging granular, jurisdiction-level employment data from LinkedIn (see Section 3.2. for details). Our approach is grounded in the notion that a firm's human capital stock is a fundamental factor of production and a central component of economic substance determination under international tax guidelines. Indeed, the OECD's guidance on profit attributions emphasizes that employee presence is essential for establishing economic substance and that tangible assets without employees (e.g., automated equipment) do not justify claims of economic substance (OECD, 2010).[4] Therefore, in addition to

[3] Under European Union (EU) Directive 2021/2101, multinational firms with consolidated revenues exceeding 750 million euros are required to disclose tax information disaggregated by country, beginning with the fiscal year 2025.
[4] Further underscoring the central role of human capital, Petruzzi and Myzithra (2020) review the OECD's guidelines on determining the economic substance of permanent establishments (PEs), noting that "in case the PE is lacking people performing functions in it, no risks and assets shall be attributed to the PE" (p.435).

directly indicating substantial operations, jurisdiction-specific human capital has a key role in meeting the OECD's substantive criteria for profit allocations.

Although our primary objective is to validate a novel measure of economic substance in tax havens rather than test theoretical relations, our analysis is guided by theory and the institutional features of economic substance. Specifically, if *Haven Substance* captures economic substance in tax havens that enable a firm to attribute profits to these locations, we expect these tax benefits to result in lower ETRs. Accordingly, we state our primary hypothesis in alternative form as follows:

Hypothesis: *Haven Substance is negatively associated with both cash and GAAP ETRs.*

3. Sample Selection, Measurement, and Descriptive Statistics

3.1. Sample Selection

Table 1 presents our sample selection criteria. We begin the sample with U.S.-incorporated multinational public firms in Compustat with assets over $10 million, and positive pre-tax income, market-to-book ratios, and sales. We focus on multinational firms, identified through foreign income (Compustat: PIFO), because domestic-only firms face distinct tax-related risks and opportunities (Klassen and Laplante, 2012).[5] Similarly, we remove loss-firms due to their distinct tax incentives (De Simone et al., 2017). Consistent with prior studies, we exclude financial firms (SIC 6000–6999) and utilities (SIC 4900–4949). We also require non-missing Exhibit 21 data to identify the jurisdiction of incorporation for significant subsidiaries.

We rely on detailed workforce data from Revelio to develop our proxies for economic substance. Revelio is a leading provider of workforce data that contains employment information from over 380 million public profiles on online professional platforms such as LinkedIn. Detailed

[5] We repeat our analyses without removing domestic-only firms. The results are robust and similar (untabulated).

LinkedIn data is available from 2008 and covers more than 5,000 public firms globally (Li et al., 2022; Liang, et al., 2023). We remove firms with missing Revelio data and provide details about Revelio and our methodology in Section 3.2. Finally, we remove observations with insufficient data to measure control variables and ETRs. Our final sample consists of 7,734 firm-year observations, spanning the period from 2008 through 2021 (the period we obtain from Revelio).

3.2. Measuring Economic Substance

3.2.1. Validation: LinkedIn data coverage

We collect individual-level employment data from Revelio, which includes the comprehensive work history of individuals, detailing job start and end dates, titles, geographic locations, and company names from LinkedIn. Prior studies validate the representative coverage of this data in various contexts, such as general labor markets (Li, 2024); employee demographics (Liang et al., 2023; Berry et al., 2024); and engineering, scientific, accounting, and compliance occupations, among others (Ge et al., 2016; Renschler et al., 2023).[6] Reflecting its wide coverage, LinkedIn is the most popular networking platform among employees (Garg and Telang, 2018).

As our primary objective is to use employees' location information to capture economic substance that is relevant for tax considerations, we further validate the coverage of LinkedIn in the specific context of our study. To this end, we first compare our sample with that of Williams (2018), who present the top 20 countries ranked by employment offshoring based on a small sample covered by the Department of Labor's Trade Adjustment Assistance (TAA) program. Appendix A Part I presents the top 20 countries in our final sample, ranked by average annual total employee counts. Even though Williams (2018) only observes offshoring due to layoffs, whereas

[6] Prior research uses LinkedIn data to assess the link between firm performance and various human capital attributes, such as turnover, diversity, and expertise (Li et al., 2022; Basu et al., 2023; Liang, et al., 2023). Recent studies also use this data to assess financial reporting quality (Renschler et al., 2023) and identify industry-peers (Li, 2024).

we measure firms' country-specific human capital stock, we find that 15 out of 20 countries appear in both lists, corroborating that the top offshore destinations based on LinkedIn align with those based on TAA data.

Second, we follow Williams (2018) and analyze the association between foreign statutory tax rates and employment in a firm-country-year level determinants analysis. Table A1 in Appendix A, Part II presents the results and details about the research design. We find a significantly negative association between a country's tax rate and foreign employment ($p < 0.10$), consistent with Williams (2018).

Third, we augment the firm-country-year level determinants test by including an indicator variable equal to one if a firm mentions the specific country in its Exhibit 21 in a year and zero otherwise. We find a significantly positive association between a firm's presence in a country based on Exhibit 21 and the number of its foreign employees located there ($p < 0.01$), further corroborating the coverage of the LinkedIn data.

Finally, in Figure 1, we plot the geographic distribution of employee concentrations by country in our final sample. The figure indicates that the coverage of Revelio is extensive, capturing a wide range of developed and developing countries. The highest concentrations of employees are in countries with large economic footprints, such as China and India, and in countries with close economic or cultural ties to the U.S., such as Canada, Mexico, and Brazil.

Collectively, our validation tests indicate that: (1) the top offshore locations based on LinkedIn and the TAA program are similar; (2) the country-specific factors influencing foreign employment are consistent with expectations; (3) disclosures in Exhibit 21 align with employee locations in LinkedIn; and (4) LinkedIn has a global coverage of employee locations. While these tests enhance our confidence in the representative coverage of the LinkedIn data, we acknowledge

that potential data incompleteness could introduce noise into our analyses. Such noise is unlikely to systematically bias our analyses but may limit the utility of LinkedIn data in developing and validating a measure of economic substance.

3.2.2. Measuring economic substance

To develop a proxy for economic substance that reflects tax-motivated variation in the scale of offshore operations, we begin by aggregating the individual-level LinkedIn data at the firm-year level based on employees' historical job-specific location information.[7] Before aggregating the data, we remove internships and jobs with missing start dates and job titles, consistent with prior research (Chen et al., 2021; Babina et al., 2024). We also exclude jobs that lack an end date and have not been updated for over five years to reduce noise from outdated or abandoned profiles. We aggregate this data at the firm-year level, including all current employees, to measure a firm's total human capital stock (*Emp LinkedIn*).

We calculate *Haven Substance* as the number of employees in tax haven countries scaled by the total number of employees of the firm. We focus on tax havens due to their very low corporate tax rates and other tax advantages that enable multinational corporations to significantly reduce their overall tax burdens (Dharmapala and Hines, 2009; Olbert et al., 2024). Following Dyreng and Lindsey (2009), we classify a country as a tax haven if it is identified as a haven country by at least three of the four following sources: (1) Organisation for Economic Co-operation and Development, (2) the U.S. Stop Tax Havens Abuse Act, (3) the International Monetary Fund, and (4) the Tax Research Organization. Appendix B presents the list of countries classified as tax havens. Due to the need to align economic activities with taxation to justify profit allocations,

[7] Revelio's Workforce Dynamics data also provides employee counts at various levels of aggregation. We rely on the individual-level data because Revelio retroactively adjust employee counts in Workforce Dynamics after mergers and acquisitions (Revelio, 2024). Therefore, numbers from Workforce Dynamics may not reflect actual employee counts in a given year. In addition, the individual data allows us to remove interns and perform additional data cleaning.

along with the central role of employees as value producing assets, we expect *Haven Substance* to reflect real investments that are associated with tax incentives.

Figure 2 plots *Haven Substance* by year, indicating that the concentration of employees in tax havens has increased in recent years. Appendix C presents the top 20 tax haven countries ranked by average annual total employee counts. The tax haven countries with the highest employee counts are Singapore, the Netherlands, Ireland, Switzerland, and Costa Rica.

To compare the tax benefits of haven operations with those of other foreign operations, we construct two other variables, *Low-Tax Substance* and *High-Tax Substance*, which quantify employee concentrations in low-tax and high-tax jurisdictions, respectively. We define a country as low-tax (high-tax) if it is a non-haven foreign jurisdiction with an average statutory tax rate below (above) the median statutory tax rate across the entire sample period.[8] These foreign countries do not design their tax policies to attract multinational firms by facilitating very low tax rates. Therefore, we do not expect economic substance within low-tax and high-tax jurisdictions to be associated with tax incentives.

Additionally, we use two indicators from prior research to assess the extent to which *Haven Substance* provides incremental information over the presence of tax haven subsidiaries. Following Dyreng and Lindsay (2009), we define *Exh21 Tax Haven* as an indicator equal to one if a firm discloses a significant subsidiary in a tax haven in its Exhibit 21 filing. Similarly, we follow Law and Mills (2022) and define *Active Haven* as an indicator equal to one if a firm engaged in an intercompany transaction with a haven subsidiary.[9]

[8] We classify countries using the average statutory tax rate across the entire sample period to ensure that variation in *Low-Tax Substance* and *High-Tax Substance* reflects changes in employee concentrations rather than changes in country classifications. The median statutory tax rate is 25% in our final sample.

[9] We thank the authors for generously sharing the data. *Active Haven* is available until 2016. Therefore, in analyses employing *Active Haven*, we have a reduced sample size.

3.3. Descriptive Statistics

Table 2 Panel A presents descriptive statistics. *Haven Substance*, *Low-Tax Substance*, and *High-Tax Substance* have a mean value of 0.031, 0.122, and 0.237, respectively. These proportions indicate that the average firm has approximately 39 (= 3.1 + 12.2 + 23.7) percent of its employees located in foreign jurisdictions, similar to the 45 percent reported by Drake et al. (2022), corroborating the convergent validity of the measures. *Exh21 Tax Haven* and *Active Haven* have means of 0.833 and 0.530, respectively. Panel B presents average *Haven Substance* by Fama-French 12 industry. Most industries have an average *Haven Substance* between 0.020 and 0.040. The highest values of *Haven Substance* are in the Healthcare and Chemicals industries, with 0.055 and 0.047 respectively. Table 3 presents pairwise correlations. *Haven Substance* is significantly and negatively correlated with both cash and GAAP effective tax rates (*Cash ETR* and *GAAP ETR*) and significantly and positively correlated with the percentage of foreign income (*Foreign Income*) and R&D intensity (*RD Intensity*) ($p < 0.05$). These correlations provide initial validation that *Haven Substance* reflects tax motivated real operations.

4. Economic Substance and Tax Avoidance

4.1. Tax Avoidance: Research Design

We examine the relation between economic substance and tax avoidance using an ordinary least squares (OLS) regression with standard errors clustered by firm, presented in Equation 1:

$$\begin{aligned} \textit{Effective Tax Rate} = {} & \beta_0 + \beta_1 \textit{Haven Substance} + \beta_2 \textit{Low-Tax Substance} \\ & + \beta_3 \textit{High-Tax Substance} + \Sigma\beta_i \textit{Controls} + \textit{Fixed Effects} + \varepsilon \end{aligned} \quad (1)$$

The dependent variable is either *Cash ETR* or *GAAP ETR*, which represent tax avoidance with the firm's effective tax rates. We calculate *Cash ETR* (*GAAP ETR*) as total cash taxes paid during the year (total tax expense), divided by pretax book income less special items. We trim both ETR

variables at [0,1] to eliminate outliers, consistent with prior research (Gupta and Newberry, 1997; Williams, 2018).[10] The independent variables of interest—*Haven Substance*, *Low-Tax Substance*, and *High-Tax Substance*—represent economic substance through the concentration of employees in tax haven, low-tax, and high-tax jurisdictions. We define all variables in Appendix D.

We include several control variables following prior research (Dyreng and Lindsey, 2009; Law and Mills, 2022). First, we control for firm size and growth opportunities with total assets (*LnSize*) and market-to-book ratio (*MB*). Second, we control for profitability and risk using return on assets (*ROA*), leverage ratio (*Leverage*), stock compensation expense (*Stock Comp*), deferred revenue (*Deferred Income*), other comprehensive income (*Comph Income*), the presence and change in tax loss carryforwards (*NOL* and *NOL Change*), and the presence of negative retained earnings (*Deficit*). Third, we control for book-tax differences in the reporting environment using equity method income (*Equity Earnings*), R&D expenses (*RD Intensity*), intangible assets (*Intangible*), and long-term fixed assets (*PPE*), all scaled by total assets. Fourth, we control for tax planning opportunities and the scale of international operations with the proportion of foreign income (*Foreign Income*), number of foreign countries with a significant subsidiary (*LnCountries*), and number of geographic segments (*LnGeo Segments*). Finally, as our independent variables of interest proxy for economic substance through human capital, we also control for human capital stock using the total number of employees of the firm (*LnEmp*). We further include year fixed effects to control for time-specific factors, such as regulatory scrutiny, and industry fixed effects to control for industry-specific factors that may influence foreign employment and tax planning.

4.2. Tax Avoidance: Results

Table 4 presents the results of estimating the relation between economic substance and tax

[10] Our untabulated results are robust and similar when we winsorize negative ETRs and ETRs over one to fall between zero and one.

avoidance. The dependent variable in the first (last) three columns is *Cash ETR* (*GAAP ETR*). Columns 1 and 4 present the results of estimating Equation 1 with *Haven Substance*, Columns 2 and 5 present the results with *Low-Tax Substance* and *High-Tax Substance*, and Columns 3 and 6 present the results with all three proxies included in the model. The results indicate a significantly negative relation between *Haven Substance* and both *Cash ETR* ($p < 0.01$) and *GAAP ETR* ($p < 0.01$). Economically, these results indicate that a one standard deviation increase in *Haven Substance* is associated with a 0.7 percentage point decrease in *Cash ETR* and a 1.1 percentage point decrease in *GAAP ETR*.[11] In contrast, while the association between *Low-Tax Substance* and *GAAP ETR* is significantly negative, the variable is not significantly associated with *Cash ETR*. This pattern is consistent with the results of Williams (2018). The coefficients on *High-Tax Substance* are also inconsistent, indicating a significantly positive association with *Cash ETR* ($p < 0.01$) but an insignificant association with *GAAP ETR*. Collectively, these results corroborate the notion that not all offshore operations are tax motivated and highlight the importance of economic substance in tax havens for facilitating tax avoidance.

4.3. The Presence and Substance of Tax Haven Operations

Next, we investigate whether *Haven Substance* provides incremental explanatory power in explaining tax avoidance beyond subsidiary-related measures of tax haven presence proposed by Dyreng and Lindsay (2009) and Law and Mills (2022). To assess the incremental explanatory power of *Haven Substance,* we re-estimate Equation 1 after replacing *Low-Tax Substance* and *High-Tax Substance* with either *Exh21 Tax Haven* or *Active Haven*.

Table 5 Panel A presents the results when *Exh21 Tax Haven* is included in the model. In Columns 1 and 3, we report the baseline regression results for *Cash ETR* and *GAAP ETR*,

[11] We calculate economic significance as follows: 0.040 x -0.177 = -0.007 for *Cash ETR* and 0.040 x -0.263 = -0.011 for *GAAP ETR*.

respectively, without *Haven Substance* in the model. Consistent with Dyreng and Lindsay (2009), *Exh21 Tax Haven* is significantly and negatively associated with both *Cash ETR* and *GAAP ETR* ($p < 0.05$). Next, we include *Haven Substance* and re-estimate the model for *Cash ETR* and *GAAP ETR* in Columns 2 and 4, respectively. The coefficient on *Haven Substance* remains significantly negative in both models ($p < 0.01$). This result is consistent with the scale of haven operations providing incremental information about firms' tax avoidance strategies, beyond that contained in significant subsidiary disclosures. Moreover, in both the *Cash ETR* and *GAAP ETR* specifications, the coefficient on *Haven Substance* is significantly more negative than that on *Exh21 Tax Haven* ($p < 0.01$). This result suggests that *Haven Substance* provides greater explanatory power concerning firms' tax burdens than Dyreng and Lindsay's (2009) haven indicator.

Table 5 Panel B presents the results of estimating Equation 1 when *Active Haven* is included in the model. Consistent with Law and Mills (2022), we find that *Active Haven* is significantly and negatively associated with both *Cash ETR* ($p < 0.05$ in Column 1) and *GAAP ETR* ($p < 0.01$ in Column 3), suggesting that firms that mention transactions with tax haven subsidiaries enjoy lower ETRs. When we include *Haven Substance* in the model in Columns 2 and 4, it remains significantly and negatively associated with both ETR measures ($p < 0.01$). Moreover, in both models, the coefficient on *Haven Substance* is significantly more negative than that on *Active Haven* ($p < 0.01$), further highlighting that information about the scale of haven operations is incrementally informative in understanding multinationals' tax burdens.

Combined, the results in Table 5 indicate that *Haven Substance* outperforms tax haven indicators from prior literature in explaining variations in firms' ETRs. Moreover, these results are consistent with our measure capturing a distinct, previously unobservable dimension of tax haven operations. By moving beyond mere legal presence, *Haven Substance* addresses a key limitation

of existing measures and enhances our understanding of how multinational firms utilize haven jurisdictions to reduce their tax burdens.

5. Additional Analyses

5.1. Income Shifting

In this section, we further validate *Haven Substance* by focusing on a specific tax planning strategy: income shifting. This analysis is motivated by the widespread use of income shifting as a predominant tax strategy and the introduction of regulatory measures such as the OECD's BEPS project, which aims to restrict income shifting through tax havens by ensuring that reported profits correspond with the economic substance in these locations. To investigate income shifting, we follow Dyreng and Markle (2016) and estimate the following system of equations:

$$\Delta PIFO = \alpha + (1-\gamma)\rho_f \Delta SALEFO + \theta\rho_d \Delta SALEDOM + \varepsilon \quad (2a)$$

$$\Delta PIDOM = \beta + \gamma\rho_f \Delta SALEFO + (1-\theta)\rho_d \Delta SALEDOM + u \quad (2b)$$

PIFO represents foreign pretax income, *PIDOM* represents domestic pretax income, *SALEFO* represents foreign sales, *SALEDOM* represents domestic sales, and Δ denotes a first difference. We estimate the outbound (θ) and inbound (γ) income transfer parameters, along with the foreign (ρ_f) and domestic (ρ_d) return on sales parameters, as follows:

$$\theta = \theta_0 + \theta_1 \textit{ Haven Substance} + \theta_2 \textit{ Exh21 Tax Haven} + \Sigma\, \theta_n \textit{ Controls} \quad (2c)$$

$$\gamma = \gamma_0 + \gamma_1 \textit{ Haven Substance} + \gamma_2 \textit{ Exh21 Tax Haven} + \Sigma\, \gamma_n \textit{ Controls} \quad (2d)$$

$$\rho_f = \rho_{f0} + \rho_{f1} \textit{ Haven Substance} + \rho_{f2} \textit{ Exh21 Tax Haven} + \Sigma\, \rho_{fn} \textit{ Controls} \quad (2e)$$

$$\rho_d = \rho_{d0} + \rho_{d1} \textit{ Haven Substance} + \rho_{d2} \textit{ Exh21 Tax Haven} + \Sigma\, \rho_{dn} \textit{ Controls} \quad (2f)$$

Conceptually, economic substance enhances a firm's ability to justify allocating profits to tax havens. Therefore, if *Haven Substance* captures economic substance, we expect θ_1 to be significantly positive, where θ_1 represents tax-motivated outbound income shifting attributable to

the concentration of employees in tax havens. Consistent with Dyreng and Markle (2016), we control for the presence of a significant subsidiary in tax havens (*Exh21 Tax Haven*), return on sales (*ROS*), R&D expenses (*RD Intensity*), advertising expenses (*Adv Intensity*), intangible assets (*Intangible*), cash (*Cash*), debt (*Leverage*), firm size (*Size*), and the ratio of foreign sale to total sales (*Foreign Sales*).

Table 6 Panel A presents the results of estimating Equations 2a and 2b. We find that θ_1 is significantly positive across all specifications ($p < 0.01$), consistent with our expectation that economic substance, as captured by *Haven Substance*, facilitates tax-motivated outbound income shifting. This result supports the notion that *Haven Substance* reflects the level of substantial operations that enable a firm to reduce its tax burdens, further validating our measure. Similar to Dyreng and Markle (2016), we find a significantly positive outbound transfer coefficient on *Exh21 Tax Haven* in Columns 2 ($p < 0.05$) and 3 ($p < 0.10$), indicating that firms utilize significant tax haven subsidiaries for income shifting. Notably, in Column 3, where both variables are included, the coefficient on *Haven Substance* is significantly larger than on *Exh21 Tax Haven* ($p < 0.05$). This result implies that economic substance in the form of employee concentrations facilitates income shifting above and beyond the mere presence of a significant subsidiary.

Next, we include Law and Mills' (2022) *Active Haven* indicator in Panel B of Table 6. We find a similar pattern of results. Specifically, the outbound transfer coefficient on *Haven Substance* (θ_1) remains significantly positive ($p < 0.05$) in Column 3, even after including *Active Haven* in the model. In contrast, the outbound transfer coefficient on *Active Haven* (θ_3) becomes insignificant in Column 3. Moreover, F-tests indicate that θ_1 is significantly larger than the coefficient for outbound transfers on *Active Haven* ($p < 0.10$). Taken together, the income shifting results validate *Haven Substance* beyond the context of ETRs, indicating that it is a better predictor

of income shifting than measures capturing legal presence through subsidiary incorporations.

5.2. Tax Settlements

Thus far, our evidence suggests that economic substance in tax havens, as reflected by *Haven Substance*, facilitates tax avoidance and income shifting. We now turn to assessing the extent to which such tax avoidance impacts firms' tax uncertainty. We proxy for uncertainty using settlements recorded under FIN 48 and estimate the following OLS regression:

$$Settlement = \beta_0 + \beta_1 Avoider \times Haven\ Substance + \beta_2 Haven\ Substance + \beta_3 Avoider + \Sigma\beta_i Controls + \Sigma\beta_i Controls \times Haven\ Substance + Fixed\ Effects + \varepsilon \quad (3)$$

The dependent variable is one of three proxies for tax settlements: (1) an indicator equal to one if the firm has a settlement in a given year (*Settlement Indicator*), (2) the natural log of one plus settlements (*LnSettlements*), and (3) the ratio of settlements to total assets (*Settlements/AT*). Our variable of interest in Equation 3 is the interaction between *Avoider* and *Haven Substance*. Similar to Dyreng et al. (2019), we define *Avoider* as an indicator variable equal to one if the firm's *GAAP ETR* is in the bottom tercile of the sample.[12] Consistent with locating employees in tax havens enhancing the economic substance of a firm's tax positions, we expect a significantly negative β_1 coefficient. We include the same controls in Equation 3 as in Equation 1 and, following Drake et al. (2022), we interact them with *Haven Substance*.

Table 7 presents the results of estimating Equation 3. Consistent with prior research indicating that tax avoidance increases tax uncertainty (Dyreng et al., 2019), the coefficients on *Avoider* are significantly positive across all three model specifications ($p < 0.05$). Moreover, the coefficients on the interaction term between *Avoider* and *Haven Substance* are significantly negative in all specifications ($p < 0.05$ in Column 1; $p < 0.01$ in Columns 2 and 3), suggesting that

[12] In untabulated robustness tests, we define *Avoider* using *Cash ETR* and find robust and similar results.

economic substance in tax havens mitigates the risk of challenges by tax authorities for low ETR firms. We confirm that the results are robust and similar when including *Exh21 Tax Haven* or *Active Haven* in the models (untabulated). Collectively, these results further validate *Haven Substance* and highlight the role of economic substance in mitigating the risks associated with tax avoidance strategies.

5.3. Employee Occupation and Tax Avoidance

One advantage of relying on Revelio to identify employment in tax havens is the ability to disaggregate employment by occupation. This allows us to explore whether all types of employees in tax havens enhance economic substances or whether our tax avoidance results are primarily attributable to a subset of occupations. Using job roles identified by Revelio, we disaggregate *Haven Substance* into two distinct variables: *Core Haven Substance*, reflecting the concentration of employees performing core functions (i.e., scientific, operational, or engineering roles), and *Supportive Haven Substance*, reflecting the concentration of employees in supportive roles (i.e., marketing, sales, administrative, or financial roles).

Table 8 presents the results of estimating Equation 1 after replacing *Haven Substance* with *Core Haven Substance* and *Supportive Haven Substance*. In Column 1, when *Cash ETR* is the dependent variable, the coefficient on *Core Haven Substance* is significantly negative ($p < 0.01$), while the coefficient on *Supportive Haven Substance* is statistically insignificant. The results are similar in Column 2, when *Exh21 Tax Haven* is included in the model. Moreover, in both columns, the coefficients on *Core Haven Substance* are significantly larger in magnitude than on *Supportive Haven Substance* ($p < 0.01$; untabulated), consistent with locating core occupations in tax havens enhancing the economic substance of tax positions. However, in the *GAAP ETR* specifications in Columns 3 and 4, we find significantly negative coefficients on both

Core Haven Substance ($p < 0.01$) and *Supportive Haven Substance* ($p < 0.10$). Although the coefficient on *Core Haven Substance* appears larger in magnitude, it is not significantly different from the coefficient on *Supportive Haven Substance*. Thus, we conclude that, in certain contexts, core occupations may provide incrementally greater economic substance to tax haven operations.

6. Robustness Tests

We conduct various untabulated robustness tests to assess the sensitivity of our findings. First, following Schwab et al. (2020), we re-estimate Equation 1 after excluding firms with ETRs below 5 percent and above 40 percent. Our results are robust and similar to those reported in Table 5. Second, we replace the tax haven subsidiary indicator from Dyreng and Lindsay (2009) with a continuous measure based on the percentage of foreign subsidiaries in tax havens and repeat our analyses. We continue to find that *Haven Substance* provides significant incremental explanatory power for tax avoidance and income shifting compared to the Exhibit 21-based measure. This finding further highlights that the jurisdiction-specific data currently available from SEC filings are insufficient for capturing the economic substance of haven operations.

Finally, to evaluate the sensitivity of our income shifting results to the chosen estimation method, we repeat the analysis using the Klassen and Laplante (2012) approach. Specifically, we regress five-year average foreign pretax income scaled by foreign sales on the average tax rate differential between foreign and U.S. tax rates over the same period (*FTR*), *Haven Substance*, their interaction, and a set of control variables following prior research. The coefficient on the interaction between *FTR* and *Haven Substance* is significantly negative, consistent with greater income shifting ($p < 0.01$). We further verify that *Haven Substance* exhibits a significantly stronger association with income shifting than the tax haven indicators from prior research, consistent with the results presented in Table 6.

7. Conclusion

Despite widespread interest in understanding the scale of firms' foreign operations, to date, researchers are limited to simple indicators of whether or not firms have a subsidiary in a given jurisdiction. This paucity of information is particularly salient in the context of tax haven jurisdictions, which offer sizeable tax benefits and potentially decrease revenue in higher tax jurisdictions. To address this gap in the literature, we develop a novel proxy, *Haven Substance*, which reflects economic substance through the percentage of a firm's employees located in tax havens based on data from employees' online professional profiles.

We provide three sets of analyses to validate *Haven Substance*. First, we assess tax avoidance, providing evidence that the associations between *Haven Substance* and ETR measures are both statistically and economically significant. Moreover, the association between *Haven Substance* and tax avoidance is stronger than the associations identified using traditional tax haven indicators. This finding indicates that our measure captures a distinct dimension of firms' haven activities beyond their legal presence, providing incremental insights into tax strategies. Second, we find that *Haven Substance* is significantly associated with tax-motivated outbound income shifting, highlighting its relevance to a specific and prevalent tax strategy. Third, we find that firms achieving lower tax burdens face smaller tax settlements when their tax haven operations possess greater economic substance, as captured by *Haven Substance*. This finding further validates our measure, indicating that economic substance mitigates tax uncertainty. Collectively, these results corroborate that *Haven Substance* reflects economic substance in tax havens, rather than the legal incorporation of subsidiaries, thereby enhancing our understanding of how firms leverage haven operations to reduce their tax burdens.

We recognize that our measure faces various limitations. First, as our analyses primarily aim to validate *Haven Substance*, our evidence is associational rather than causal in nature. Second, although our validation tests presented in Appendix A corroborate the representative coverage of our LinkedIn data, we acknowledge that *Haven Substance* may be subject to measurement error due to the voluntary nature of employees' disclosures on their online profiles. However, unlike disclosures by managers, disclosures by employees are unlikely to be systematically influenced by tax incentives, mitigating the threat of biased inferences. Third, while our focus on human capital aligns with the OECD's emphasis on employees for profit attribution and substance determination, our measure reflects economic substance through human capital alone due to data limitations. Subject to these limitations, our study introduces and validates a novel measure of economic substance in tax havens, providing researchers and regulators with a new tool to assess the scale of a firm's operations in these increasingly important jurisdictions.

References

Babina, T., Fedyk, A., He, A. and Hodson, J., 2024. Artificial intelligence, firm growth, and product innovation. Journal of Financial Economics, 151, p.103745.

Basu, S., Ma, X. and Shen, M., 2023. The Value of Human Capital for Firm Performance: Roles of Individual and Group Expertise. Available at SSRN 4914722.

Berry, A., Maloney, E.M. and Neumark, D., 2024. The Missing Link? Using LinkedIn Data to Measure Race, Ethnic, and Gender Differences in Employment Outcomes at Individual Companies (No. w32294). National Bureau of Economic Research.

Cazier, R., Rego, S.O, Tian, X. and Wilson, R.J. 2010. Early evidence on the determinants of unrecognized tax benefits, Working Paper. Available at SSRN 1578485.

Chen, X., Cheng, Q., Chow, T. and Liu, Y., 2021. Corporate in-house tax departments. Contemporary Accounting Research, 38(1), p.443-482.

De Simone, L., Klassen, K.J., and Seidman, J.K. 2017. Unprofitable affiliates and income shifting behavior, 92(3): p.113-136.

De Simone, L., Klassen, K.J. and Seidman, J.K., 2022. The effect of income-shifting aggressiveness on corporate investment. Journal of Accounting and Economics, 74(1), p.101491.

De Simone, L., and Olbert, M. 2022. Real effects of private country-by-country disclosure. The Accounting Review, 97(6), p. 201-232.

Dharmapala, D., and Hines, J. 2009. Which countries become tax havens? Journal of Public Economics, 93(9), p. 1058-1068.

Drake, K.D., Goldman, N.C. and Murphy, F., 2022. Foreign employment, income shifting, and tax uncertainty. The Accounting Review, 97(2), p.183-212.

Dyreng, S. and Hanlon, M., 2021. Tax avoidance and multinational firm behavior. In chapter in the International Tax Policy Forum/Brookings Institution book: Global Goliaths: Multinational Corporations in the 21st Century Economy.

Dyreng, S.D. and Lindsey, B.P., 2009. Using financial accounting data to examine the effect of foreign operations located in tax havens and other countries on US multinational firms' tax rates. Journal of Accounting Research, 47(5), p.1283-1316.

Dyreng, S.D. and Markle, K.S. 2016. The effect of financial constraints on income shifting by U.S. multinationals. The Accounting Review, 91(6): 1601-1627.

Dyreng, S.D., Hanlon, M., and Maydew, E.L., 2019. When does tax avoidance result in tax uncertainty? The Accounting Review, 94(2): p.179-203

Dyreng, S.D., Hanlon, M., Maydew, E.L. and Thornock, J.R., 2017. Changes in corporate effective tax rates over the past 25 years. Journal of financial economics, 124(3), pp.441-463.

Dyreng, S., Hoopes, J., Langetieg, P., Wilde, J. 2020. Strategic Subsidiary Disclosure. Journal of Accounting Research, 58(3), p. 643-692.

Garg, R., and R. Telang. 2018. To be or not to be linked: Online social networks and job search by unemployed workforce. Management Science 64 (8): p. 3926–3941.

Ge, C., Huang, K.W. and Png, I.P., 2016. Engineer/scientist careers: Patents, online profiles, and misclassification bias. Strategic Management Journal, 37(1), p. 232-253.

Gramlich, J.D. and Whiteaker-Poe, J., 2013. Disappearing subsidiaries: the cases of Google and Oracle. Working Paper. Available at SSRN 2229576.

Grubert, H. 2003. Intangible income, intercompany transactions, income shifting, and the choice of location. National Tax Journal, 56(1): p. 221-242.

Gupta S., and K. Newberry. 1997. Determinants of the variability in corporate effective tax rates: Evidence from longitudinal data. Journal of Accounting and Public Policy 16 (1): p.1–34.
Hepfer, B.F., Wilde, J.H. and Wilson, R.J., 2023. Hidden in Plain Sight: Havens and Captives. Working Paper.
Honigsberg, C. and Rajgopal, S., 2022. Wage wars: The battle over human capital accounting. Harv. Bus. L. Rev., 12, p.275.
Joshi, P. 2020. Does private country-by-country reporting deter tax avoidance and income shifting? Evidence from BEPS Action Item 2013. Journal of Accounting Research, 58(2): p. 333-381.
Kim, J. 2024. Income-shifting by U.S. multinational corporations using stock-based compensation. Working Paper, University of Arizona.
Klassen, K.J. and Laplante, S.K. 2012. Are U.S. multinational corporations becoming more aggressive income shifters? Journal of Accounting Research, 50(5): p. 1245-1285.
Law, K.K. and Mills, L.F., 2022. Taxes and haven activities: Evidence from linguistic cues. The Accounting Review, 97(5), p. 349-375.
Lester, R., and M. Olbert. 2024. Firms' real and reporting responses to taxation: A review. Available at SSRN 4779893.
Li, N., 2024. Labor market peer firms: understanding firms' labor market linkages through employees' internet "also viewed" firms. Review of Accounting Studies, p.1-52.
Li, Q., Lourie, B., Nekrasov, A. and Shevlin, T., 2022. Employee turnover and firm performance: Large-sample archival evidence. Management Science, 68(8), p.5667-5683.
Liang, C., Lourie, B., Nekrasov, A. and Yoo, I.S., 2023. Voluntary disclosure of workforce gender diversity. Available at SSRN 3971818.
MacCarthy, B.L. and Atthirawong, W., 2003. Factors affecting location decisions in international operations–a Delphi study. International journal of operations & production management, 23(7), p.794-818.
Mintz, J. and Smart, M., 2004. Income shifting, investment, and tax competition: theory and evidence from provincial taxation in Canada. Journal of Public Economics, 88(6), p.1149-1168.
Olbert, M., Spengel, C. and Weck, S., 2024. Multinational Firms in Tax Havens–Corporate Motives, Regulatory Countermeasures, and Recent Statistics. Available at SSRN
Organisation for Economic Co-operation and Development (OECD). 2010. Report on the Attribution of Profits to Permanent Establishments. Available at: https://www.oecd.org/en/publications/2010/07/2010-report-on-the-attribution-of-profits-to-permanent-establishments_d659f672.html
Petruzzi, R. and Myzithra, A., 2020. Substance in transfer pricing in a post-BEPS world and beyond. International Transfer Pricing Journal, 27(6), p. 430-447.
Renschler, M.E., Ahn, J., Hoitash, R. and Hoitash, U., 2023. Internal Audit Competency and Financial Reporting Quality: Evidence from LinkedIn Human Capital Data. Auditing: A Journal of Practice & Theory, 42(3), p. 107-136.
Revelio. 2024. FAQs.
Available at: https://www.data-dictionary.reveliolabs.com/faq.html#modeling.
Schwab, C.M, Stomberg, B., and Xia, J. 2022. What determines effect tax rates? The relative influence of tax and other factors. Contemporary Accounting Research, 39(1), p. 459-497.
Tørsløv, T., Wier, L. and Zucman, G., 2023. The missing profits of nations. The Review of Economic Studies, 90(3), p.1499-1534.

Williams, B.M., 2018. Multinational tax incentives and offshored US jobs. The Accounting Review, 93(5), p.293-324.

Appendix A. Validation of LinkedIn Data

Part I: Sample Details

Appendix A, Part I presents country-level summary statistics for the top 20 foreign countries, ranked by average total foreign employees. This ranking is based on the average of annual total employee counts, where we calculate annual totals for each year in our final sample using data from Revelio. The table also includes mean values for gross domestic product (GDP), GDP per capita, and the statutory tax rate for each country over the sample period. GDP and GDP per capita are adjusted for purchasing power parity (PPP).

Rank	Country	Employees	GDP (in billions)	GDP Per Capita	Tax Rate
1	India	276,346.214	7,367.725	5,552.970	33.119
2	United Kingdom	128,157.571	2,764.109	42,537.995	22.357
3	Canada	93,596.071	1,632.836	45,547.603	27.431
4	Brazil	77,165.357	3,051.920	14,929.061	34.000
5	Mexico	63,663.357	2,243.511	18,773.341	29.714
6	China	53,307.571	18,231.373	13,198.978	25.000
7	Australia	51,922.214	1,107.555	46,530.770	30.000
8	France	45,858.929	2,797.488	42,135.977	34.830
9	Germany	41,003.357	3,961.993	48,287.358	30.093
10	Philippines	30,457.643	727.968	7,022.794	30.000
11	Singapore	28,959.286	486.433	89,613.601	17.143
12	Pakistan	28,425.286	961.152	4,560.959	32.643
13	The Netherlands	28,194.714	890.720	52,407.239	25.107
14	Ireland	27,546.643	315.982	66,079.896	12.500
15	Italy	26,550.143	2,363.168	39,492.075	30.079
16	Spain	26,516.286	1,686.893	36,101.746	27.714
17	Poland	20,655.571	1,040.551	27,397.181	19.000
18	Malaysia	19,715.000	760.482	24,503.810	24.643
19	Argentina	18,024.071	895.581	20,757.958	33.929
20	Belgium	17,654.071	529.919	47,110.403	32.076

Appendix A. (Continued)

Part II: Determinants of Foreign Employment

Appendix A, Part II presents the analysis of the determinants of foreign employment at the firm-country-year level. We estimate the following regression specification:

$$\begin{aligned} \textit{Foreign Employment}_{ijt} = {} & \beta_0 + \beta_1\, \textit{Tax Rate}_{jt} + \beta_2\, \textit{Exhibit21 Country}_{ijt} \\ & + \Sigma\, \beta_C\, \textit{Country Controls}_{jt} + \Sigma\, \beta_F\, \textit{Firm Controls}_{it} + \textit{YearFE} \\ & + \textit{IndustryFE} + \textit{RegionFE} + \varepsilon_{ijt} \end{aligned} \tag{A1}$$

We use two proxies for foreign employment as dependent variables. In Columns 1 and 2, the dependent variable is *CountryEmpPct*, calculated as the proportion of employees of firm *i* in country *j* in year *t*. In Columns 3 and 4, the dependent variable is *LnCountryEmp*, calculated as the log of one plus the number of employees of firm *i* in country *j* in year *t*. The main variables of interest are $\textit{Tax Rate}_{jt}$ and $\textit{Exhibit21 Country}_{ijt}$. $\textit{Tax Rate}_{jt}$ represents the statutory tax rate of country *j* in year *t*, obtained from the Tax Foundation. $\textit{Exhibit21 Country}_{ijt}$ is an indicator variable equal to one if firm *i* discloses a significant subsidiary located in country *j* in year *t*, and zero otherwise.

Following prior research (Williams, 2018), $\textit{Country Controls}_{jt}$ is a vector of country-level characteristics that capture economic, political, and labor-related factors that may influence foreign employment. Specifically, we control for the log of GDP in billions (*LnGDP*), *GDP Per Capita*, the percentile rank of the perceived quality of legal rights and enforcement (*Rule of Law*), the percentile rank of the perceived quality of public policies and government (*Government Effectiveness*), taxes on goods and services as a percentage of total government revenue (*Indirect Tax*), income, profit, and capital taxes as a percentage of total government revenue (*Income Tax*), the annual average of the official exchange rate relative to the U.S. dollar (*Exchange Rate*), the ratio of a purchasing power parity conversion factor to the exchange rate (*Price Level Ratio*), the

direct investment equity flows as a share of GDP (*FDI*), the sum of exports and imports of goods and services as a share of GDP (*Trade*), and the annual average of effectively applied tariff rates weighted by the product import shares of each partner country (*Tariffs*). We obtain these variables from the World Bank's World Development Indicators (WDI) database.

Firm Controls$_{it}$ is a vector of firm-level controls, including *LnSize*, *MB*, *RD Intensity*, *ROA*, *Leverage*, *Stock Comp*, *Deferred Income*, *Comph Income*, *Foreign Income*, *Equity Income*, *Intangible*, *PPE*, *LnCountries*, *LnGeo Segments*, and *LnEmp*, defined in Appendix D. In addition, we include industry, year, and region fixed effects to control for sector-, time-, and region-specific unobservable factors that might influence foreign employment. Standard errors are clustered by country, following Williams (2018). Table A1 presents the results of estimating Equation A1.

This analysis validates the use of LinkedIn data to assess the geographic distribution of employees in two regards. First, the negative and significant coefficients on *Tax Rate* in Columns 1, 3, and 4 ($p < 0.10$) align with the small sample evidence of Williams (2018), corroborating the convergent validity of our methodology. Second, the significantly positive coefficients on *Exhibit21 Country* in Columns 1–4 ($p < 0.01$) indicate that countries where the firm has a significant subsidiary have a greater concentration of employees, corroborating that employee locations based on Revelio data align with the location of firms' significant subsidiaries based on their Exhibit 21 disclosures.

Appendix A. (Continued)

Table A1
Determinants of Employee Location

	CountryEmpPct		*LnCountryEmp*	
	(1)	(2)	(3)	(4)
Tax Rate	-0.138*	-0.107	-0.013*	-0.008*
	(-1.69)	(-1.62)	(-1.96)	(-1.80)
Exhibit21 Country		11.740***		1.736***
		(9.63)		(26.23)
LnGDP	3.153***	1.904***	0.485***	0.300***
	(6.02)	(5.73)	(11.04)	(9.99)
GDP Per Capita	-0.000***	-0.000***	-0.000***	-0.000***
	(-3.30)	(-3.97)	(-2.82)	(-3.54)
Rule of Law	0.087**	0.072**	0.007*	0.005*
	(2.17)	(2.25)	(1.92)	(1.85)
Government Effectiveness	-0.046	-0.045	-0.000	-0.000
	(-1.34)	(-1.63)	(-0.13)	(-0.13)
Indirect Tax	-0.011	-0.024	-0.002	-0.003
	(-0.21)	(-0.59)	(-0.34)	(-1.10)
Income Tax	0.113*	0.095*	0.014***	0.011***
	(1.91)	(1.96)	(2.77)	(3.12)
Exchange Rate	-0.000	-0.000*	-0.000	-0.000
	(-1.60)	(-1.69)	(-1.43)	(-1.44)
Price Level Ratio	3.807	1.747	0.411**	0.106
	(1.50)	(0.86)	(1.99)	(0.76)
FDI	0.017	0.012	0.002	0.001
	(1.53)	(1.29)	(1.13)	(0.71)
Trade	0.015*	0.006	0.002*	0.000
	(1.69)	(0.78)	(1.70)	(0.50)
Tariffs	0.043	0.049	-0.010	-0.009
	(0.30)	(0.44)	(-0.69)	(-0.93)
LnSize	0.033	-0.062	0.207***	0.193***
	(0.36)	(-0.78)	(14.56)	(14.86)
MB	0.002	-0.005*	0.005***	0.004***
	(0.61)	(-1.76)	(9.23)	(8.32)
RD Intensity	3.060	3.414	1.570***	1.622***
	(1.20)	(1.49)	(6.54)	(8.14)
ROA	-5.896***	-4.390***	-1.015***	-0.792***
	(-5.69)	(-4.48)	(-9.31)	(-8.18)
Leverage	0.119	0.186	-0.426***	-0.416***
	(0.45)	(0.76)	(-13.57)	(-13.72)
Stock Comp	-5.893	-3.093	0.371	0.785**
	(-1.63)	(-0.86)	(0.97)	(2.32)
Deferred Income	0.125	0.232**	0.052***	0.068***
	(1.37)	(2.55)	(5.09)	(7.29)
Comph Income	5.046***	3.604**	1.103***	0.890***
	(2.94)	(2.38)	(7.78)	(8.22)
Foreign Income	15.208***	13.592***	3.314***	3.075***

	(6.50)	(5.99)	(12.50)	(12.56)
Equity Earnings	-5.549	3.089	3.792***	5.069***
	(-0.72)	(0.43)	(4.03)	(6.44)
Intangible	-0.483***	-0.654***	-0.396***	-0.421***
	(-2.88)	(-3.95)	(-12.33)	(-12.98)
PPE	-1.165***	-0.496	-0.485***	-0.386***
	(-2.76)	(-1.43)	(-10.02)	(-9.25)
LnCountries	0.776***	-0.861***	0.128***	-0.114***
	(5.56)	(-5.95)	(7.52)	(-12.60)
LnGeo Segments	0.298***	0.339***	-0.009	-0.003
	(2.79)	(3.18)	(-1.19)	(-0.44)
LnEmp	0.122	0.039	0.175***	0.163***
	(1.09)	(0.39)	(10.35)	(10.13)
Observations	585,289	585,289	585,289	585,289
Adjusted R^2	0.180	0.278	0.416	0.533
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Region FE	YES	YES	YES	YES

Appendix B. List of Tax Haven Countries

Tax Haven Countries	
Andorra	Lebanon
Anguilla	Liberia
Antigua and Barbuda	Liechtenstein
Aruba	Luxembourg
Bahamas	Macao
Bahrain	Maldives
Barbados	Malta
Belize	Marshall Islands
Bermuda	Mauritius
Botswana	Monaco
British Virgin Islands	Montserrat
Brunei Darussalam	Nauru
Cabo Verde	Netherlands, The
Cayman Islands	Panama
Cook Islands	Saint Kitts and Nevis
Costa Rica	Saint Lucia
Cyprus	Saint Vincent and the Grenadines
Dominica	Samoa
Gibraltar	Seychelles
Grenada	Singapore
Guernsey	Switzerland
Ireland	U.S. Virgin Islands
Isle of Man	Uruguay
Jersey	Vanuatu
Latvia	

Appendix C. Top Tax Haven Countries by Number of Employees

Appendix C presents country-level summary statistics for the top 20 tax haven countries, ranked by average total foreign employees. We base this ranking on the average of annual total employee counts, where annual totals are calculated for each year in our final sample using data from Revelio. The table also includes mean values for GDP, GDP per capita, and the statutory tax rate for each country over the sample period.

Rank	Haven Country	Employees	GDP (in billions)	GDP Per Capita	Tax Rate
1	Singapore	28,959.286	309.441	56,988.678	17.143
2	The Netherlands	28,194.714	880.823	51,951.484	25.107
3	Ireland	27,546.643	310.834	65,281.946	12.500
4	Switzerland	16,193.429	688.212	83,679.613	21.054
5	Costa Rica	11,170.643	51.583	10,542.198	30.000
6	Panama	1,665.286	49.147	12,323.691	25.893
7	Uruguay	1,019.786	53.954	15,881.570	25.000
8	Luxembourg	961.714	65.227	115,621.506	27.800
9	Lebanon	747.643	42.667	7,548.567	15.429
10	Bermuda	518.643	6.812	105,453.313	0.000
11	Bahrain	420.500	32.070	23,964.489	0.000
12	Latvia	413.929	30.841	15,481.754	16.429
13	Cyprus	339.143	24.980	29,287.497	11.607
14	Malta	291.000	12.080	26,409.446	35.000
15	Macao	138.429	40.316	65,781.275	12.000
16	Mauritius	124.786	12.004	9,529.275	15.000
17	Liberia	119.143	2.895	632.186	27.143
18	Botswana	113.214	14.602	6,344.648	22.643
19	Aruba	107.500	2.857	27,604.979	26.714
20	Dominica	82.786	0.522	7,461.543	27.714

Appendix D. Variable Definitions

Tax Haven Variables	
Haven Substance	The number of employees located in tax haven countries scaled by the total number of employees of the firm in a given year, as identified through data from employees' online professional profiles (*Tax Haven Emp* ÷ *Emp LinkedIn*). (Data source: Revelio).
Tax Haven Emp	The number of employees located in tax haven countries for a firm in a given year. (Data source: Revelio).
Exh21 Tax Haven	An indicator variable equal to one if a firm mentions the name of a tax haven country in its Exhibit 21 in a given year, zero otherwise. (Data source: Exhibit 21 filings).
Active Haven	An indicator variable equal to one if the firm mentions the name of a tax haven country in its Exhibit 21 and in its 10-K near words that are indicative of active operations, such as "sales," "market," and "import;" zero otherwise. (Data source: Law and Mills, 2022).
Other Variables	
Cash ETR	Cash effective tax rate: cash taxes paid, divided by pretax income minus special items (TXPD ÷ (PI - SPI)), truncated at [0, 1].
GAAP ETR	Effective tax rate: total tax expense, divided by pretax income minus special items (TXT ÷ (PI - SPI)), truncated at [0, 1].
Low-Tax Substance	The number of employees located in low-tax foreign jurisdictions, excluding tax havens, scaled by the total number of employees of the firm in a given year. We classify a country as "low-tax" if its average statutory tax rate over the sample period, obtained from the Tax Foundation, is below 25 percent, which is the median average tax rate of all countries in our sample over the same period. (Data source: Revelio).
High-Tax Substance	The number of employees located in high-tax foreign jurisdictions, excluding tax havens, scaled by the total number of employees of the firm in a given year. We classify a country as "high-tax" if its average statutory tax rate over the sample period, obtained from the Tax Foundation, is above 25 percent, which is the median average tax rate of all countries in our sample over the same period. (Data source: Revelio).
Emp LinkedIn	The total number of employees of the firm based on aggregating data from employees' online professional profiles. (Data source: Revelio).
Size	Total assets (AT). (Data source: Compustat).
MB	The ratio of market value of equity to book value of equity. (Data source: Compustat).
RD Intensity	Research and development expense (XRD) divided by total assets (AT). (Data source: Compustat).
ROA	Return on assets, calculated as income before extraordinary items (IB) divided by total assets (AT).
Leverage	The ratio of long-term debt (DLTT) to total assets (AT). (Data source: Compustat).
NOL	An indicator variable equal to one if tax loss carry-forward (TLCF) is non-zero, and zero otherwise.
NOL Change	Change in tax loss carry forward (TLCF) from prior year divided by lagged total assets (AT). (Data source: Compustat).
Foreign Income	Foreign pretax income (PIFO) divided by lagged total assets (AT). (Data source: Compustat).

Appendix D. (Continued)

Equity Earnings	The absolute value of equity in earnings (loss) (ESUB) divided by lagged total assets (AT). (Data source: Compustat).
Intangible	Intangible assets (INTAN) divided by lagged total assets (AT). (Data source: Compustat).
PPE	Property, plant, and equipment (PPENT) divided by lagged total assets (AT). (Data source: Compustat).
Deficit	An indicator variable equal to one if the firm reported an accumulated deficit (i.e., negative value for its retained earnings, REUNA) at the end of the prior year, zero otherwise. (Data source: Compustat).
Deferred Income	An indicator variable equal to one if deferred revenue (DRC + DRLT) is non-zero, and zero otherwise. (Data source: Compustat).
Stock Comp	An indicator variable equal to one if stock compensation expense (STKCO) is non-zero, zero otherwise. (Data source: Compustat).
Comph Income	The absolute value of accumulated other comprehensive income (loss) (ACOMINC) divided by lagged total assets (AT). (Data source: Compustat).
Countries	The number of unique countries a firm mentions in its Exhibit 21 of a given year. (Data source: Exhibit 21 filings).
Geo Segments	The number of geographical segments as of the current fiscal year end. (Data source: Compustat Segment Database).
Emp	The total number of employees reported by the firm in a given year. (Data source: Compustat).
Avoider	An indicator variable equal to one if the firm's *GAAP ETR* is in the bottom tercile of the sample, zero otherwise.
Settlement Indicator	An indicator variable equal to one if the firm reports a non-zero settlement with tax authorities (TXTUBSETTLE) in a given year, zero otherwise. (Data source: Compustat).
LnSettlements	The log of one plus the settlements with tax authorities (TXTUBSETTLE) in a given year. (Data source: Compustat).
Settlements / AT	Settlements with tax authorities (TXTUBSETTLE) scaled by total assets (AT). (Data source: Compustat).

Continuous, non-logged variables are winsorized at 1 and 99 percent. *Ln* prefix indicates natural log.

Figure 1
Geographic Distribution of Employees

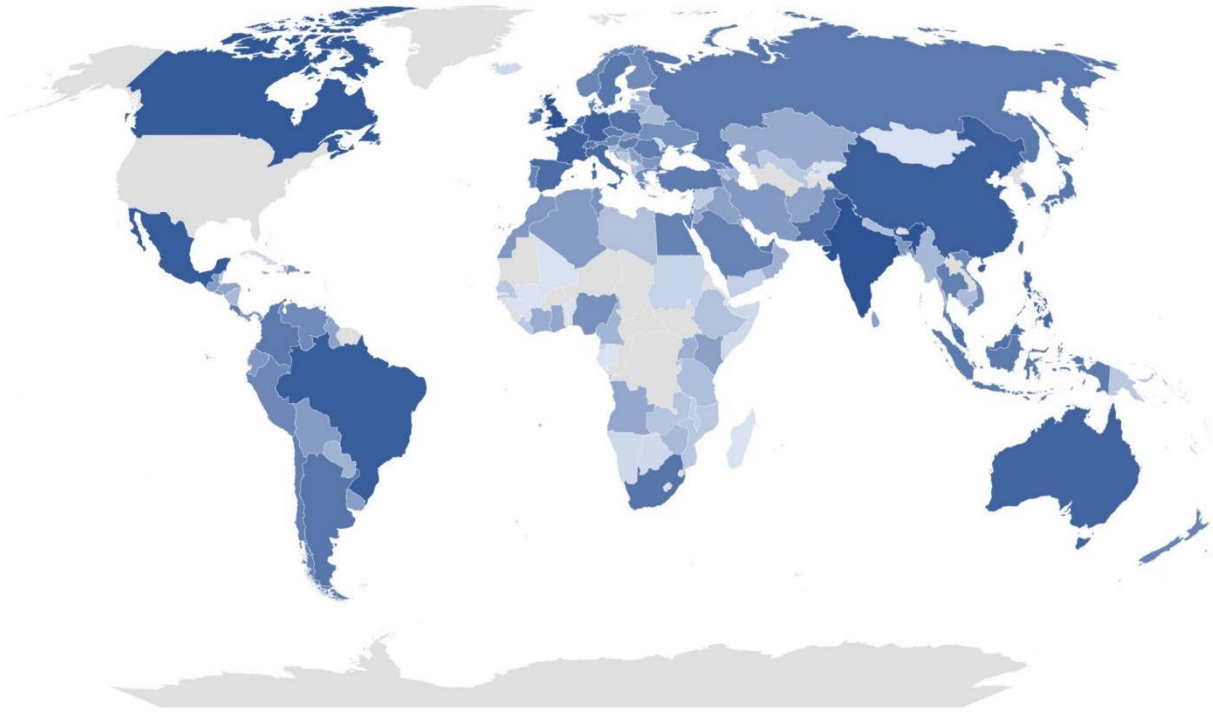

Figure 1 presents the geographic distribution of employee concentration by foreign country, shaded according to the average proportion of firms' employees in a country relative to their total workforce. To create the figure, we first calculate the proportion of employees by country for each firm-year observation using data from Revelio Labs. Next, we average these proportions across all firm-year observations in our final sample to calculate average proportions by country. The shading of the colors denotes the proportion of employees, with darker colors representing a higher concentration of employees. The countries with the highest proportions of employees are India, United Kingdom, Canada, Brazil, and Mexico.

Figure 2
Trend in Haven Substance

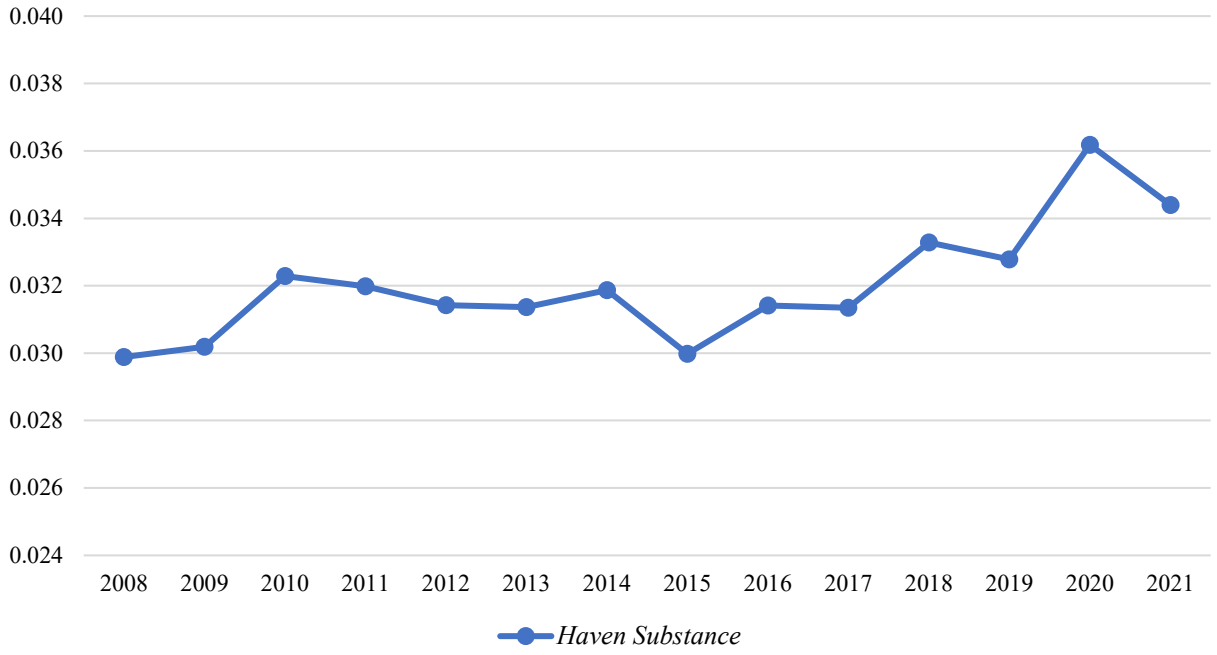

Figure 2 presents the mean value of *Haven Substance* by year from 2008 through 2021, requiring at least 5 years for each firm to ensure consistency. See Appendix D for variable definitions.

Table 1
Sample Selection

	Firm-Year Observations
U.S.-incorporated multinational public firms in Compustat with assets over $10 million and positive pre-tax income, market-to-book ratios, and sales, from 2008 to 2021.	15,067
Less: Financial and utility firms.	(1,548)
Less: Observations missing Exhibit 21 data.	(1,183)
Less: Observations missing Revelio Labs data.	(2,021)
Less: Observations with insufficient data and extreme ETRs.	(2,581)
Final Tax Avoidance Sample	7,734
Unique Firms	1,247

Table 1 presents the sample selection process. We begin with a sample of profitable, U.S.-incorporated multinational firms in Compustat with assets over $10 million, and positive market-to-book ratios and sales, for the period from 2008 to 2021. We remove financial and utility firms, and observations missing Exhibit 21 filings, merged to Compustat using CIKs. We require non-missing data in Revelio Labs to construct *Haven Substance*. Finally, we remove observations with insufficient data to estimate Equation 1, and observations with cash or GAAP ETRs below zero or above one. The final sample for the tax avoidance analysis consists of 7,734 firm-year observations. See Section 3.1 for details about sample selection.

Table 2
Summary Statistics

Panel A: Summary statistics

	N	Mean	SD	P25	P50	P75
Cash ETR	7,734	0.230	0.125	0.145	0.220	0.302
GAAP ETR	7,734	0.254	0.118	0.178	0.252	0.323
Haven Substance	7,734	0.031	0.040	0.002	0.016	0.044
Low-Tax Substance	7,734	0.122	0.101	0.042	0.103	0.175
High-Tax Substance	7,734	0.237	0.154	0.111	0.214	0.331
Exh21 Tax Haven	7,734	0.833	0.373	1.000	1.000	1.000
Active Haven	5,685	0.530	0.499	0.000	1.000	1.000
Emp LinkedIn	7,734	5,694.066	14,022.915	539.000	1,561.500	4,611.000
Emp	7,734	21,963.013	42,549.786	2,725.000	7,300.000	19,588.000
Size	7,734	10,397.609	24,244.323	930.946	2,674.496	7,739.481
MB	7,734	4.519	6.395	1.775	2.756	4.622
RD Intensity	7,734	0.031	0.045	0.000	0.012	0.047
ROA	7,734	0.072	0.048	0.038	0.062	0.095
Leverage	7,734	0.198	0.153	0.062	0.190	0.304
Deferred Income	7,734	0.516	0.500	0.000	1.000	1.000
Comph Income	7,734	0.038	0.046	0.006	0.020	0.052
Stock Comp	7,734	0.012	0.013	0.004	0.007	0.014
NOL	7,734	0.703	0.457	0.000	1.000	1.000
NOL Change	7,734	0.001	0.042	-0.003	0.000	0.003
Foreign Income	7,734	0.045	0.039	0.013	0.034	0.067
Equity Earnings	7,734	0.001	0.004	0.000	0.000	0.000
Intangible	7,734	0.298	0.237	0.101	0.259	0.447
PPE	7,734	0.214	0.179	0.090	0.162	0.275
Deficit	7,734	0.133	0.340	0.000	0.000	0.000
Countries	7,734	18.220	16.223	6.000	14.000	26.000
Geo Segments	7,734	3.844	2.567	2.000	3.000	5.000

Table 2 (Continued)

Panel B: By industry

	N	*Haven Substance*	*Tax Haven Emp*
Business equipment	2,109	0.038	318.656
Chemicals	467	0.047	325.561
Consumer durables	350	0.018	46.697
Consumer nondurables	626	0.022	124.367
Energy	199	0.021	203.266
Healthcare	646	0.055	424.226
Manufacturing	1,462	0.026	112.511
Other	983	0.020	129.260
Telephone and TV	103	0.017	114.524
Wholesale, retail	789	0.020	181.024
Total	7,734	0.031	217.088

Table 2, Panel A presents descriptive statistics for the primary dependent, explanatory, and control variables. Panel B reports the distribution by industry. See Appendix D for variable definitions.

Table 3
Pairwise Correlations

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(1)	*Haven Substance*		**-0.104**	**-0.227**	**0.343**	**0.165**	**0.154**	**0.324**	**0.078**	-0.017	**0.086**	**0.049**
(2)	*Cash ETR*	**-0.105**		**0.388**	**-0.051**	**-0.076**	**-0.068**	**-0.177**	-0.003	**-0.093**	**-0.054**	**0.063**
(3)	*GAAP ETR*	**-0.171**	**0.343**		**-0.140**	**-0.159**	**-0.070**	**-0.220**	0.002	**-0.150**	**-0.089**	**-0.039**
(4)	*Exh21 Tax Haven*	**0.227**	**-0.046**	**-0.109**		**0.222**	**0.106**	**0.153**	**0.045**	**0.032**	**0.031**	**0.143**
(5)	*LnSize*	**0.070**	**-0.079**	**-0.127**	**0.216**		**0.183**	**-0.117**	**0.025**	**0.393**	**-0.055**	**0.098**
(6)	*MB*	**0.045**	**-0.053**	**-0.055**	**0.045**	**0.105**		**0.213**	**0.449**	**0.197**	**0.144**	**0.053**
(7)	*RD Intensity*	**0.210**	**-0.155**	**-0.161**	**0.090**	**-0.141**	**0.095**		**0.150**	**-0.215**	**0.172**	**-0.039**
(8)	*ROA*	**0.053**	**-0.063**	**-0.060**	**0.049**	0.016	**0.306**	**0.162**		**-0.215**	**0.033**	-0.011
(9)	*Leverage*	**-0.024**	**-0.090**	**-0.115**	**0.024**	**0.341**	**0.286**	**-0.262**	**-0.197**		**-0.114**	**0.172**
(10)	*Deferred Income*	**0.038**	**-0.048**	**-0.079**	**0.031**	**-0.057**	**0.088**	**0.227**	**0.049**	**-0.101**		**-0.128**
(11)	*Comph Income*	0.014	-0.008	**-0.030**	**0.104**	**0.072**	**0.088**	**-0.117**	**-0.029**	**0.137**	**-0.126**	
(12)	*Stock Comp*	**0.104**	**-0.109**	**-0.053**	**0.038**	**-0.269**	**0.179**	**0.584**	**0.211**	**-0.291**	**0.287**	**-0.231**
(13)	*NOL*	**0.093**	**-0.099**	**-0.080**	**0.073**	0.002	-0.002	**0.085**	**-0.091**	**0.053**	**0.067**	-0.010
(14)	*NOL Change*	**0.030**	**0.063**	-0.021	**0.030**	**0.045**	0.012	**-0.050**	**-0.059**	**0.044**	-0.016	-0.012
(15)	*Foreign Income*	**0.310**	**-0.120**	**-0.134**	**0.216**	**0.102**	**0.112**	**0.220**	**0.438**	**-0.119**	-0.003	**0.111**
(16)	*Equity Earnings*	**-0.033**	-0.012	**0.027**	**0.033**	**0.190**	0.007	**-0.104**	**0.026**	**0.062**	**-0.055**	**0.079**
(17)	*Intangible*	-0.012	-0.003	**-0.115**	0.010	**0.145**	0.017	**-0.085**	**-0.158**	**0.329**	**0.077**	-0.018
(18)	*PPE*	**-0.084**	**-0.064**	**0.053**	**-0.046**	**0.152**	**-0.037**	**-0.263**	**-0.037**	**0.216**	**-0.171**	**0.057**
(19)	*Deficit*	-0.008	**-0.101**	-0.021	**-0.073**	**-0.157**	**0.063**	**0.155**	**-0.102**	**0.070**	**0.086**	**-0.063**
(20)	*LnCountries*	**0.243**	0.022	**-0.117**	**0.589**	**0.315**	**0.086**	**0.040**	0.012	**0.101**	-0.021	**0.220**
(21)	*LnGeo Segments*	**0.151**	-0.021	**-0.084**	**0.157**	0.003	**-0.066**	**0.184**	**0.032**	**-0.108**	**0.032**	**0.150**
(22)	*LnEmp*	-0.013	0.015	**-0.065**	**0.185**	**0.803**	**0.100**	**-0.259**	**0.037**	**0.256**	**-0.073**	**0.148**

		(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)
(1)	*Haven Substance*	**0.177**	**0.105**	0.017	**0.385**	**-0.034**	**0.030**	**-0.146**	-0.021	**0.400**	**0.188**	**0.088**
(2)	*Cash ETR*	**-0.087**	**-0.109**	**0.073**	**-0.113**	-0.016	0.008	**-0.036**	**-0.124**	0.017	**-0.040**	**0.037**
(3)	*GAAP ETR*	**-0.049**	**-0.105**	0.003	**-0.195**	**0.024**	**-0.116**	**0.041**	**-0.040**	**-0.160**	**-0.119**	**-0.079**
(4)	*Exh21 Tax Haven*	**0.048**	**0.073**	**0.040**	**0.260**	**0.033**	**0.025**	**-0.032**	**-0.073**	**0.558**	**0.152**	**0.190**
(5)	*LnSize*	**-0.356**	0.021	**0.073**	**0.101**	**0.189**	**0.159**	**0.092**	**-0.148**	**0.349**	0.008	**0.807**
(6)	*MB*	**0.300**	0.020	-0.002	**0.194**	**-0.033**	**0.124**	**-0.087**	**0.027**	**0.152**	**-0.049**	**0.151**
(7)	*RD Intensity*	**0.432**	**0.071**	**-0.052**	**0.283**	**-0.109**	-0.001	**-0.233**	**0.083**	**0.144**	**0.271**	**-0.228**
(8)	*ROA*	**0.262**	**-0.082**	**-0.043**	**0.387**	-0.010	**-0.112**	0.007	**-0.146**	0.019	0.021	**0.049**
(9)	*Leverage*	**-0.325**	**0.053**	**0.036**	**-0.105**	**0.114**	**0.335**	**0.187**	**0.039**	**0.139**	**-0.084**	**0.292**
(10)	*Deferred Income*	**0.301**	**0.067**	-0.017	-0.009	**-0.094**	**0.063**	**-0.213**	**0.086**	-0.019	**0.033**	**-0.076**
(11)	*Comph Income*	**-0.262**	-0.010	0.020	**0.169**	**0.119**	**0.056**	**0.157**	**-0.097**	**0.277**	**0.167**	**0.195**
(12)	*Stock Comp*		**0.069**	**-0.047**	**0.174**	**-0.216**	0.019	**-0.285**	**0.153**	**-0.066**	0.011	**-0.336**
(13)	*NOL*	**0.078**		0.020	**0.050**	**-0.030**	**0.058**	**-0.060**	**0.081**	**0.082**	**0.071**	-0.015
(14)	*NOL Change*	-0.017	**0.045**		0.020	0.006	**0.048**	**0.040**	**-0.100**	**0.070**	0.010	**0.077**
(15)	*Foreign Income*	**0.135**	**0.029**	**0.036**		-0.006	**-0.124**	0.005	**-0.066**	**0.319**	**0.337**	**0.095**
(16)	*Equity Earnings*	**-0.139**	**-0.028**	0.001	**0.028**		**-0.037**	**0.160**	**-0.081**	**0.064**	**0.043**	**0.164**
(17)	*Intangible*	**-0.030**	**0.050**	**0.086**	**-0.145**	**-0.088**		**-0.391**	-0.021	**0.151**	**-0.103**	**0.111**
(18)	*PPE*	**-0.235**	**-0.041**	**0.048**	-0.006	**0.181**	**-0.361**		**-0.059**	**-0.078**	**0.033**	**0.143**
(19)	*Deficit*	**0.220**	**0.081**	**-0.083**	**-0.058**	**-0.063**	-0.019	**-0.031**		**-0.122**	**-0.026**	**-0.164**
(20)	*LnCountries*	**-0.065**	**0.080**	**0.060**	**0.249**	**0.033**	**0.113**	**-0.123**	**-0.125**		**0.212**	**0.354**
(21)	*LnGeo Segments*	-0.002	**0.066**	**0.029**	**0.270**	**0.053**	**-0.107**	-0.010	**-0.026**	**0.213**		**-0.023**
(22)	*LnEmp*	**-0.312**	-0.020	**0.051**	**0.082**	**0.140**	**0.084**	**0.104**	**-0.175**	**0.332**	**-0.028**	

Table 3 presents Pearson correlations below the diagonal and Spearman correlations above the diagonal. Correlation coefficients that are significantly different from 0 at the 5 percent level or better are in boldface. See Appendix D for variable definitions.

Table 4
Economic Substance in High-Tax, Low-Tax, and Haven Jurisdictions and Tax Avoidance

	Cash ETR			*GAAP ETR*		
	(1)	(2)	(3)	(4)	(5)	(6)
Haven Substance	-0.178***		-0.177***	-0.262***		-0.263***
	(-3.00)		(-2.99)	(-5.39)		(-5.50)
Low-Tax Substance		0.029	0.028		-0.061***	-0.063***
		(1.17)	(1.14)		(-2.97)	(-3.06)
High-Tax Substance		0.050***	0.050***		-0.023	-0.023
		(2.99)	(3.03)		(-1.59)	(-1.61)
LnSize	-0.018***	-0.018***	-0.017***	-0.007***	-0.008***	-0.007***
	(-6.28)	(-6.06)	(-5.81)	(-2.62)	(-3.30)	(-2.88)
MB	0.000*	0.000*	0.001*	0.001***	0.001***	0.001***
	(1.76)	(1.75)	(1.78)	(2.78)	(2.74)	(2.80)
RD Intensity	-0.271***	-0.286***	-0.274***	-0.353***	-0.363***	-0.345***
	(-4.15)	(-4.35)	(-4.20)	(-6.57)	(-6.85)	(-6.46)
ROA	-0.121**	-0.081	-0.094*	-0.110**	-0.123**	-0.142***
	(-2.30)	(-1.51)	(-1.74)	(-2.18)	(-2.35)	(-2.73)
Leverage	-0.050***	-0.053***	-0.052***	-0.048***	-0.050***	-0.049***
	(-2.80)	(-2.91)	(-2.89)	(-3.13)	(-3.25)	(-3.21)
Stock Comp	-0.409**	-0.396**	-0.399**	0.176	0.176	0.171
	(-2.03)	(-1.97)	(-1.98)	(0.99)	(1.03)	(0.98)
Deferred Income	-0.002	-0.002	-0.002	-0.007*	-0.006*	-0.006*
	(-0.37)	(-0.48)	(-0.50)	(-1.87)	(-1.73)	(-1.78)
Comph Income	-0.065	-0.081*	-0.088*	0.001	0.027	0.016
	(-1.36)	(-1.72)	(-1.86)	(0.04)	(0.68)	(0.42)
NOL	-0.016***	-0.018***	-0.017***	-0.007*	-0.007**	-0.006*
	(-3.81)	(-4.15)	(-3.99)	(-1.93)	(-2.14)	(-1.84)
NOL Change	0.179***	0.181***	0.184***	-0.049	-0.053	-0.048
	(4.18)	(4.22)	(4.26)	(-1.19)	(-1.28)	(-1.17)
Foreign Income	-0.250***	-0.376***	-0.331***	-0.171***	-0.160**	-0.094
	(-3.51)	(-5.15)	(-4.36)	(-2.76)	(-2.49)	(-1.41)
Equity Earnings	-0.705	-0.659	-0.707	0.189	0.290	0.218
	(-1.29)	(-1.19)	(-1.30)	(0.43)	(0.63)	(0.50)
Intangible	-0.009	-0.005	-0.007	-0.045***	-0.046***	-0.048***
	(-0.86)	(-0.50)	(-0.63)	(-4.63)	(-4.64)	(-4.88)
PPE	-0.080***	-0.075***	-0.076***	-0.010	-0.013	-0.014
	(-4.42)	(-4.11)	(-4.22)	(-0.72)	(-0.91)	(-1.01)
Deficit	-0.026***	-0.027***	-0.027***	-0.011**	-0.009*	-0.010*
	(-3.83)	(-3.98)	(-4.01)	(-2.00)	(-1.76)	(-1.80)
LnCountries	0.011***	0.007***	0.009***	-0.001	-0.001	0.001
	(4.30)	(2.78)	(3.22)	(-0.44)	(-0.41)	(0.39)
LnGeo Segments	0.005	0.003	0.004	-0.004	-0.004	-0.003
	(1.11)	(0.68)	(0.79)	(-1.19)	(-0.97)	(-0.76)
LnEmp	0.013***	0.012***	0.011***	-0.001	0.002	0.000
	(4.00)	(3.75)	(3.48)	(-0.22)	(0.58)	(0.14)
Observations	7,734	7,734	7,734	7,734	7,734	7,734
Adjusted R^2	0.135	0.135	0.138	0.197	0.193	0.199
Industry FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES

Table 4 presents the results of estimating Equation 1 with *Cash ETR* (*GAAP ETR*) as the dependent variable in Columns 1-3 (4-6). Continuous, non-logged variables are winsorized at 1% and 99%. ***, **, and * denote statistical significance in two-tailed tests at the 1%, 5%, and 10% levels. Coefficient t-statistics are in parentheses, calculated with standard errors clustered by firm. See Appendix D for variable definitions.

Table 5
Incremental Effect of Haven Substance on Tax Avoidance

Panel A: Controlling for the mention of a tax haven in Exhibit 21

	Cash ETR		*GAAP ETR*	
	(1)	(2)	(3)	(4)
Haven Substance		-0.170***		-0.256***
		(-2.86)		(-5.27)
Exh21 Tax Haven	-0.013**	-0.011*	-0.010**	-0.008
	(-1.96)	(-1.71)	(-2.07)	(-1.54)
Observations	7,734	7,734	7,734	7,734
Adjusted R^2	0.134	0.136	0.191	0.197
Controls Included	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
F-Test p-values				
Haven Substance = Exh21 Tax Haven		0.008***		0.000***

Panel B: Controlling for the presence of an active haven following Law and Mills (2022)

	Cash ETR		*GAAP ETR*	
	(1)	(2)	(3)	(4)
Haven Substance		-0.228***		-0.252***
		(-3.07)		(-4.18)
Active Haven	-0.012**	-0.010*	-0.012***	-0.009**
	(-2.42)	(-1.88)	(-2.84)	(-2.13)
Exh21 Tax Haven	-0.013	-0.011	-0.009	-0.008
	(-1.60)	(-1.45)	(-1.45)	(-1.23)
Observations	5,685	5,685	5,685	5,685
Adjusted R^2	0.140	0.143	0.148	0.153
Controls Included	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
F-Test p-values				
Haven Substance = Active Haven		0.004***		0.000***
Haven Substance = Exh21 Tax Haven		0.004***		0.000***
Active Haven = Exh21 Tax Haven	0.980	0.871	0.757	0.898

Table 5 presents the results of assessing the incremental explanatory power of *Haven Substance* in explaining firms' ETRs, relative to tax haven indicators from prior research. In both panels, the dependent variable is *Cash ETR* in Columns 1-2 and *GAAP ETR* in Columns 3-4. In Panel A, we estimate Equation 1 with *Exh21 Tax Haven*. Following Dyreng and Lindsay (2009), *Exh21 Tax Haven* is an indicator equal to one if the firm discloses the presence of a significant subsidiary in a tax haven country in its Exhibit 21 in a given year, zero otherwise. Columns 1 and 3 present baseline results, excluding proxies for economic substance, and Columns 2 and 4 present results when *Haven Substance* is also included in the model. In Panel B, we include Law and Mills' (2022) *Active Haven* indicator, equal to one if a firm has a significant tax haven subsidiary that is more than a shell company, and zero otherwise. *Active Haven* is only available until 2016, reducing the sample size. Similar to Panel A, we report baseline regression results in Columns 1 and 3 and results after including *Haven Substance* in Columns 2 and 4. F-tests comparing coefficients are also presented in the table. Continuous, non-logged variables are winsorized at 1% and 99%. ***, **, and * denote statistical significance in two-tailed tests at the 1%, 5%, and 10% levels. Coefficient t-statistics are in parentheses, calculated with standard errors clustered by firm. See Appendix D for variable definitions.

Table 6
Income Shifting

Panel A: Full Sample

		DV = ΔPIDOM and ΔPIFO		
		(1)	(2)	(3)
Outbound Transfers				
θ0		-0.181	-0.240	-0.213
		(-1.04)	(-1.34)	(-1.27)
θ1	*Haven Substance*	2.122***		1.889***
		(3.01)		(2.65)
θ2	*Exh21 Tax Haven*		0.123**	0.107*
			(2.09)	(1.80)
Inbound Transfers				
γ0		0.579**	0.552**	0.548**
		(2.37)	(2.04)	(2.11)
γ1	*Haven Substance*	0.255		0.184
		(0.30)		(0.24)
γ2	*Exh21 Tax Haven*		0.043	0.051
			(0.28)	(0.35)
Observations		7,047	7,047	7,047
Adjusted R^2 - ΔPIFO Eqn.		0.257	0.258	0.258
Adjusted R^2 - ΔPIDOM Eqn.		0.185	0.183	0.185
Controls Included		YES	YES	YES
F-Test p-values				
θ1 = θ2				0.014**

Table 6 (Continued)

Panel B: Controlling for the presence of an active haven following Law and Mills (2022)

		DV = ΔPIDOM and ΔPIFO		
		(1)	(2)	(3)
Outbound Transfers				
θ0		-0.110	-0.068	-0.087
		(-0.42)	(-0.28)	(-0.35)
θ1	*Haven Substance*	2.500***		1.718**
		(3.37)		(2.14)
θ2	*Exh21 Tax Haven*		0.041	0.033
			(0.43)	(0.37)
θ3	*Active Haven*		0.175**	0.145
			(2.05)	(1.64)
Inbound Transfers				
γ0		0.684***	0.656***	0.589***
		(2.85)	(2.73)	(2.65)
γ1	*Haven Substance*	0.995		0.913
		(1.43)		(1.17)
γ2	*Exh21 Tax Haven*		0.199	0.190
			(1.30)	(1.20)
γ3	*Active Haven*		-0.077	-0.096
			(-1.04)	(-1.49)
Observations		5,091	5,091	5,091
Adjusted R^2 - ΔPIFO Eqn.		0.270	0.275	0.274
Adjusted R^2 - ΔPIDOM Eqn.		0.165	0.162	0.166
Controls Included		YES	YES	YES
F-Test p-values				
θ1 = θ2				0.043**
θ1 = θ3				0.055*
θ2 = θ3			0.378	0.448

Table 6 presents the results of validating *Haven Substance* in the context of tax motivated outbound income shifting. Following Dyreng and Markle (2016), we estimate Equations 2a and 2b using a seemingly unrelated regression analysis. The dependent variables, ΔPIDOM and ΔPIFO, represent the change in domestic pretax income and the change in foreign pretax income, respectively. Panel A presents full sample results after requiring non-missing controls, and Panel B presents the results after including *Active Haven* in the model. *Active Haven* is only available until 2016, resulting in a lower sample size. F-tests comparing outbound income shifting coefficients are also presented in the table. Continuous, non-logged variables are winsorized at 1% and 99%. ***, **, and * denote statistical significance in two-tailed tests at the 1%, 5%, and 10% levels. Coefficient t-statistics are reported in parentheses. Consistent with Dyreng and Markle (2016), standard errors are clustered by firm and year (results are similar when clustering by firm). See Appendix D for variable definitions.

Table 7
Settlements with Tax Authorities

	Settlement Indicator	*LnSettlements*	*Settlements / AT*
	(1)	(2)	(3)
Avoider x Haven Substance	-0.900**	-2.167***	-0.273***
	(-2.40)	(-3.01)	(-2.81)
Haven Substance	-1.437	-12.909**	-0.604
	(-0.76)	(-2.52)	(-1.27)
Avoider	0.049**	0.163***	0.029***
	(2.55)	(3.98)	(4.80)
Observations	6,762	6,762	6,762
Adjusted R^2	0.199	0.425	0.061
Controls	YES	YES	YES
Controls x *Haven Substance*	YES	YES	YES
Industry FE	YES	YES	YES
Year FE	YES	YES	YES

Table 7 presents the results of estimating Equation 3. The dependent variable is one of three alternative proxies for tax settlements: *Settlement Indicator* in Column 1, *LnSettlements* in Column 2, and *Settlements / AT* in Column 3. All control variables from Equation 1, along with their interactions with *Haven Substance*, are included in the model. Continuous, non-logged variables are winsorized at 1% and 99%. ***, **, and * denote statistical significance in two-tailed tests at the 1%, 5%, and 10% levels. Coefficient t-statistics are in parentheses, calculated with standard errors clustered by firm and year. See Appendix D for variable definitions.

Table 8
Disaggregation of Haven Substance by Core and Supportive Occupations

	Cash ETR		*GAAP ETR*	
	(1)	(2)	(3)	(4)
Core Haven Substance	-0.373***	-0.366***	-0.356***	-0.353***
	(-3.66)	(-3.59)	(-3.85)	(-3.79)
Supportive Haven Substance	0.073	0.080	-0.238*	-0.234*
	(0.46)	(0.50)	(-1.69)	(-1.67)
Exh21 Tax Haven		-0.008		-0.004
		(-1.12)		(-0.71)
Observations	5,649	5,649	5,649	5,649
Adjusted R^2	0.136	0.136	0.207	0.207
Controls Included	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Table 8 presents the results of estimating Equation 1 after disaggregating *Haven Substance* into *Core Haven Substance* and *Supportive Haven Substance*. *Core Haven Substance* is calculated as the number of employees in tax havens engaged in scientific, operational, or engineering roles, scaled by the total number of employees of the firm, using data from Revelio. *Supportive Haven Substance* is calculated as the number of employees in tax havens engaged in marketing, sales, administrative, and financial roles, scaled by the total number of employees of the firm. The dependent variable is *Cash ETR* in Columns 1 and 2 and *GAAP ETR* in Columns 3 and 4. To ensure sufficient variation in job categories, we remove firm-year observations with haven employees that do not have at least 20 employees located in haven countries in a given year. All continuous, non-logged variables are winsorized at 1% and 99%. Coefficient t-statistics are in parentheses, calculated with standard errors clustered by firm. ***, **, and * denote statistical significance in two-tailed tests at the 1%, 5%, and 10% levels, respectively. See Appendix D for variable definitions.