How Do Features of Tax Enforcement Impact Effective Tax Rates Globally?

Lisa De Simone
Associate Professor
University of Texas at Austin
Email: Lisa.DeSimone@mccombs.utexas.edu

Bridget Stomberg
Professor
Kelley School of Business
Email: bstomber@indiana.edu

Brian Williams
Professor
Mays Business School
Email: bwilliams@mays.tamu.edu

January 2024

Keywords: Tax avoidance; Tax enforcement; Tax administration; Tax policy

Abstract: Tax enforcement is an important tool that governments use to ensure compliance, minimize tax avoidance, and increase tax collections. Prior work finds that corporate tax avoidance is decreasing in tax enforcement spending. However, countries face significant political and budgetary obstacles in dedicating funds to tax enforcement, and spending is only one element of a tax enforcement system. We use a new "glass-box" machine learning approach to understand which features of tax enforcement are most important for explaining country-level GAAP effective tax rates. Using OECD data for 136 country-years, we find that tax enforcement personnel is the most important feature, followed by technological advancements such as artificial intelligence. In particular, having a greater percentage of staff within the large taxpayer office of the tax authority is an important driver of corporate tax avoidance, more so than staff tenure or turnover. Artificial intelligence and machine learning are especially important for firms likely subject to the pending global minimum tax under Pillar 2 as well as for intangible-intensive firms. Our results inform tax authorities, corporate taxpayers, and researchers about which elements of global tax enforcement systems are most associated with corporate tax avoidance.

We thank workshop participants at the University of Notre Dame and reviewers for the 2024 Midyear Meeting of the American Taxation Association for helpful comments and suggestions. We thank our institutions for financial support. Bridget Stomberg gratefully acknowledges the Weimar Fellowship, and Brian Williams gratefully acknowledges the PWC Professorship.

1. Introduction

Tax enforcement is a key lever that governments exploit to mitigate corporate tax avoidance and increase tax collections. Prior work finds that corporate tax avoidance responds negatively to increases in enforcement spending (e.g., De Simone, Stomberg, and Williams 2023; Nessa, Schwab, Stomberg and Towery 2020). However, countries face political and budgetary obstacles in earmarking funds for enhanced enforcement. Moreover, monetary resources are but one element of a tax enforcement system, and tax authorities face challenges in choosing how to allocate funds across various enforcement tools. In the face of resource constraints and allocation decisions, it is important for tax authorities to understand how best to spend enforcement funds. Learning which elements of a tax enforcement system are most useful in explaining tax avoidance is also of interest to multinational corporations as they weigh the costs and benefits of doing business and avoiding taxes in different jurisdictions. In this paper, we provide new and novel evidence on which features of country-year tax enforcement explain the most variation in country-year effective tax rates (ETRs).

We aim to evaluate the relative importance of different features of tax enforcement systems on country-year tax avoidance. To do so, we utilize an explainable boosting machine (EBM) learning algorithm originally developed by researchers at Microsoft Research (Nori et al. 2019). This algorithm is an example of a "glass-box" machine learning prediction model and is a relatively recent development in explainable artificial intelligence (AI). These models estimate predictions with high interpretability and allow the user to understand what leads the AI model to generate the prediction. Not only are these models highly interpretable, but they also offer high dimensionality – that is for each observation we can quantify the role of each individual feature in the prediction for that observation. We utilize EBM to estimate the relative contribution of various features of country-year tax enforcement and their pairwise interactions on country-year corporate tax avoidance.

We examine various features of countries' tax enforcement systems using publicly available annual data from the Organization for Economic Cooperation and Development (OECD). We investigate three groups of features: (1) those related to financial resources, (2) those related to human capital at the tax authority, and (3) those related to technological innovations at the tax authority. First, we examine financial resources using the country-year tax enforcement budget as De Simone et al. (2023) and Nessa et al. (2020) demonstrate that tax enforcement spending is associated with corporate tax avoidance. Second, we explore features related to human capital. These features capture the labor that tax authorities use to facilitate enforcement of the tax laws in their country. We use country-year measures of the percentage of all tax authority staff allocated to the audit function. We also use the percentage of all tax authority staff allocated to the large taxpayer division, which may be a more direct measure of staff influence over the large corporations in our sample. Additionally, we measure the average annual tenure and turnover of all staff at the tax authority. We include these variables as Seidman, Sinha and Stomberg (2022) report that U.S. tax executives often characterize low-quality tax agents and frequent agent turnover as barriers to effective enforcement. Third, we include features that capture the tax authority's use of AI and machine learning as well as robotic process automation to understand the role that new technology can play in helping tax enforcement (Joshi and Weng 2022).

In addition to features directly related to tax enforcement, we examine tax policy features and economic features within the country-year. Doing so allows us to explore the effects of tax enforcement features in the presence of various tax policy or economic features. We include the statutory tax rate because it captures general incentives for tax avoidance within the country — taxpayers facing higher statutory tax rates have greater incentives for tax avoidance, all else equal. We also include measures of GDP and corruption, which can influence incentives and opportunities for tax avoidance.

We follow Shevlin et al. (2019) and measure corporate tax avoidance using the asset-weighted country-year average GAAP ETR. The measure allows for the tax avoidance of the largest firms in a country-year to have a bigger impact than that of smaller firms. Our sample is 136 country-year observations from 39 unique countries from 2016 through 2019. We begin in 2016 and end in 2019 as this is the period with detailed data on tax enforcement available from the OECD. We use a random subset of 25 percent of the sample (34 country-year observations) to train the machine learning model. This step results in a sample of 102 country-year observations for use in further analysis. Given our small sample, EBM is an especially powerful tool because it avoids overfitting in small datasets.

We begin by comparing the ability of the machine learning model to predict country-year asset-weighted GAAP ETRs using the tax enforcement, policy, and economic features described above to that of an OLS prediction model of asset-weighted GAAP ETRs as a function of the same features (Guenther, Peterson, Searcy, and Williams 2023). The EBM model's prediction is superior to that of the OLS model; the predicted value of country-year asset-weighted GAAP ETRs from the machine learning model explains 41 percent of the variation in the actual values whereas the predicted value from the OLS model explains only 21.4 percent. We conclude that our EBM model outperforms OLS in its predictive abilities.

We next address our research question, which asks which tax enforcement features are most important in understanding country-year corporate tax avoidance. We examine which features contribute the most to the machine learning model's prediction of country-year tax avoidance. The machine learning model determines tax enforcement features to be highly important in predicting country-year asset-weighted GAAP ETRs, as these features in sum account for 51.3 percent of the machine learning model's prediction. Within tax enforcement features, we estimate that features related to tax authority human capital are most important (27.1 percent), followed by features related to tax authority technology (14.3 percent) and finally the tax authority budget (9.8 percent). We also estimate that a significant amount of country-year corporate tax avoidance is explained by features outside of the tax enforcement system. The statutory tax rate accounts for 16.8 percent of the contribution whereas GDP and corruption contribute the remaining 31.9 percent.

We also examine the contributions of these tax enforcement features separately based on whether firms are likely targets of recent enhanced enforcement efforts, particularly around cross-border income shifting. We do so because curbing tax-motivated income shifting by multinationals has been a keen focus of recent global tax enforcement efforts. We first identify firms likely to be subject to the new global minimum tax under Pillar 2 of the OECD/G20's Base Erosion and Profit Shifting initiative as those with sales above €750 million. We then compute country-year asset-weighted GAAP ETRs for the subsample of firms subject to Pillar 2, as well as the subsample of firms that are not subject to Pillar 2. We conduct another analysis that splits firms with relatively high levels of intangible assets from those with low levels of intangible assets. We do this because prior literature shows the intangible assets can facilitate income shifting (e.g., De Simone, Mills and Stomberg 2019; Klassen and Laplante 2012). We then estimate separate machine learning models for each of these four subsamples. These analyses reveal that AI and machine learning are significantly more important features of tax enforcement for the subsample of firms subject to Pillar 2 than for the subsample not subject to Pillar 2, as well as for high intangible versus low intangible firms. Additionally, human capital features such as staff tenure and staff turnover have a greater contribution to country-year asset-weighted GAAP ETRs for high intangible firms relative to low intangible firms. These results are important for tax authorities worldwide that are striving to effectively target base erosion and profit shifting among these firms.

Lastly, we utilize the dimensionality of the interpretable machine learning values (i.e., feature-by-observation level estimates) to explore how tax enforcement features influence other tax enforcement features. Results reveal a few key insights. First, when tax enforcement expenditures increase, AI and machine learning have a larger effect on country-level tax avoidance. Second, having more staff working in the large taxpayer division also allows AI and machine learning to have a greater effect on corporate tax avoidance. These results reveal complementarities between financial resources, human capital, and automation. Third, having more staff working in the audit function or in the large taxpayer division is associated with tax expenditures having *less* of an effect on corporate tax avoidance. Thus, human tax agents compensate for a lack of tax authority resources. Finally, when the statutory tax rate is lower, the tax enforcement budget and staff turnover have a greater impact on corporate tax avoidance. These findings underscore the relation between incentives for tax enforcement (as reflected in the statutory tax rate) impacting the role of tax authorities' financial resources and human capital on corporate tax avoidance.

Our study informs policymakers about the effectiveness of various enforcement efforts to mitigate corporate tax avoidance. Our findings suggest that although technological innovations using AI and machine learning as well as robotic process automation are important drivers of corporate tax avoidance, the role of human capital currently dominates. Thus, investments to improve technology should not be made at the expense of investments in human capital. These results should provide comfort to those in the tax profession who fear that AI is cannibalizing human jobs. Our findings also reveal that the relative importance of tax enforcement features varies with firm-level characteristics. These findings are particularly important for tax authorities looking to deploy targeted enforcement mechanisms to achieve higher returns on audits.

Our study also adds to the literature examining the relation between tax enforcement and corporate tax avoidance (Atwood et al. 2012; Beuselinck, Deloof, and Vanstraelen 2015; De Simone et al. 2023; Gupta and Lynch 2016, Hoopes et al. 2012; Hoopes et al. 2018; Nessa et al. 2020). Using Shapley values in a sample of U.S. publicly traded firms, Belnap, Kroeger, and Thornock (2023) find that receipt of an SEC comment letter and firm size (their proxies for enforcement) explain little variation in ETRs. They conclude that tax enforcement is a relatively less important determinant of corporate tax avoidance for U.S. firms. These results stand in apparent contrast to those in other studies that find significant effects of tax enforcement spending and other elements of tax enforcement on the tax avoidance of large multinational corporations. We extend the work in Belnap et al. (2023) by analyzing tax enforcement in a global setting, utilizing machine learning to identify nonlinear relationships between tax enforcement and tax avoidance, and measuring multiple elements of tax enforcement — enforcement spending, staffing metrics, use of AI and automated processes in enforcement efforts, etc. — to further our understanding of the relative importance of these components of tax enforcement in explaining global effective tax rates.

Of course, our EBM estimation comes with the caveat that our machine learning model cannot explain 59 percent of the variation in country-level tax enforcement. However, the goal of our paper is not to identify all the determinants of country-year tax avoidance, but rather to understand how and when tax enforcement can be effective in mitigating country-year tax avoidance. Further, the variation left unexplained by our EBM estimation is smaller than the variation left unexplained by an OLS prediction model.

2. Related literature

All else equal, governments can increase tax revenues by raising the tax rate or increasing enforcement (Keen and Slemrod 2017). Enforcement is a key component of any tax system because it outlines the rules and procedures necessary to ensure compliance (Slemrod and Gillitzer 2014). The importance of corporate tax enforcement has increased in the last ten to fifteen years as regulators, tax authorities, and the public have become more aware and more critical of the tax avoidance activities of large corporations, particularly of large multinational corporations. Tax administrators worldwide have increased disclosure requirements such as Schedule UTP and country-by-country reporting to combat aggressive corporate tax avoidance. Other initiatives include expanding the number of large corporate audits and implementing advanced analytic techniques such as AI in audit selection processes. For example, the IRS used a portion of its special allocation from the Inflation Reduction Act to audit sixty of the largest corporate taxpayers that were selected for audit using a combination of AI and subject matter experts. Chi et al. (2022) provide evidence that foreign tax authorities download the annual reports of U.S. publicly traded multinationals to aid in their enforcement efforts.

Several studies examine the relation between tax enforcement and tax avoidance, particularly among U.S. corporations. Hoopes et al. (2012) find that public U.S. corporate tax avoidance is decreasing with their expected probability of an IRS audit. Nessa et al. (2020) use confidential IRS enforcement data on enforcement spending and audit hours to find a positive association between greater enforcement and aggregate collections from large public corporate taxpayers. At the U.S. state level, Gupta and Lynch (2016) find further evidence that increased enforcement spending is associated with decreased corporate tax avoidance.

Other studies examine the role of tax enforcement globally. For example, Atwood et al. (2012) find less tax avoidance in countries where constituents perceive tax enforcement to be stronger. Beuselinck et al. (2015) examine the differential effects of corporate tax enforcement on public versus private firms. Relatedly, Hoopes et al. (2018) provide evidence that private firms increased tax payments in response to public disclosure of their tax information in Australia whereas public firms decreased tax payments. De Simone et al. (2023) examine the effect of tax enforcement spending on corporate tax avoidance, finding differential effects of home-country tax enforcement on domestic and multinational firms. We extend this literature by providing new and novel evidence on which features of tax enforcement systems worldwide contribute the most to corporate-level tax avoidance within countries.

3. Research design and sample

3.1. Research Design

We aim to evaluate the relative importance of different features of country-level tax policy, tax enforcement, and the economic environment on the asset-weighted average effective tax rate reported by publicly traded firms in the country. To do so, our research design relies on relatively recent developments in explainable AI. Specifically, we make use of "glass-box" machine learning prediction models. These models estimate predictions with high interpretability, allowing the user to understand what leads the AI model to produce the prediction. In our setting, this interpretability allows us to understand the relative importance of and interactions between various components of country-level tax enforcement and other country-level characteristics on country-level tax planning measured with ETRs.

In our paper, we utilize an explainable boosting machine (EBM) learning algorithm originally developed by researchers at Microsoft Research (Nori et al. 2019). As noted by Nori et al. (2019), "EBM is a glass-box model, designed to have accuracy comparable to state-of-the-art machine learning methods like Random Forest and Boosted Trees, while being highly intelligible and explainable" (p. 3). EBM models estimate the nonlinear effects of each feature, or independent variable, in the model as well as pairwise interaction terms between all features in the model. To our knowledge we are the first paper in accounting to utilize the glass-box EBM approach, although we note that we are far from the first paper to utilize techniques to understand the importance of features to the predictions of a machine learning algorithm in a general sense, whether that be the importance of textual or quantitative features (see, for example, Chen, Cho, Dou, and Lev 2022; Donovan, Jennings, Koharki, and Lee 2021; Erel, Stern, Tan and Weisbach 2021; Guenther, Peterson, Searcy and Williams 2023; and Jennings, Lee, and Towery 2021).

EBM's can be especially powerful to avoid overfitting in small datasets due to their additive nature. This feature of EBM is particularly important for our setting because our full sample consists of only 136 country-year observations. As noted in Nori et al. (2019), "Because EBM is an additive model, each feature contributes to predictions in a modular way" (p. 3). EBM learns the association between each feature and the outcome or target variable through machine learning techniques. It restricts the boosting algorithm to train the model on one feature at a time in a round-robin fashion (to mitigate effects of collinearity) while using a low learning rate. Thus, the order in which the contribution of each feature is estimated is irrelevant.

Specifically, EBM is a form of a generalized additive model that estimates function *f* for feature X_j when predicting the value of a target or outcome *Y* such that:

$$(1) \qquad (E[Y]) = B_0 + \sum f_j\left(X_j\right)$$

EBM is also expanded to estimate pairwise interactions of the form:

$$(2) \qquad (E[Y]) = B_0 + \sum f_j\left(X_j\right) + \sum f_{ij}\left(X_i, X_j\right)$$

We estimate our machine learning models using a combination of code provided by InterpretML as well as our own proprietary code.

3.2. An Example of the Glass-Box Machine Learning Model

The output of the interpretable machine learning model is multi-fold. First, the model uses training data to understand the nonlinear relationships between the features, pairwise interactions of the features, and the targets. Second, the model then uses the information it learns about these relationships in the training sample to predict the target in the test sample, given the available features and their interaction.

In our study, the target is country-level tax avoidance measured using asset-weighted country-year average GAAP ETR following Shevlin et al. (2019). The features encompass country-year tax enforcement variables. We also include country-level economic and other tax policy variables to understand both their direct and interactive effect on country-level tax planning. Our approach uses the actual asset-weighted average GAAP ETR and feature values (tax enforcement variables, etc.) in the training data to map the relationship between these features and the asset-weighted average GAAP ETR. We then utilize these mappings to generate a *predicted value* of the asset-weighted average GAAP ETR (target) for each country-year observation in our test sample. These test sample results, which are akin to out of sample tests, form the basis for our analysis. As an additive glass-box model, EBM also provides the contribution of each feature to its prediction of the target value for each observation in the sample (i.e., each country-year). These interpretable machine learning values for each feature, pairwise interaction, plus a constant sum to the final predicted value for that observation.

For example, assume we estimate an EBM model that predicts Y given X_1 and X_2. The model would generate $\hat{y}$, the predicted value of Y, as well as interpretable machine learning values for X_1 , X_2, and the interaction of X_1,X_2. Using example numbers, if $\hat{y}$ were 0.20, the sum of machine learning features and interactions (plus a constant) add up to 0.20. It could be the case that if the constant is zero, the interpretable machine learning values for X_1 , X_2, and the interaction X_1,X_2 are 0.10, 0.03, and 0.07, respectively. To interpret the relative contribution of each feature in a pairwise interaction, we allocate half of the machine learning value for the interaction to each component feature. To continue our example, we attribute half of the interpretable machine learning value of X_1,X_2 to X_1 and X_2, so that the final interpretable machine learning value for X_1 is $(0.1+\frac{.07}{2}) = 0.135$ and X_2 is $(0.03+\frac{.07}{2}) = 0.065$.

In this example, each feature increases the predicted value of the target, but a feature can also decrease the predicted value of the target. In this case, the country-year feature leads the algorithm to predict a lower value of the target Y. When calculating the relative contributions of each feature to the prediction, we take the absolute value of the interpretable machine learning value and scale it by the sum of the absolute value of all interpretable machine learning values. Continuing our above example, to understand the contribution of X_1 to ŷ, we calculate absolute value (0.135) / [absolute value (0.135) + absolute value (0.065)] = 67.5%.

Finally, we note that the interpretable machine learning values are created for each feature-observation combination. That is, we can determine for each observation the unique contribution of each feature to that particular observation's predicted value. This allows, for example, the algorithm to take all data and functional forms into context that could result in a feature having a different effect on the algorithm's predicted value across countries, or even across years within the same country. Thus in a dataset with *N* observations and *Z* features, our final calculation results in *N***Z* interpretable machine learning values.[1] We then either aggregate these across types of observations for the respective sample (e.g. Figures 1 and 2; Table 4), or utilize them individually at the observation-year level for the respective sample (e.g. Tables 5-7), depending upon the analysis.

3.3. Enforcement and Country Characteristics

We utilize data from the OECD's tax administration database to select the features of tax enforcement that we examine. The OECD provides several variables across multiple datasets that offer details on tax enforcement annually across countries. We investigate three groups of features: (1) those related to financial resources, (2) those related to human capital at the tax authority, and (3) those related to technological innovations at the tax authority.

First, we measure financial resources. We use two measures. The first measure captures overall spending by calculating the total enforcement budget as a percentage of GDP *(Tax Expenditure/GDP).* The second measure is the inflation-adjusted percentage change in enforcement budget *(% Change Enforce)* from De Simone et al. (2023), which captures the year-over-year trend in spending on tax enforcement.

Second, we examine factors related to human capital at the tax authority. One set of measures captures how tax authority staff are allocated within the organization. *% Staff – Audit Function* is the percentage of all tax authority staff working in the audit function and *% Staff –*

[1] As noted above EBM utilizes the role of each feature as well as each pairwise interaction between features. Thus in calculation of the *Z* interpretable machine learning values noted above, *Z* represents the total contribution of each feature, which as mentioned previously we calculate as the contribution of each individual feature *Z*, plus 0.5* the sum the contribution of each of the interactions of *Z* with other features.

Large Taxpayer is the percentage of all tax authority staff working in the division that monitors large taxpayers. These variables allow us to understand the extent to which having tax authority staff engaged in auditing taxpayers or interacting directly with large taxpayers influences the effectiveness of enforcement. The inclusion of these variables is important as prior research documents that tax authority audits are associated with decreased tax avoidance for US firms (Hoopes et. al 2012), while the role of large taxpayer programs such as the Coordinated Industry Case program on corporate tax avoidance is somewhat mixed (Ayers, Seidman, and Towery 2019). We also include *Staff Tenure* as a proxy for experience and subject matter expertise and a measure of staff turnover (*% Staff Departing*). Interviews with tax executives in the U.S. reveal that both characteristics of tax agents are salient elements of the audit process for corporate taxpayers (Seidman, Sinha and Stomberg 2022).

Third, to understand the role that new technology can play in tax enforcement, we examine the use of *AI and Machine Learning* as well as the use of *Robotic Process Automation.* Although these variables specifically measure the use of automation and AI or machine learning, they are likely correlated with the use of other advanced analytics by the tax authority. We code these variables using a three-point scale: not in use (0), planning on being used/implemented (1), or actively in use (2).

Finally, we also examine the role of other country characteristics. We include the statutory tax rate (*STR*), as it is likely important for understanding country-level tax planning. Further, the machine learning model may uncover types of enforcement that are more (or less) effective when the country has a higher versus lower statutory rate.[2] We also include other

[2] In examining cash tax payments in the United States, Dyreng, Hanlon, and Maydew (2008) document that a significant portion of corporations are able to maintain long-term corporate tax rates that are substantially below the statutory tax rate.

country-level macro-environment variables such as *Ln(GDP), Ln(GDP per Capita), GDP Growth,* and *Control of Corruption,* as each of these features of a macro-environment may influence the effectiveness of tax enforcement.[3]

Our measure of corporate tax planning is the asset-weighted country-year average GAAP ETR. Our asset-weighted approach follows Shevlin et al. (2019) and allows for the tax planning of the largest firms in a country-year to have a bigger impact than that of smaller firms. In addition to overall tax planning, we also wish to understand how tax enforcement can differently affect various types of firms. In particular, there has been a recent focus on curbing the tax avoidance activities of multinationals that utilize income shifting. To examine how different types of tax enforcement might be more (or less) effective in curbing the avoidance of these firms, we also separately estimate the country-year asset-weighted average GAAP ETRs for those firms likely versus unlikely to be subject to Pillar 2 legislation, as well as those with relatively high versus low intangible assets. We identify Pillar 2 (non-Pillar 2) firm-years as those with sales above (below) €750 million. We identify high (low) intangible asset firm-years based on the median level of intangible assets within that country-year. For example, the *Asset-Weighted GAAP ETR – Pillar 2* of a country year is the asset-weighted average GAAP ETR for all firm-years with sales above €750 million. Appendix A provides detailed variable definitions for all variables.

3.4. Sample

Our sample begins with data from 136 country-year observations from 39 unique countries 2016-2019. We begin in 2016 and end in 2019 as this is the time period with detailed

[3] These characteristics could also influence tax avoidance itself. For example, financial constraints that occur in poor macroeconomic conditions are associated with tax avoidance (Edwards, Schwab, and Shevlin 2016), while DeBacker, Heim, and Tran (2015) find that higher levels of corruption are associated with increased tax evasion for US firms.

data on tax enforcement available from the OECD.[4] We calculate asset-weighted average ETRs for each country year following Shevlin et al. (2019). In calculating these asset-weighted ETRs, we require the underlying firm-year observations to have positive pretax income and sales, as well as non-missing data for pretax income, sales, and assets. We use Compustat NA to obtain data for firms incorporated in the United States or Canada, and Compustat Global to obtain data for firms incorporated in other countries. We use a random subset of 25 percent of the sample (34 country-year observations) to train the machine learning model (our training sample). This step results in a sample of 102 country-year observations for use in further analysis (our test sample). Table 1 provides details on the country-year observations used to train the machine learning model, as well as those that are used as a test sample for all subsequent analysis.

[Insert Table 1 Here]

4. Results

4.1. Descriptive Statistics

Table 2 Panel A presents descriptive statistics for the full sample. We find a mean asset-weighted GAAP ETR of 23.9 percent, with significant interquartile variation ranging from 19 percent to 28.4 percent. Average country-level tax enforcement expenditures are approximately 0.165 percent of GDP, and the average inflation adjusted change in tax enforcement spending is four percent. We find that about 26 percent of all tax authority staff works in the audit function, and about 3.5 percent of all tax authority staff work in the large taxpayer office. Approximately 7.4 percent of staff leaves the tax authority on average each year. Average tenure of all staff is 15.669 years. The average values of *AI and Machine Learning* and *Robotic Process Automation*

[4] Data on overall spending are available over a longer time period (e.g., De Simone, Stomberg, and Williams 2023), but data on items such as use of robotic process automation or machine learning and AI are available only in this shorter window.

suggest the average tax authority is planning on using AI and Machine learning techniques, whereas Robotic Process Automation is in earlier stages.

Table 2 Panel B presents correlations for the variables we use in our analysis. We present Pearson correlations on the bottom left, and Spearman on the top right. Correlations that are statistically significant at a 10% level or better are in bold. We find that our proxies for enforcement measure distinct constructs. The largest correlation is between *AI and Machine Learning* and *Robotic Process Automation* (correlation around 0.45), the second largest is between *Staff Tenure* and *% Staff Departing* (correlation around 0.40), and the third highest is between *AI and Machine Learning* and *STR* (correlation around 0.25). Most correlations have an absolute value of less than 0.2. The largest correlation between any two variables is between *Ln(GDP per Capita)* and *Control of Corruption,* (correlation around 0.87), which is perhaps not surprising as economic development is often correlated with higher quality institutions. Taken together, we interpret the correlation matrices as indicating that our measures of enforcement capture related, but distinct, constructs.

[Insert Table 2 Here]

4.2. Validating the EBM Prediction Model

As a first step, we examine the ability of the machine learning model to understand and predict country-year asset-weighted GAAP ETRs in an out-of-sample test (i.e., the test sample). This test uses what the model learned from the training data predict what leads to higher or lower country-year GAAP ETRs in the test sample. To evaluate the effectiveness of the machine learning prediction model, we borrow from Guenther, Peterson, Searcy, and Williams (2023) and examine the relation between the actual value of the target (*Y* = asset-weighted country-year

GAAP ETR) and the machine learning model's predicted value of the target ($\hat{y}$) by estimating the following OLS regression:

(3) *Asset-Weighted GAAP ETR* = $\alpha + \beta_1$*(Predicted Value)* + *e*

where *Asset Weighted GAAP ETR* is the actual country-year value and *Predicted Value* is either (1) the predicted value from an OLS regression of *Asset-Weighted GAAP ETR* as a function of all features or (2) the predicted value from the machine learning model. Both the OLS and EBM prediction models use the same training data and the same set of features. We compare the R-squared values from these two models to evaluate the predictive ability of the EBM model. These tests effectively compare the ability of the EBM model to predict *Asset-Weighted GAAP ETR* to that of an OLS regression.

Table 3 presents the results of this estimation. We test the differences in R-squared using a Vuong (1989) test.[5] We estimate results for the full test sample, as well as for the sub-samples of firm-years classified as *Pillar 2 (Not Pillar 2),* and *High (Not High) Intangibles*. Across all specifications, we find that the EBM model's prediction is superior to that of the OLS model. For example, for the full test sample, we find that the predicted value of *Asset-Weighted GAAP ETR* from the machine learning model explains 41 percent of the variation in the actual *Asset-Weighted GAAP ETR.* In comparison, the OLS model explains only 21.4 percent of the variation in *Asset-Weighted GAAP ETR.* We therefore conclude that our EBM model outperforms OLS in its predictive abilities.

[Insert Table 3 here.]

4.3. Understanding the Relation Between Enforcement Types and Avoidance

[5] We thank Jeff Woolridge for providing the code on which we base our Vuong (1989) tests.

Based on the conclusion that our EBM model outperforms OLS in predictive ability, we now evaluate our main research question: which features are most important in understanding country-year tax avoidance? To answer this question, we examine which features contribute most to the machine learning model's predicted country-year tax avoidance. The interpretable machine learning values for *Asset-Weighted GAAP ETR* provide insights into how the features of the model interact to explain the 41 percent of observed variation in the country-year *Asset-Weighted GAAP ETR* we find above. Table 4 presents the results of this estimation. Panel A presents the features aggregated by enforcement feature type (i.e., human capital, technology, and budget), while Panel B presents each feature's individual contribution.

[Insert Table 4 here.]

In Table 4 Panel A, we find in the full test sample that the machine learning model determines the features related to tax enforcement to be relatively important in predicting country-year asset-weighted GAAP ETRs. Specifically, these features account for between 48.4 percent to 57.0 percent of the contribution to the predicted asset-weighted GAAP ETRs. When breaking down the tax enforcement features by type, we find that features related to tax authority human capital are most important (27.1 percent), followed by features related to tax authority technology (14.3 percent) and tax authority budget (9.8 percent). Features related to the tax authority budget are likely correlated with items that are not directly measured in the other tax authority attributes such as staff education or the number of offices that the tax authority has within a country.[6] When we allow for variation across firm types, we find that tax authority human capital and budget are most important for firms that are not likely subject to Pillar 2,

[6] Kubick, Lockhart, Mills, and Robinson (2017) document that in the United States, a closer geographic proximity between the tax authority office (i.e. IRS office) and the corporate taxpayer is associated with lower levels of corporate tax avoidance.

while the use of technology is most important for firms with high intangible assets. Perhaps not surprisingly, we also find that non-tax attributes are important in understanding country-year tax planning, accounting for between 25.3 percent to 41.7 percent of the predicted value of country-year asset-weighted GAAP ETRs.

In Table 4 Panel B, we present the detailed contribution of each feature to the machine learning model's prediction by firm type. Figure 1 depicts the relative importance of the features for all firm-years, while Figure 2 Panels A and B graph the contributions based on based on firm type. We find that the statutory tax rate is generally the most important feature, which provides additional comfort around our EBM model given we are predicting GAAP ETRs. We find that the control of corruption is the second most important for the overall GAAP ETR (All Firms), and it is also the most important feature for firms with low intangibles. In terms of tax enforcement variables, we generally find that the use of *AI and Machine Learning* is the most important tax enforcement variable, but that there is significant variation based on the complexity of the firm. In particular, AI *and Machine Learning* is significantly more important for Pillar 2 (9.8 percent) versus non-Pillar 2 firms (3.2 percent), and for high intangible (12.6 percent) versus low intangible firms (5.5 percent). We find the *% Staff – Large Taxpayer* is the second most important tax enforcement variable, although it exhibits less variation in importance than *AI and Machine Learning* across different firm types. The next most important tax enforcement features are *% Staff – Audit Function, % Staff Departing, Tax Expenditure/GDP, Staff Tenure, % Change Enforce,* and *Robotic Process Automation.*

The pattern of results for *Robotic Process Automation* is opposite that for *AI and Machine Learning.* In particular, the use of *Robotic Process Automation* is most useful for *less* complex firms, accounting for 8.4 percent (3.6 percent) of the contribution to the prediction for firms that are not Pillar 2 (are Pillar 2), and similarly 8.7 percent (5.4 percent) for Low Intangible Asset (High Intangible Asset) firms. This result suggests that not all technology has the same effect. In particular, automating lower skill tasks appears to help more in mitigating avoidance by less complex firms, while implementing more advanced analytics helps more in mitigating avoidance by more complex firms.

4.4. What Influences the Effectiveness of Enforcement Mechanisms?

In our next set of analysis, we investigate what leads to relatively higher or lower importance of tax enforcement features. Specifically, we aim to understand how individual tax enforcement features influence each other and what role country-level characteristics have on the contributions of the tax enforcement features. We conduct this analysis by using the interpretable machine learning values for each tax enforcement feature, in turn, as the dependent variable in the regression. We include as independent variables the actual (not interpretable machine learning) values of all of the other features. For example, we estimate regressions of the form:

(4) Abs. Value (Interpretable ML Value of *AI and Machine Learning*) = α+ β_1*(STR)* + β_2*(AI and Machine Learning)* + β_3*(% of Staff – Large Taxpayer)* + β_4*(GDP Growth)* + β_5*(% of Staff – Audit Function)* + β_6*(Ln(GDP))* + β_7*(% Staff Departing)* + β_8*(ln(GDP per Capita)))* + β_9*(Tax Expenditure / GDP)* + β_{10}*(Staff Tenure)* + β_{11}*(% Change Enforce)* + β_{12}*(Robotic Process Automation)* + *Country Fixed Effects* + *Year Fixed Effects* + *e*

The coefficients in these regressions reveal how the actual value of each feature is associated with the contribution of the others. We include country and year fixed effects in these regressions to control for any time or country-level correlated omitted variables.[7]

Table 5 presents the results. The first dependent variables we examine relate to the tax enforcement budget. In Columns (1) and (2), the dependent variables are the absolute value of the interpretable machine learning values for *Tax Expenditures/GDP* and *% Change Enforce*,

[7] We do not include country or year fixed effects in our first step of machine learning training due to the small size of our training sample (34 observations).

respectively. We find that *% Staff – Audit Function* is negative and significant in both columns. These findings reveal that when the percentage of staff in the audit function increases, the tax enforcement budget has a smaller impact on tax avoidance. In terms of economic significance, a one standard deviation increase in *% Staff – Audit Function* (0.095) is associated with approximately a 19.8 percent decrease in the expected effect of *Tax Expenditure/GDP* on the predicted country-year *GAAP ETR.*[8] Thus, audit staff and financial resources can be substitutes. In contrast, we find that *Staff Tenure* is positive and significant, which means when staff tenure increases, the tax enforcement budget has a larger impact on tax avoidance. Thus, having staff with a longer tenure (and likely more expertise) complements a higher budget in mitigating tax avoidance. Interestingly, we also find that the overall budget (*Tax Expenditure/GDP*) is predicted to have a greater impact on tax avoidance when *STR* is lower (Column 1), but that the impact of a change in the enforcement budget (Column 2) does not vary with the *STR*.

We next examine the impact of country-level features on the machine learning model's prediction of the effectiveness *AI and Machine Learning* and *Robotic Process Automation.* We find that country-level features are associated with the predicted effectiveness of these features somewhat differently. For example, we find that the expected impact of *AI and Machine Learning* is higher when the total tax enforcement budget is larger, when there are more staff in the large taxpayer division, and when the country-year observation has higher GDP per capita. In contrast, we find that *Robotic Process Automation* matters more when there is more corruption given the negative coefficient on *Control of Corruption.* We also find that the use of *AI and Machine Learning* increases the predicted effectiveness of *Robotic Process Automation,* however

[8] We multiply all coefficients by 100 for convenience. Thus, the coefficient on *% Staff – Audit Function* of -0.461 * a one standard deviation change of *% Staff – Audit Function* (0.095) = 0.0004375 change in the role of *Tax Expenditure/GDP* on *Asset-Weighted GAAP ETR.* This value equates to 0.0004375 divided by mean *abs. iML Tax Expenditure/GDP* of approximately 0.0022125 (untabulated).

we do not find that *Robotic Process Automation* increases the predicted effectiveness of *AI and Machine Learning.*

Columns 5-8 present results using the interpretable machine learning values *(iML)* for *% Staff – Audit Function, % Staff – Large Taxpayer, % Staff Departing,* and *Staff Tenure.* Interestingly, we find that while *Robotic Process Automation* is associated with an increase in the predicted impact of the *% Staff – Large Taxpayer,* it is associated with a *decrease* in the predicted impact of the *% Staff – Audit Function.* We find that a higher or increased tax enforcement budget is associated with an increase in the predicted impact of *% Staff – Large Taxpayer,* as well as an increase in the impact of % *Staff Departing.* Interestingly, we find that having a higher proportion of staff in the audit function is predicted to have a larger impact when there is less control of corruption. Finally, having a higher statutory rate is associated with a higher predicted importance of the *Staff Tenure*, and a lower predicted importance of *Staff Departing*. The wide variation in explanatory power of these tests, as measured using the adjusted R^2 values, suggests similarly wide variation in our ability to explain the relative importance of each feature.

[Insert Table 5 here.]

4.5. How Do Tax Enforcement Features' Contributions Differ across Subsamples of Firms?

Given the focus on curbing international tax avoidance, we next investigate potential differences in the predicted impact of tax enforcement features on country-year tax avoidance separately estimated for *Pillar 2* and *Non-Pillar 2* firms. We first present estimations based on Tax Enforcement Budget (Table 6, Panel A). The estimated differences in the features associated with higher or lower predicted contributions of various tax enforcement features on ETRs are striking. In particular, we find that the contribution is much more contextual for *Pillar 2* firms, with several features influencing the predicted impact of *Tax Expenditure/GDP* (Panel A Column 1) as well as *% Change Enforce* (Panel A Column 2) for *Pillar 2* firms, compared to only one feature for *Non-Pillar 2* firms (Panel A Column 4). For example, for *Pillar 2* firms we find that *Staff Tenure* is associated with an increase in the predicted contribution of tax enforcement budget features whereas *% Staff – Audit Function* is associated with a decrease in the predicted contribution of tax enforcement budget features.

In terms of the use of technology (Table 6 Panel B), we find more similarities between *Pillar 2* and *Non-Pillar 2* firms. Specifically, we find that for both sets of firms the use of *AI and Machine Learning* is associated with an increase in the predicted contribution of *Robotic Process Automation* on ETRs (Columns 2 and 4), Similarly, for both sets of firms we find that having less control over corruption is associated with an increase in the predicted contribution of *Robotic Process Automation* on ETRs (i.e. negative coefficient on *Control of Corruption* in Columns 2 and 4). In terms of differences, we find that the total budget *(Tax Expenditure/GDP)* is associated with an increase in the predicted contribution of *AI and Machine Learning* for *Pillar 2* firms, but not *for Non-Pillar 2* firms.

We next evaluate human capital (Table 6 Panel C). Here we once again find substantial differences between *Pillar 2* and *Non-Pillar 2* firms. The budgetary variables have a larger effect on the predicted contribution of human capital related measures for *Pillar 2* firms versus *Non-Pillar 2* firms, with a higher budget being predicted to increase the contribution of *% Staff – Large Taxpayer* as well as leading to a higher predicted contribution of *% Staff Departing* for *Pillar 2* firms. For *Non-Pillar 2* firms, *Staff Tenure* appears to be the most contextual human capital variable, with four of the features leading to a higher or lower predicted contribution of *Staff Tenure* on effective rates of *Non-Pillar 2* firms (Column 8).

[Insert Table 6 here.]

In our final set of tests, we investigate potential differences in the predicted contribution of tax enforcement features on country-year tax avoidance separately estimated for firms with high versus low intangible assets. We present these results in Table 7. In Panel A, we focus on features related to the Tax Enforcement Budget — *Tax Expenditure/GDP* and *% Change Enforce.* We find substantial differences in contributions between high and low intangible firms. For example, an increase in *Robotic Process Automation* is associated with *% Change Enforce* being a more important determinant of *GAAP ETR* for intangible intensive firms whereas an increase in *% Staff – Audit Function* makes *% Change Enforce* a more important determinant of *GAAP ETR* among firms that are not intangible-intensive. Furthermore, we find that *Staff Tenure* has opposite effects on the contribution of *% Change Enforce* for high vs. low intangible intensive firms, with *Staff Tenure* being associated with an increase (decrease) in the predicted contribution of *% Change Enforce* for high (low) intangible firms.

One more consistent finding in Table 7 Panel A is that lower statutory tax rates are associated with relatively higher importance of the tax enforcement budget for both high and low intangible firms (*STR* – Columns 2 and 3). However, we also find differences in that for high intangible firms, having more staff departures and higher staff tenure is associated with an increased importance in the change in budget, whereas for low intangible firms, staff tenure is *negatively* associated with the importance of the change in the tax enforcement budget.

In Table 7 Panel B, we explore the contribution of tax enforcement technology features. We continue to find substantial differences between high and low intangible firms. For example, we find that the use of *Robotic Process Automation* appears to be a complement to the importance of *AI and Machine Learning* for high intangible firms (Column 1), but a substitute for low intangible firms (Column 3). Similarly, whereas a higher control of corruption leads to increased importance for *AI and Machine Learning* for high intangible firms, it is associated with lower importance of *AI and Machine Learning* for low intangible firms. These findings suggest that the corruption in a country can have a significant influence on the relative importance of the use of analytics, particularly as it relates to firms with relatively higher or lower levels of intangible assets.

We next examine the relative importance of human capital. We start with similarities between the high and low intangible firms. Specifically, we find that for both sets of firms, the use of robotic process automation is predicted to increase in the contribution of *% Staff – Large Taxpayer* (Columns 2 and 6). Similarly, *AI and Machine Learning* is predicted to decrease the importance of *Staff Tenure* (Columns 4 and 8). However, once again we find far more differences than similarities in how features of tax enforcement influence the predicted impact of other features of tax enforcement. For example, we find that there are more interactive effects with the change in budget as well as macroeconomic features *(Ln(GDP), Ln(GDP per Capita)* for firms with high intangible assets relative to firms with low intangible assets. Overall, we take the results in Table 7 to suggest that the predicted contribution of various features of tax enforcement is contextual and substantially different between firms with high versus low levels of intangible assets.

[Insert Table 7 here.]

5. Conclusion

This study uses recent developments in machine learning and interpretable artificial intelligence to understand how features of a country's tax enforcement, tax policy, and economic environment impact country-level tax avoidance measured using asset-weighted GAAP ETRs.

Specifically, we utilize an explainable boosting machine (EBM) learning algorithm that estimates predictions with high interpretability and allow the user to understand what leads the AI model to generate the prediction. This approach also provides high dimensionality of the interpretable predictions, providing the contribution of each individual feature to each observation's predicted value. After validating that the predictive ability of this model outperforms that of standard OLS, we estimate the relative contribution of various features of country-level tax enforcement and their pairwise interactions on country-level tax planning.

We find that tax enforcement features are highly important in predicting country-year asset-weighted GAAP ETRs, accounting for 51.3 percent of the contribution to the algorithm's prediction. Within tax enforcement features, we find that features related to tax authority human capital are most important (27.1 percent), followed by features related to tax authority technology (14.3 percent) and tax authority budget (9.8 percent). Additionally, we find the statutory tax rate accounts for 16.8 percent of the contribution whereas GDP and corruption contribute the remaining 31.9 percent.

Our approach also allows us to evaluate any differential importance of features based on types of firms. For example, we find AI and machine learning are significantly more important features of tax avoidance for firms likely to be subject to the global minimum tax under Pillar 2 of the OECD/G20's Base Erosion and Profit Shifting initiative, as well as for firms with high intangible assets. Finally, our approach also allows us to understand how features interact to contribute to country-level tax avoidance. We find that complementarities between financial resources, human capital, and automation - that human tax agents compensate for a lack of tax authority resources. We also find that when the statutory tax rate is lower, the tax enforcement budget and staff turnover have a greater impact on corporate tax avoidance. Our findings are useful to policymakers, corporate taxpayers, and researchers.

References

Atwood, T.J., M. Drake, J. Myers, and L. Myers. 2012. Home country tax system characteristics and corporate tax avoidance: International evidence. *The Accounting Review* 87 (6): 1831-1860.

Ayers, B., Seidman, J., Towery, E. 2019. Tax Reporting Behavior Under Audit Certainty. *Contemporary Accounting Research.* 36 (1). 326-358.

Belnap, A., K. Kroeger, and J. Thornock. 2023. Explaining tax avoidance: Insights from thirty years of tax research. Working paper.

Beuselinck, Deloof, and Vanstraelen (2015), Cross-jurisdictional income shifting and tax enforcement: evidence from public versus private multinationals. *Review of Accounting Studies* 20. 710-746.

Chen, X., Cho, Y., Dou, Y., and Lev, B. 2022. Predicting Future Earnings Changes using Machine Learning and Detailed Financial Data. *Journal of Accounting Research.* 60 (2). 467-515.

Chi, S., A. Persson, T. Shevlin, and O. Urcan. 2022. The deterrence effect of cross-border monitoring: Evidence from foreign tax authority enforcement. Working paper.

DeBacker, J., Heim, B., Tran, A. 2015. Importing corruption culture from overseas: Evidence from corporate evasion in the United States. *Journal of Financial Economics.* 117 (1). 122-138.

Donovan, J., Jennings, J., Koharki, K., and Lee, J. 2021. Measuring credit risk using qualitative disclosure. *Review of Accounting Studies* 26. 815-863.

De Simone, L., Mills, L., and Stomberg, B. 2019. Using IRS data to identify income shifting to foreign affiliates. *Review of Accounting Studies*, 24: 694-730.

De Simone, L., Stomberg, B., and Williams, B. 2023. Does tax enforcement disparately affect domestic versus multinational corporations around the world? Forthcoming. *Contemporary Accounting Research.*

Dyreng, S., Hanlon, M, and Maydew, E. 2008. Long-Run Corporate Tax Avoidance. *The Accounting Review.* 83 (1). 61-82.

Edwards, A., Schwab, S., and Shevlin, T. 2016. Financial Constraints and Tax Savings. *The Accounting Review.* 91 (3). 859-881.

Erel, I., Stern, L., Tan, C., and Weisbach, M. 2021. Selecting directors using machine learning. *Review of Financial Studies.* 34: 3226-3264.

Guenther, D., K. Peterson, J. Searcy, and B. Williams. 2023. How useful are tax disclosures in predicting effective tax rates? A Machine Learning approach. *The Accounting Review* 98 (5). 297-322.

Gupta, S. and D. Lynch. 2016. The effects of changes in state tax enforcement on corporate income tax collections. *Journal of the American Taxation Association* 38 (1): 125-143.

Hoopes, J., D. Mescall, and J. Pittman. 2012. Do IRS audits deter corporate tax avoidance? *The Accounting Review* 87 (5): 1603-1639.

Hoopes, J., L. Robinson, and J. Slemrod. 2018. Public tax-return disclosure. *Journal of Accounting and Economics* 66: 142-162.

Jennings, J., Lee, J., and Towery, E. 2021. Estimating the likelihood of future tax settlements *Working Paper.*

Joshi, P. and Weng. 2022. Does artificial intelligence and machine learning improve the tax enforcement function? Working paper.

Keen, M. and J. Slemrod. 2017. Optimal tax administration. *Journal of Public Economics* 152: 133-142.

Klassen, K. and Laplante, S. 2012. The effect of foreign reinvestment and financial reporting incentives on cross-jurisdictional income shifting. *Contemporary Accounting Research*, 29, 928–955.

Kubick, T., Lockhart, G.B., Mills, L., and Robinson, J. 2017. IRS and corporate taxpayer effects of geographic proximity. *Journal of Accounting and Economics.* 63 (2-3). 428-453.

Nessa, M., C. Schwab, B. Stomberg, and E. Towery. 2020. How do IRS resources affect the corporate audit process? *The Accounting Review* 95(2): 311-338.

Nori, H., Jenkins, S., Koch, P., Caruana, R. 2019. InterpretML: A Unified Framework for Machine Learning Interpretability. arXiv: 19-9.09223v1

Seidman, J., R. Sinha and B. Stomberg. 2022. The Effect of Tax Authorities on Corporate Tax Planning: Insights from Tax Executives. Working paper.

Shevlin, T. Shivakumar, L., Urcan, O. 2019. Macroeconomic effects of corporate tax policy. *Journal of Accounting and Economics.* 68 (1). 1-22

Slemrod 2015. Tax System Design: The Role of Enforcement. Keynote lecture, Inaugural conference of the Welfare State and Taxation Unit of the Dondena Research Centre on Social Dynamics and Public Policy, Bocconi University,

Slemrod, J. and C. Gillitzer. 2014. *Tax Systems.* Oxford: Oxford University Press.

Vuong, Q. 1989. Likelihood Ratio Tests for Model Selection and Non-Nested Hypothesis. *Econometrica* 57 (2). 307-333

Appendix A - Variable Definitions

Variable	**Definition**
Asset-Weighted GAAP ETR	Pretax Income (PI) divided by Total Tax Expense (TXT), weighted by assets in that country-year.
Asset-Weighted GAAP ETR – Pillar 2	Asset-weighted GAAP ETR for firm-year observations with greater than or equal to €750 Million Euros in Sales (SALE)
Asset-Weighted GAAP ETR – Not Pillar 2	Asset-weighted GAAP ETR for firm-year observations with less than €750 Million Euros in Sales (SALE)
Asset-Weighted GAAP ETR – High Intangibles	Asset-weighted GAAP ETR for firm-year observations with above-median country-year Intangible Assets scaled by total assets (INTAN/AT)
Asset-Weighted GAAP ETR – Low Intangibles	Asset-weighted GAAP ETR for firm-year observations with below-median country-year Intangible Assets scaled by Total Assets (INTAN/AT)
STR	The country's statutory tax rate
Control of Corruption	Control of Corruption from the World Bank Governance Indicators. "Control of corruption captures perceptions of the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as "capture" of the state by elites and private interests." (World Bank, 2022)
% Staff – Large Taxpayer	Percentage of all tax authority staff who work in the large taxpayer division, defined as total full-time employees in large taxpayer program, divided by total full-time employees.
AI and Machine Learning	Use of Artificial Intelligence and Machine Learning by the tax authority. 0 = Not Using, 1 = Planning on Using, 2 = In Use
% Staff Departing	Number of staff that departed during the year divided by the number of staff at the end of the year.
Ln(GDP)	Natural Log of GDP (from the World Bank)

GDP Growth	Year- over-year percentage growth in GDP (from the World Bank)
Ln(GDP per Capita)	Natural Log of GDP per Capita (from the World Bank)
% Staff – Audit Function	Percent of staff in the audit function, defined as number of staff in audit, investigation, and other verification, over number of staff at the end of the year.
Robotic Process Automation	Use of Robotic Process Automation by the tax authority. 0 = Not Using, 1 = Planning on Using, 2 = In Use
Staff Tenure	Average tenure as staff, defined as: 2.5 x the % of staff with less than 5 years of service + 7.5 x the % of staff with 5-9 years of service + 15 x the % of staff with 10-19 years of service + 25 x the % of staff with 20 or more years of experience
Tax Expenditure/GDP	Total Expenditures on the Tax Administration, Divided by GDP
% Change Enforce	Inflation adjusted percent change in Tax Expenditure
Abs. iML	Absolute value of the interpretable machine learning value for the feature. For example *Abs. iML % Change Enforce* is the absolute value of the interpretable machine learning value for *% Change Enforce*

Figure 1
Relative Contribution to Predicted Values – Asset Weighted GAAP ETR (All Firms)

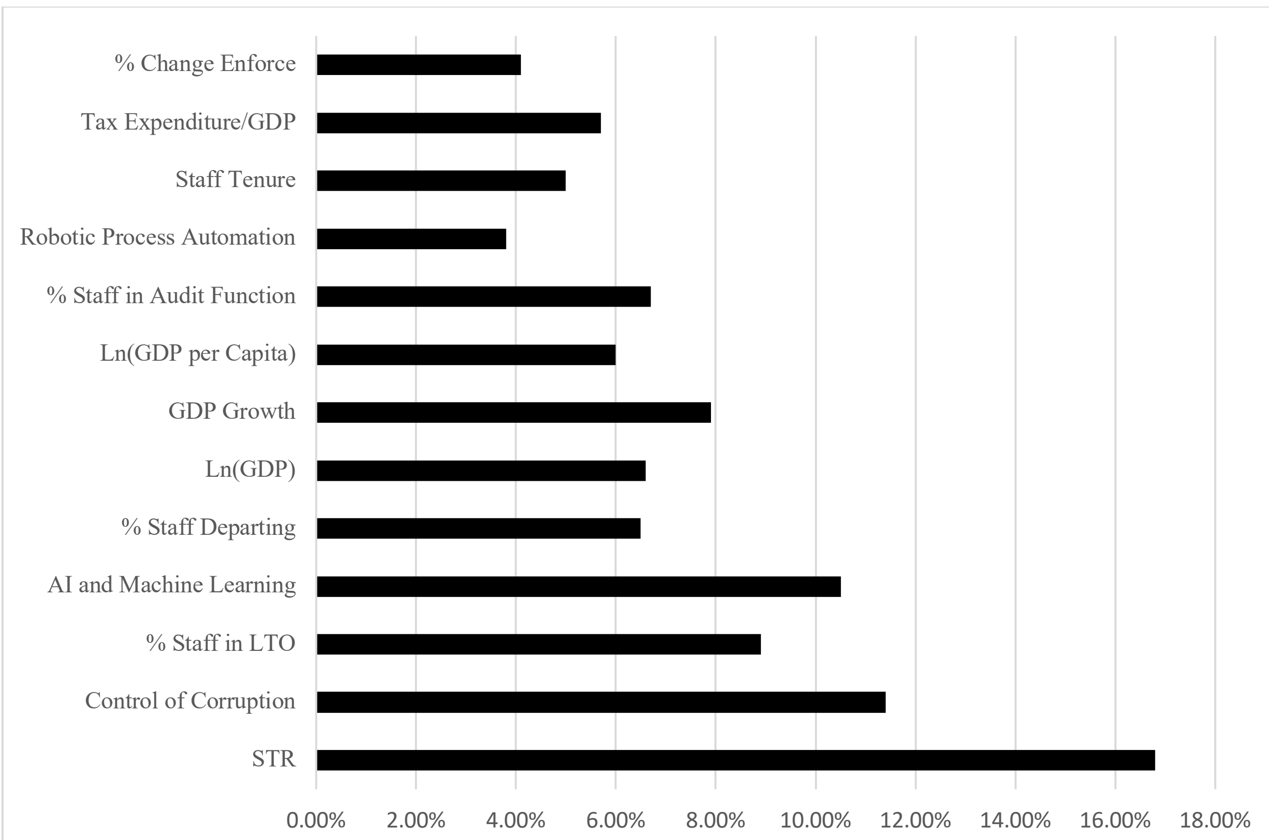
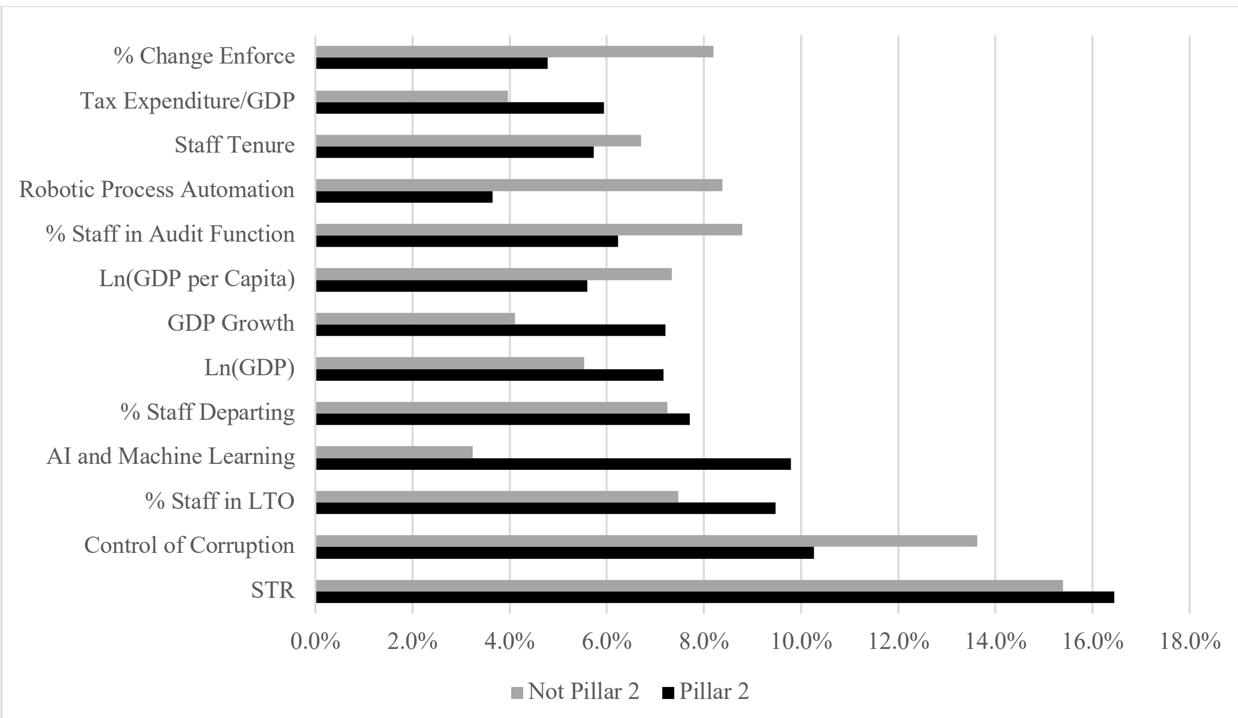
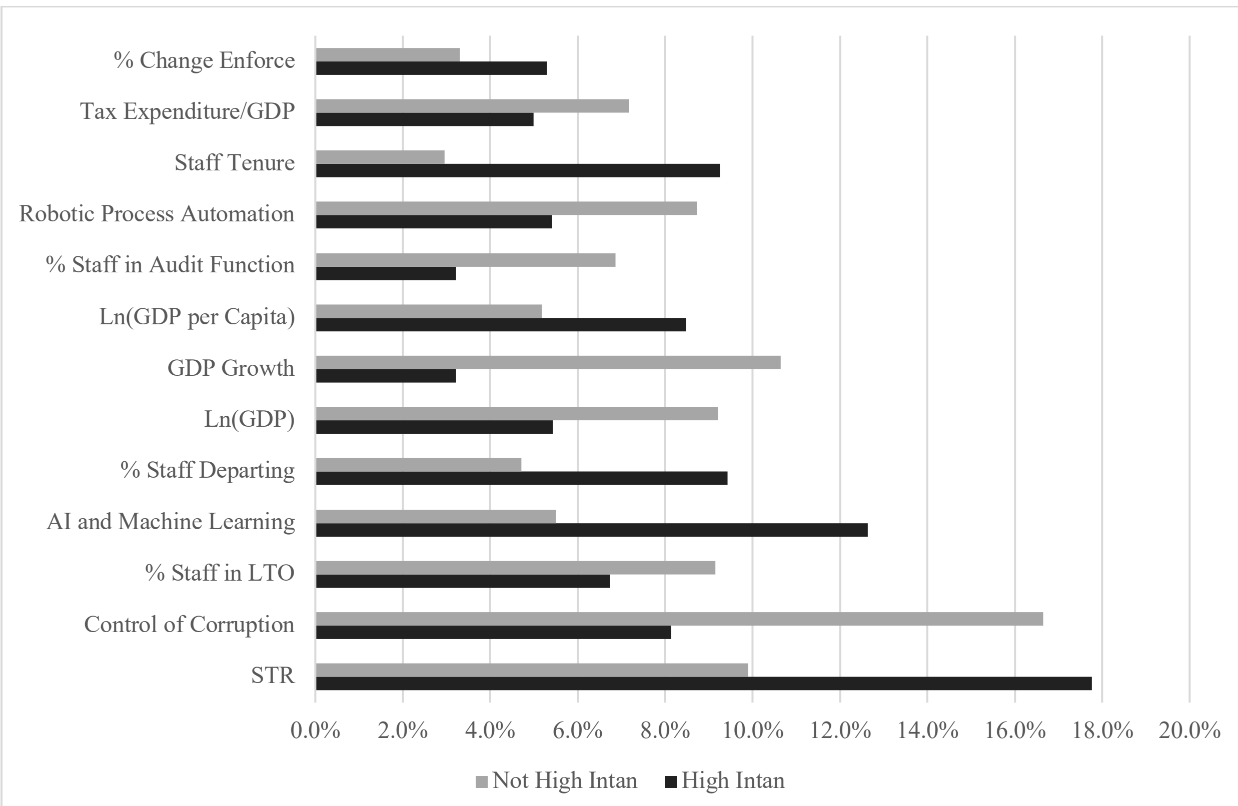


Figure 2
Relative Contribution to Predicted Values – Pillar 2 versus Not Pillar 2 Firms



This figure presents the contribution of each tax enforcement feature as well as *STR, Control of Corruption* and economic variables splitting firm-year observations into whether they are likely to have to comply with OECD's Pillar 2 (*Pillar 2*) or not (*Not Pillar 2*). We identify firms likely to be subject to Pillar 2 as those with greater than or equal to €750 Million Euros in Sales. All other variables are defined in Appendix A.

Figure 2 Panel B
Relative Contribution to Predicted Values – Based on Intangible Intensity



This figure presents the contribution of each tax enforcement feature as well as *STR, Control of Corruption* and economic variables splitting firm-year observations into whether they are intangible intensive (*High Intan*) or not (*Not High Intan*). We identify firms likely as intangible intensive as those with above-median country-year Intangible Assets as a proportion of Total Assets (INTAN/AT). All other variables are defined in Appendix A.

Table 1
Number of Country-Year Observations in Sample

Country	Total Obs.	Obs. Used for Training	Obs In Test Sample
ARGENTINA	1	1	0
AUSTRALIA	4	0	4
AUSTRIA	4	1	3
BELGIUM	4	1	3
CANADA	4	2	2
CHILE	4	0	4
COLOMBIA	4	2	2
CROATIA	4	1	3
CYPRUS	4	1	3
CZECH REPUBLIC	4	2	2
DENMARK	4	1	3
ESTONIA	4	0	4
FINLAND	4	1	3
FRANCE	4	1	3
HUNGARY	1	0	1
ICELAND	2	0	2
IRELAND	4	0	4
ISRAEL	4	1	3
ITALY	4	1	3
KENYA	2	0	2
LITHUANIA	3	2	1
LUXEMBOURG	1	0	1
MALAYSIA	4	0	4
MEXICO	4	2	2
MOROCCO	4	1	3
NETHERLANDS	4	1	3
NEW ZEALAND	4	1	3
NORWAY	4	1	3
PERU	4	2	2
POLAND	1	0	1
PORTUGAL	4	0	4
SLOVENIA	4	0	4
SOUTH AFRICA	4	2	2
SPAIN	4	0	4
SWEDEN	4	1	3
SWITZERLAND	4	1	3
THAILAND	3	1	2

UNITED KINGDOM	2	1	1
UNITED STATES	4	2	2
Total	**136**	**34**	**102**

Table 2
Panel A: Descriptive Statistics

Variable	N	Mean	SD	p25	p50	p75
GAAP ETR	136	0.239	0.068	0.190	0.238	0.284
Tax Expenditure / GDP	136	0.165	0.070	0.113	0.171	0.210
GDP Growth	136	2.851	1.585	1.915	2.566	3.771
% Change Enforce	136	0.040	0.193	-0.051	0.009	0.072
STR	136	0.240	0.059	0.200	0.240	0.280
AI and Machine Learning	136	0.985	0.720	0.000	1.000	1.500
Robotic Process Automation	136	0.654	0.734	0.000	0.500	1.000
% Staff – Audit Function	136	0.260	0.095	0.197	0.248	0.302
% Staff – Large Taxpayer	136	0.035	0.028	0.018	0.031	0.046
Control of Corruption	136	0.979	0.929	0.194	1.022	1.817
% Staff Departing	136	0.074	0.038	0.043	0.062	0.101
Staff Tenure	136	15.669	3.169	12.886	15.654	17.961
Ln(GDP)	136	26.567	1.423	26.021	26.566	27.370
Ln(GDP per Capita)	136	10.101	0.915	9.514	10.341	10.768

Table 2, Panel A presents descriptive statistics for variables used in our analysis. All variables are defined in Appendix A.

Table 2
Panel B: Correlations

	Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	GAAP ETR		-0.06	**-0.41**	-0.07	**0.54**	**0.21**	0.06	-0.09	**0.31**	-0.13	0.05	**-0.17**	**0.27**	-0.09
2	Tax Expenditure / GDP	-0.04		0.01	0.13	-0.08	**0.16**	0.11	-0.12	0.02	**0.28**	0.05	0.23	-0.12	0.14
3	GDP Growth	**-0.39**	-0.04		0.06	**-0.39**	**-0.19**	**-0.16**	-0.03	-0.06	**-0.17**	-0.11	0.03	**-0.53**	**-0.18**
4	% Change Enforce	-0.02	0.08	0.00		-0.11	**0.22**	0.01	-0.07	-0.08	0.03	-0.11	**-0.20**	**-0.17**	0.01
5	STR	**0.51**	**-0.12**	**-0.46**	0.01		**0.25**	**-0.14**	-0.12	**0.25**	**-0.18**	0.02	**-0.24**	**0.41**	**-0.20**
6	Ai and Machine Learning	**0.19**	**0.19**	**-0.19**	**0.17**	**0.25**		**0.45**	-0.05	**0.20**	**0.19**	0.03	-0.04	**0.45**	**0.24**
7	Robotic Process Automation	0.07	**0.14**	-0.06	-0.04	**-0.16**	**0.44**		0.09	**0.21**	**0.43**	**0.22**	0.04	**0.28**	**0.53**
8	% Staff in Audit Function	-0.10	-0.05	0.02	-0.12	-0.14	-0.07	0.06		**0.17**	0.11	-0.07	0.14	0.10	**0.21**
9	% Staff in Large Taxpayer Office	**0.32**	0.13	-0.04	-0.10	**0.20**	**0.23**	**0.22**	0.05		0.08	0.12	**-0.20**	**0.20**	0.09
10	Control of Corruption	**-0.16**	**0.28**	-0.09	-0.14	**-0.18**	**0.19**	**0.41**	0.11	0.06		**0.30**	-0.12	**0.18**	**0.87**
11	% Staff Departing	0.03	0.03	-0.12	-0.12	0.09	0.04	0.14	-0.12	**0.20**	**0.23**		**-0.40**	**0.26**	0.32
12	Staff Tenure	**-0.15**	**0.24**	0.05	**-0.20**	**-0.22**	-0.03	0.06	**0.18**	**-0.15**	-0.07	**-0.38**		-0.01	-0.01
13	Ln(GDP)	**0.27**	**-0.16**	**-0.45**	-0.10	**0.51**	**0.41**	**0.25**	-0.03	**0.20**	**0.15**	**0.29**	-0.08		**0.39**
14	Ln(GDP per Capita)	**-0.18**	**0.21**	-0.10	**-0.16**	**-0.23**	**0.19**	**0.47**	**0.24**	0.00	**0.87**	**0.25**	0.08	**0.33**	

Table 2, Panel B presents correlation matrices for variables used in our analysis. The bottom left (top right) present Pearson (Spearman) correlations. Correlations that are statistically significant at the 10% level or better are in bold. All variables are defined in Appendix A.

Table 3
Machine Learning Model's Ability to Explain Out of Sample Country-Year Asset Weighted GAAP ETR (using the test-sample)

	OLS Prediction Model (1)	EBM Model's Prediction (2)
Asset-Weighted GAAP ETR	0.214	0.410*
Asset-Weighted GAAP ETR – Pillar 2	0.190	0.389*
Asset-Weighted GAAP ETR – NOT Pillar 2	0.190	0.370*
Asset-Weighted GAAP ETR – High Intangibles	0.178	0.293*
Asset-Weighted GAAP ETR – Low Intangibles	0.036	0.243*

Table 3 presents results of estimating:

$$\textit{Country-Year Asset Weighted GAAP ETR} = \alpha + \beta_1(\textit{Predicted Value}) + e$$

on a testing sample of 102 country-year observations. *Predicted* value in column (1) is the estimate from an OLS regression using all the same tax enforcement features included in the machine learning model and in column (2) is the predicted value from the machine learning model. ***, **, and * indicate the EBM model has a significantly higher R-squared than the OLS model at the 10% level or better (two-tailed) using the Vuong (1989) test. All variables are defined in Appendix A.

Table 4
Contributions to Machine Learning Model's Prediction by Feature Type and Firm Type

Panel A: Proportional Contributions Aggregated by Feature Type and Firm Type

Feature Type	All Firms	Pillar 2	Not Pillar 2	High Intan	Low Intan
Tax Enforcement Human Capital	27.1%	29.1%	30.2%	28.6%	23.7%
Tax Enforcement Technology	14.3%	13.4%	11.6%	18.0%	14.2%
Tax Enforcement Budget	9.8%	10.7%	12.2%	10.3%	10.5%
Sum of Tax Enforcement Features	51.3%	53.3%	54.0%	57.0%	48.4%
Statutory Rate	16.8%	16.4%	15.4%	17.8%	9.9%
Non-Tax Characteristics	31.9%	30.2%	30.6%	25.3%	41.7%

This Panel shows the proportional contribution of the features to the machine learning model's predicted value of asset-weighted GAAP ETR for the test sample, aggregated by feature type. *All Firms* is the *asset-weighted GAAP ETR, Pillar 2* is the *asset-weighted GAAP ETR – Pillar 2, Not Pillar 2* is the *asset-weighted GAAP ETR- Not Pillar 2. High Intan* is the *asset-weighted GAAP ETR – High Intangibles. Low Intan* is the *asset-weighted GAAP ETR – Low Intangibles.*

Panel B: Each Feature's Unique Proportional Contribution to the Machine Learning Model's Prediction

Feature Type	**Feature**	**All Firms**	**Pillar 2**	**Not Pillar 2**	**High Intan**	**Low Intan**
Statutory Rate	*STR*	16.8%	16.4%	15.4%	17.8%	9.9%
Non-Tax Characteristic	*Control of Corruption*	11.4%	10.3%	13.6%	8.1%	16.6%
TE - Technology	*AI and Machine Learning*	10.5%	9.8%	3.2%	12.6%	5.5%
TE – Human Capital	*% Staff – Large Taxpayer*	8.9%	9.5%	7.5%	6.7%	9.2%
Non-Tax Characteristic	*GDP Growth*	7.9%	7.2%	4.1%	3.2%	10.6%
TE – Human Capital	*% Staff – Audit Function*	6.7%	6.2%	8.8%	3.2%	6.9%
Non-Tax Characteristic	*Ln(GDP)*	6.6%	7.2%	5.5%	5.4%	9.2%
TE – Human Capital	*% Staff Departing*	6.5%	7.7%	7.2%	9.4%	4.7%
Non-Tax Characteristic	*Ln(GDP per Capita)*	6.0%	5.6%	7.3%	8.5%	5.2%
TE - Budget	*Tax Expenditure/GDP*	5.7%	5.9%	4.0%	5.0%	7.2%
TE – Human Capital	*Staff Tenure*	5.0%	5.7%	6.7%	9.3%	3.0%
TE - Budget	*% Change Enforce*	4.1%	4.8%	8.2%	5.3%	3.3%
TE - Technology	*Robotic Process Automation*	3.8%	3.6%	8.4%	5.4%	8.7%

This Panel shows the proportional contribution of each feature to the machine learning model's predicted value of asset-weighted GAAP ETR for the test sample. *All Firms* is the *asset-weighted GAAP ETR, Pillar 2* is the *asset-weighted GAAP ETR – Pillar 2, Not Pillar 2* is the *asset-weighted GAAP ETR- Not Pillar 2. High Intan* is the *asset-weighted GAAP ETR – High Intangibles. Low Intan* is the *asset-weighted GAAP ETR – Low Intangibles.*

Table 5
What Influences the Effectiveness of Tax Enforcement Features?

VARIABLES	*Abs. iML Tax Expenditure / GDP* (1)	*Abs. iML % Change Enforce* (2)	*Abs. iML AI and Machine Learning* (3)	*Abs. iML Robotic Process Automation* (4)	*Abs. iML % Staff – Audit Function* (5)	*Abs. iML % Staff – Large Taxpayer* (6)	*Abs. iML % Staff Departing* (7)	*Abs. iML Staff Tenure* (8)
Tax Expenditure / GDP		-0.698	**1.410****	-0.414	-0.298	-1.136	**2.740*****	0.383
		(-1.52)	**(2.08)**	(-0.93)	(-0.33)	(-1.45)	**(3.45)**	(0.99)
GDP Growth	0.001	-0.007	0.003	0.005	-0.007	-0.001	-0.018	-0.016
	(0.10)	(-0.48)	(0.32)	(0.85)	(-0.44)	(-0.07)	(-1.08)	(-1.09)
% Change Enforce	0.061		-0.010	-0.030	-0.091	**0.239****	0.119	-0.018
	(1.14)		(-0.23)	(-0.61)	(-0.93)	**(2.48)**	(1.51)	(-0.39)
STR	**-2.009*****	0.529	0.540	0.119	0.949	0.248	**-1.525****	**0.720***
	(-3.84)	(0.83)	(1.29)	(0.22)	(1.56)	(0.20)	**(-2.72)**	**(1.83)**
AI and Machine Learning	-0.049	0.007		**0.056****	0.019	-0.017	-0.015	-0.030
	(-1.05)	(0.19)		**(2.73)**	(0.37)	(-0.32)	(-0.32)	(-0.93)
Robotic Process Automation	0.005	0.027	0.014		**-0.088****	**0.102****	0.014	-0.001
	(0.10)	(1.04)	(0.66)		**(-2.25)**	**(2.09)**	(0.42)	(-0.03)
% Staff – Audit Function	**-0.461****	**-0.468*****	-0.012	0.139		0.081	-0.119	-0.154
	(-2.49)	**(-4.43)**	(-0.06)	(1.05)		(0.32)	(-0.59)	(-0.68)
% Staff – Large Taxpayer	1.684	**-3.346****	**2.020*****	0.464	-0.551		-0.316	-0.207
	(1.08)	**(-2.50)**	**(2.77)**	(0.40)	(-0.50)		(-0.35)	(-0.24)
Control of Corruption	0.098	-0.049	-0.189	**-0.408*****	**-0.614***	-0.152	-0.048	-0.086
	(0.68)	(-0.26)	(-1.00)	**(-3.76)**	**(-2.03)**	(-0.66)	(-0.18)	(-0.44)
% Staff Departing	-0.046	0.024	-0.285	-0.444	-0.009	1.057		-0.327
	(-0.07)	(0.03)	(-0.37)	(-1.27)	(-0.01)	(1.22)		(-0.57)
Staff Tenure	**0.020***	**0.023****	-0.005	0.011	-0.011	-0.021	0.002	
	(1.77)	**(2.05)**	(-0.48)	(0.85)	(-0.75)	(-1.18)	(0.15)	
Ln(GDP)	3.327	0.433	-1.187	-1.026	-2.101	3.145	1.310	-1.353
	(1.54)	(0.26)	(-1.00)	(-0.88)	(-0.83)	(1.37)	(0.70)	(-0.81)
Ln(GDP per Capita)	-2.694	0.236	**1.919***	0.519	0.522	-2.871	-2.220	0.463
	(-1.25)	(0.15)	**(1.76)**	(0.42)	(0.25)	(-1.19)	(-1.09)	(0.34)
Observations	102	102	102	102	102	102	102	102
R-squared	0.841	0.542	0.714	0.801	0.786	0.853	0.680	0.882

This table presents the results of independent regressions with the absolute value of the interpretable machine learning values (abs. iML) as dependent variables, and the feature values as the independent variables.. ***, **, and * represent significance at the 1%, 5%, and 10% level, respectively, based on t-statistics clustered by country. Country-year fixed effects are included in all specifications. For convenience all coefficients are multiplied by 100, and coefficients that are statistically significant are in bold.

Table 6
How Do Tax Enforcement Features' Effectiveness Differ for Pillar 2 vs. Non-Pillar 2 Firms?

Panel A: Tax Enforcement Budget

	1	2	3	4
	GAAP ETR - Pillar 2		***GAAP ETR - Not Pillar 2***	
VARIABLES	*Abs. iML Tax Expenditure / GDP*	*Abs. iML % Change Enforce*	*Abs. iML Tax Expenditure / GDP*	*Abs. iML % Change Enforce*
Tax Expenditure / GDP		0.120		**2.938****
		(0.17)		**(2.58)**
GDP Growth	-0.008	-0.014	-0.005	-0.010
	(-0.67)	(-0.78)	(-0.82)	(-0.42)
% Change Enforce	0.068		-0.034	
	(1.09)		(-0.79)	
STR	**-2.193*****	-0.724	-0.353	-0.066
	(-4.52)	(-0.57)	(-0.60)	(-0.04)
AI and Machine Learning	-0.072	-0.025	-0.009	-0.013
	(-1.39)	(-0.44)	(-0.42)	(-0.17)
Robotic Process Automation	0.023	**0.079***	-0.007	-0.020
	(0.44)	**(2.00)**	(-0.26)	(-0.45)
% Staff – Audit Function	**-0.758*****	**-0.670*****	-0.087	0.310
	(-3.28)	**(-3.63)**	(-0.72)	(0.76)
% Staff – Large Taxpayer	1.987	**-5.512***	-0.086	2.497
	(0.98)	**(-1.73)**	(-0.11)	(1.39)
Control of Corruption	0.099	-0.053	0.045	0.077
	(0.54)	(-0.22)	(0.58)	(0.26)
% Staff Departing	0.616	0.139	0.117	-0.320
	(0.89)	(0.14)	(0.25)	(-0.29)
Staff Tenure	**0.023***	**0.051*****	-0.000	-0.020
	(1.90)	**(3.39)**	(-0.02)	(-0.73)
Ln(GDP)	**4.055***	2.338	1.332	-4.956
	(1.75)	(0.94)	(1.06)	(-1.41)
Ln(GDP per Capita)	-3.720	-0.660	-0.939	4.219
	(-1.60)	(-0.27)	(-0.71)	(1.07)
Observations	102	102	102	102
R-squared	0.824	0.582	0.734	0.454

Table 6 (continued)
How Do Tax Enforcement Features' Effectiveness Differ for Pillar 2 vs. Non-Pillar 2 Firms?

Panel B: Tax Enforcement Technology Features

	1	2	3	4
	GAAP ETR - Pillar 2		***GAAP ETR - Not Pillar 2***	
VARIABLES	*Abs. iML AI and Machine Learning*	*Abs. iML Robotic Process Automation*	*Abs. iML AI and Machine Learning*	*Abs. iML Robotic Process Automation*
Tax Expenditure / GDP	**1.771****	-0.506	-0.289	-0.079
	(2.21)	(-0.89)	(-1.08)	(-0.13)
GDP Growth	0.002	0.004	0.003	-0.003
	(0.20)	(0.61)	(0.66)	(-0.65)
% Change Enforce	-0.006	-0.045	0.021	-0.074
	(-0.12)	(-0.76)	(0.81)	(-1.69)
STR	0.453	0.079	0.123	**1.776*****
	(0.89)	(0.13)	(0.50)	**(2.99)**
AI and Machine Learning		**0.054****		**0.124*****
		(2.36)		**(3.78)**
Robotic Process Automation	**0.051***		0.021	
	(2.02)		(1.34)	
% Staff – Audit Function	-0.096	0.113	-0.024	-0.181
	(-0.32)	(0.78)	(-0.31)	(-0.84)
% Staff – Large Taxpayer	1.562	0.119	0.042	1.580
	(1.66)	(0.08)	(0.04)	(1.41)
Control of Corruption	-0.070	**-0.460*****	-0.133	**-0.228***
	(-0.33)	**(-3.89)**	(-1.02)	**(-1.70)**
%Staff Departing	-0.250	-0.442	**0.656***	-0.720
	(-0.29)	(-1.06)	**(2.01)**	(-1.39)
Staff Tenure	-0.004	0.014	-0.012	0.010
	(-0.28)	(0.99)	(-1.43)	(0.79)
Ln(GDP)	-1.684	-0.495	-0.485	-2.222
	(-1.08)	(-0.39)	(-0.58)	(-1.41)
Ln(GDP per Capita)	1.969	0.018	0.258	2.664
	(1.22)	(0.01)	(0.33)	(1.58)
Observations	102	102	102	102
R-squared	0.727	0.793	0.823	0.814

Table 6 (cont'd) - How Do Tax Enforcement Features' Effectiveness Differ for Pillar 2 vs. Non Pillar 2 Firms
Panel C: Tax Enforcement Human Capital Features

	1	2	3	4	5	6	7	8
		GAAP ETR - Pillar 2				***GAAP ETR - Not Pillar 2***		
VARIABLES	*Abs. iML* % Staff – Audit Function	*Abs. iML* % Staff – Large Taxpayer	*Abs. iML* % Staff Departing	*Abs. iML* Staff Tenure	*Abs. iML* % Staff – Audit Function	*Abs. iML* % Staff – Large Taxpayer	*Abs. iML* % Staff Departing	*Abs. iML* Staff Tenure
Tax Expenditure / GDP	-0.563	-1.626	**2.481****	0.261	-0.221	0.081	**-0.725***	-0.264
	(-0.57)	(-1.51)	**(2.55)**	(0.42)	(-0.09)	(0.33)	**(-1.78)**	(-0.46)
GDP Growth	-0.010	-0.003	-0.023	**-0.030***	-0.022	0.002	0.007	-0.004
	(-0.64)	(-0.11)	(-1.27)	**(-1.70)**	(-1.07)	(0.24)	(1.13)	(-0.30)
% Change Enforce	-0.099	**0.299****	**0.216****	0.005	-0.191	0.027	0.080	-0.032
	(-0.81)	**(2.35)**	**(2.08)**	(0.07)	(-0.87)	(0.76)	(1.32)	(-0.53)
STR	0.896	0.210	**-1.484***	**1.023****	0.684	0.517	-0.129	1.279
	(1.35)	(0.14)	**(-1.97)**	**(2.05)**	(0.55)	(1.18)	(-0.22)	(1.64)
AI and Machine Learning	0.002	-0.059	-0.044	-0.041	-0.030	0.029	0.033	**-0.055***
	(0.03)	(-0.90)	(-0.78)	(-0.97)	(-0.50)	(1.30)	(1.08)	**(-1.75)**
Robotic Process Automation	**-0.083***	**0.143****	0.020	0.020	0.079	-0.004	0.004	**0.107*****
	(-1.80)	**(2.24)**	(0.51)	(0.41)	(1.61)	(-0.17)	(0.15)	**(2.94)**
% Staff – Audit Function		-0.181	-0.314	-0.130		-0.001	**0.259****	0.305
		(-0.57)	(-1.01)	(-0.38)		(-0.01)	**(2.38)**	(1.37)
% Staff – Large Taxpayer	-1.352		-0.140	-1.663	**-5.919*****		-1.074	**-1.761***
	(-1.05)		(-0.12)	(-1.38)	**(-3.15)**		(-0.95)	**(-1.84)**
Control of Corruption	**-0.610****	-0.183	-0.115	-0.143	-0.228	0.145	0.126	**0.356***
	(-2.06)	(-0.58)	(-0.42)	(-0.53)	(-0.76)	(0.85)	(0.84)	**(1.87)**
% Staff Departing	0.224	1.235		0.465	0.822	**0.640***		-0.203
	(0.25)	(1.15)		(0.69)	(0.77)	**(1.97)**		(-0.36)
Staff Tenure	-0.009	-0.013	0.004		-0.001	0.001	**-0.017***	
	(-0.56)	(-0.55)	(0.29)		(-0.07)	(0.10)	**(-1.71)**	
Ln(GDP)	-1.325	3.166	2.400	-0.724	-0.003	-0.526	-0.677	-0.164
	(-0.48)	(1.09)	(1.06)	(-0.36)	(-0.00)	(-0.54)	(-0.65)	(-0.09)
Ln(GDP per Capita)	-0.176	-3.499	-3.379	-0.516	-1.458	1.007	0.122	-0.629
	(-0.07)	(-1.18)	(-1.34)	(-0.32)	(-0.46)	(1.20)	(0.11)	(-0.41)
Observations	102	102	102	102	102	102	102	102
R-squared	0.782	0.846	0.697	0.860	0.826	0.878	0.765	0.855

Table 6 (cont'd) - How Do Tax Enforcement Features' Effectiveness Differ for Pillar 2 vs. Non-Pillar 2 Firms

This table presents the results of independent regressions with the absolute value of the interpretable machine learning values *(abs. iML)* as dependent variables, and the feature values as the independent variables. Country-year fixed effects are included in all specifications. We analyze subsamples of firms subject to the upcoming global minimum tax under Pillar 2 of the OECD/G20 Base Erosion and Profit Shifting plan (Columns 1 through 4) separately from those not subject tax (Columns 5 through 8) ***, **, and * represent significance at the 1%, 5%, and 10% level, respectively, based on t-statistics clustered by country. For convenience all coefficients are multiplied by 100.

Table 7: How Do Tax Enforcement Features' Effectiveness Differ for High vs. Low Intangible Firms

Panel A: Tax Enforcement Budget Features

VARIABLES	1	2	3	4
	GAAP ETR – High Intan		***GAAP ETR – Low Intan***	
	Abs. iML Tax Expenditure / GDP	*Abs. iML* % Change Enforce	*Abs. iML* Tax Expenditure / GDP	*Abs. iML* % Change Enforce
Tax Expenditure / GDP		1.125		-0.058
		(1.59)		(-0.05)
GDP Growth	0.005	-0.023	-0.039	**-0.046*** **
	(0.55)	(-1.40)	(-1.37)	**(-2.87)**
% Change Enforce	-0.015		0.025	
	(-0.49)		(0.20)	
STR	-0.553	**-1.649***	**-2.340***	1.046
	(-1.50)	**(-1.79)**	**(-1.91)**	(1.18)
AI and Machine Learning	0.012	-0.057	**-0.167***	0.016
	(0.42)	(-1.01)	**(-1.72)**	(0.41)
Robotic Process Automation	0.010	**0.122*** **	0.031	0.033
	(0.42)	**(2.95)**	(0.33)	(1.03)
% Staff – Audit Function	-0.050	-0.206	-0.059	**0.511** **
	(-0.17)	(-1.11)	(-0.17)	**(2.10)**
% Staff – Large Taxpayer	2.480	-3.848	3.146	-0.087
	(1.12)	(-1.43)	(1.18)	(-0.03)
Control of Corruption	-0.008	0.020	0.405	-0.089
	(-0.07)	(0.08)	(1.11)	(-0.42)
% Staff Departing	1.287	**1.962** **	1.042	1.539
	(1.36)	**(2.16)**	(0.61)	(1.60)
Staff Tenure	-0.017	**0.044*** **	-0.014	**-0.047** **
	(-1.43)	**(2.85)**	(-0.54)	**(-2.36)**
Ln(GDP)	-0.441	1.467	1.155	0.574
	(-0.30)	(0.55)	(0.24)	(0.28)
Ln(GDP per Capita)	0.854	-0.519	-2.148	-0.138
	(0.59)	(-0.20)	(-0.46)	(-0.08)
Observations	102	102	102	102
R-squared	0.800	0.632	0.782	0.592

Table 7 (continued): How Do Tax Enforcement Features' Effectiveness Differ for High vs. Low Intangible Firms

Panel B: Tax Enforcement Technology Features

	1	2	3	4
	GAAP ETR – High Intan		***GAAP ETR – Low Intan***	
VARIABLES	*Abs. iML* AI and Machine Learning	*Abs. iML* Robotic Process Automation	*Abs. iML* AI and Machine Learning	*Abs. iML* Robotic Process Automation
Tax Expenditure / GDP	**2.275****	0.055	-0.329	**-1.732****
	(2.41)	(0.15)	(-0.98)	**(-2.28)**
GDP Growth	0.002	0.009	-0.003	0.002
	(0.15)	(1.64)	(-0.27)	(0.14)
% Change Enforce	0.039	**-0.050****	-0.057	0.090
	(0.65)	**(-2.24)**	(-0.84)	(1.30)
STR	0.019	0.128	0.600	0.278
	(0.03)	(0.33)	(0.62)	(0.29)
AI and Machine Learning		**-0.031****		0.023
		(-2.15)		(0.57)
Robotic Process Automation	**0.129*****		**-0.080*****	
	(4.87)		**(-3.48)**	
% Staff – Audit Function	0.236	**0.310*****	-0.256	-0.091
	(0.78)	**(3.80)**	(-1.05)	(-0.49)
% Staff – Large Taxpayer	-0.176	0.453	**3.425****	**-3.222***
	(-0.14)	(0.97)	**(2.62)**	**(-1.78)**
Control of Corruption	**0.281***	0.054	**-0.676****	0.149
	(1.78)	(0.60)	**(-2.20)**	(0.40)
% Staff Departing	0.160	-0.433	-0.392	**-1.568***
	(0.20)	(-1.56)	(-0.40)	**(-1.95)**
Staff Tenure	-0.003	0.000	-0.004	-0.000
	(-0.16)	(0.07)	(-0.23)	(-0.02)
Ln(GDP)	-0.736	0.094	-0.430	4.675
	(-0.35)	(0.09)	(-0.21)	(1.65)
Ln(GDP per Capita)	0.443	-0.001	2.491	-3.630
	(0.19)	(-0.00)	(1.42)	(-1.44)
Observations	102	102	102	102
R-squared	0.807	0.932	0.841	0.854

Table 7 (continued): How Do Tax Enforcement Features' Effectiveness Differ for High vs. Low Intangible Firms
Panel C: Tax Enforcement Human Capital Features

	1	2	3	4	5	6	7	8
	GAAP ETR – High Intan				***GAAP ETR – Low Intan***			
VARIABLES	*Abs. iML* % Staff – Audit Function	*Abs. iML* % Staff – Large Taxpayer	*Abs. iML* % Staff Departing	*Abs. iML* Staff Tenure	*Abs. iML* % Staff – Audit Function	*Abs. iML* % Staff – Large Taxpayer	*Abs. iML* % Staff Departing	*Abs. iML* Staff Tenure
Tax Expenditure / GDP	-0.925	0.012	**4.085*****	0.219	-1.380	**-2.205***	-0.438	-0.200
	(-1.09)	(0.02)	**(5.89)**	(0.34)	(-0.80)	**(-1.83)**	(-0.33)	(-0.48)
GDP Growth	-0.005	-0.005	-0.011	0.030	-0.019	-0.012	0.006	0.000
	(-0.59)	(-0.38)	(-0.50)	(1.45)	(-1.18)	(-0.45)	(0.42)	(0.05)
% Change Enforce	**-0.108***	**0.211*****	-0.060	-0.004	-0.189	0.213	-0.063	0.033
	(-1.91)	**(3.12)**	(-0.50)	(-0.04)	(-1.13)	(1.67)	(-0.96)	(0.87)
STR	0.178	-1.027	**-2.790*****	0.732	0.473	0.124	0.146	**-1.318****
	(0.36)	(-1.66)	**(-3.20)**	(1.08)	(0.32)	(0.08)	(0.15)	**(-2.43)**
AI and Machine Learning	0.021	**-0.083****	-0.017	**-0.103*****	-0.060	0.005	0.071	**-0.068****
	(0.69)	**(-2.72)**	(-0.34)	**(-3.05)**	(-1.08)	(0.07)	(1.35)	**(-2.51)**
Robotic Process Automation	0.036	**0.105*****	0.028	0.011	0.077	**0.132***	0.015	**0.089*****
	(1.00)	**(3.52)**	(0.57)	(0.24)	(1.66)	**(1.99)**	(0.30)	**(5.64)**
% Staff – Audit Function		-0.288	0.226	-0.059		0.462	**0.982****	**0.258****
		(-1.06)	(0.67)	(-0.23)		(1.09)	**(2.70)**	**(2.20)**
% Staff – Large Taxpayer	**-2.142****		-0.121	-0.368	**-5.353*****		**-5.586*****	**-1.758****
	(-2.11)		(-0.05)	(-0.21)	**(-2.93)**		**(-2.77)**	**(-2.38)**
Control of Corruption	-0.036	0.196	-0.238	0.227	-0.333	0.045	0.001	**0.278****
	(-0.22)	(1.06)	(-0.96)	(0.66)	(-1.21)	(0.16)	(0.01)	**(2.34)**
% Staff Departing	0.376	**2.150*****		-1.599	1.304	**2.529***		0.614
	(0.69)	**(2.94)**		(-1.66)	(1.39)	**(1.95)**		(1.06)
Staff Tenure	**-0.028***	-0.013	-0.027		0.003	**-0.039***	-0.012	
	(-1.80)	(-1.01)	(-1.23)		(0.13)	**(-1.75)**	(-0.52)	
Ln(GDP)	0.080	**3.847***	0.562	-2.609	2.737	4.876	-1.586	0.865
	(0.05)	**(2.03)**	(0.19)	(-1.01)	(1.19)	(1.59)	(-0.66)	(0.69)
Ln(GDP per Capita)	-0.627	**-4.445****	-1.063	0.975	-1.822	-4.351	1.014	-1.257
	(-0.39)	**(-2.35)**	(-0.36)	(0.51)	(-0.80)	(-1.30)	(0.42)	(-1.06)
Observations	102	102	102	102	102	102	102	102
R-squared	0.658	0.832	0.680	0.834	0.823	0.886	0.712	0.894

Table 7 (continued): How Do Tax Enforcement Features' Effectiveness Differ for High vs. Low Intangible Firms

This table presents the results of independent regressions with the absolute value of the interpretable machine learning values *(abs. iML)* as dependent variables, and the feature values as the independent variables. ***, **, and * represent significance at the 1%, 5%, and 10% level, respectively, based on t-statistics clustered by country. For convenience all coefficients are multiplied by 100. *GAAP ETR High Intangible* represents the asset-weighted country-year *GAAP ETR* for High Intangible firms, while *GAAP ETR Low Intangible* represents the asset-weighted country-year *GAAP ETR* for Low Intangible firms.