Political Economy of Cross-Border Income Shifting: A Protection Racket Approach*

Maxim Ananyev

September 2018

Abstract

Multinational firms often shift their incomes to low-tax jurisdictions, thus robbing host states of tax revenue. I offer a new theory to explain why some firms engage in such activity while others do not. I argue that firms that are more vulnerable to expropriation by the government are less likely to engage in income shifting, since complying fully with the tax law makes firms more valuable to the government. I explore the empirical implications of my theory using a registry-based data set that allows for tracing connections between firms and tax havens. I set up a Bayesian multilevel model, allowing for random country-level and sector-level intercepts, and find that firms that – conditional on other firm-level covariates – have more concentrated fixed assets are less likely to utilize offshore havens. I explore a set of alternative explanations (reverse causality, statutory tax rate and political connections as omitted factors, among others) and carry out a sensitivity analysis to quantify the stability of my results with regard to potential omitted variable bias. My results challenge existing theories of the political economy of development. An influential "pillars of prosperity" theory asserts that successful states tend to simultaneously develop protection of property rights and fiscal capacity. My results imply that perfect property rights protection can actually harm fiscal capacity. In a world of frictionless international capital flows, some level of expropriation risk may be necessary for states to survive.

*I thank Leslie Johns and Daniel Treisman for reading and commenting on multiple iterations of this paper. I also thank John Ahlquist, Alma Bezares, Ashley Blum, Kirill Borusyak, Darin Christensen, John Christensen, Mikhail Galashin, Olga Gasparyan, Barbara Geddes, Sergei Guriev, Chad Hazlett, Natalia Lamberova, Yegor Lazarev, Jeff Lewis, Peter Lorentzen, Zsuzsanna Magyar, Tatiana Maiskaya, Natalia Naumenko, Eugenia Nazrullaeva, Imil Nurutdinov, John Nye, Francesca Parente, Joe Perry, Sergey Popov, Daniel Posner, Michael Poyker, Ron Rogowski, Arturas Rozenas, Arseny Samsonov, Anton Sobolev, Tim Frye, Ekaterina Volkova, John Zaller, seminar participants at UCLA, and conference participants at PolMeth 2016, SC^{2}PI 2017, HSE-EACES 2017, APSA 2017 for their excellent suggestions. I gratefully acknowledge the financial support from Columbia University Harriman Institute, George Mason University Institute for Humane Studies, UCLA Institute for European and Russian Studies, and UCLA International Institute. Correspondence: ananyev@ucla.edu

1. Introduction

The ability of the state to collect taxes (fiscal capacity) is an important determinant of economic development (Dincecco and Prado [2012]), technological progress (Acemoglu et al. [2016]), redistribution (Scheve and Stasavage [2010]) and is one of the plausible preconditions for democracy and accountability (Fortin [2012], Gottlieb and Hollenbach [2018]). While many rich states rely mostly on individual taxes to fund their policies and infrastructural projects, many poor and middle-income countries depend on corporate taxes. Thus, one of the biggest threats to fiscal capacity in many countries is corporate tax avoidance through tax havens (Palan et al. [2013]). According to the most reliable recent calculations, 40 percent of multinational corporations engage in tax avoidance leading to an annual loss of 10 percent of global taxes (Tørsløv et al. [2018]). This tax avoidance robs states of revenue and damages their ability to provide public goods and services.

One of the puzzling aspects of corporate tax avoidance is its within-country variation: Some firms are more aggressive in shifting their profits than the others (Dyreng et al. [2008]). Better understanding of the reasons behind the firms (non-)compliance is necessary to help poor and middle-income states build fiscal capacity, maintain order, and implement effective policies. In this paper, I attempt to explain this variation by looking at companies' infrastructural vulnerability to expropriation. I argue – both theoretically and empirically – that firms that are more vulnerable to expropriation (due to, for example, having more fixed assets in the country), will show more tax compliance.

My theory is inspired by Charles Tilly's famous assertion that the state operates as a protection racket (Tilly [1985]). I start with an assumption that when deciding whether to expropriate a firm or not, the government considers a cost-benefit analysis, where the cost of expropriation is the stream of tax revenue from a firm. Regardless of the potential benefits of expropriation, the government loses the taxes the firm would have paid. It follows immediately that, given the non-zero expropri-

ation risk, higher levels of tax compliance would decrease the probability of expropriation; therefore, companies are motivated to pay more taxes as a way to shield themselves from expropriation. If the expropriation risk is zero, this motivation is absent: Firms face no consequences of tax avoidance. If a state cannot enforce taxes through its legal system, an implicit threat of "lawless" expropriation might motivate the firm to pay more taxes.

Using a formal model, I derive the main testable implication of my theory: a firm's tax compliance should be inversely correlated with their concentration of assets, since companies behave strategically and anticipate the actions of the government. In other words, if a firm has more assets in a single country or in a small number of countries, it has more to lose from expropriation and, thus, should be more tax compliant. Also, firms that come from more vertically integrated sectors should be more tax compliant because a single act of expropriation does more damage for integrated firms than for non-integrated firms.

I explore the empirical implications of my theory using a proprietary Orbis data set that contains a registry-based cross-section of firms, as of the year 2014, from 77 countries and 699 four-digit industry codes. These data allow me to control for a variety of firm-level covariates (total assets, total revenue, industry, and number of foreign branches). Most importantly, the data set contains an identifier of the ultimate parent firm in the multinational corporate structure. This allows me to observe whether two firms in different countries belong to one ultimate parent, and, specifically, to observe whether a firm has a branch in a tax haven.

To test my theory, I estimate a firm-level logistical regression, where the indicator for the presence of a tax-haven affiliate is a dependent variable and the concentration of fixed assets is the main explanatory variable.[1] I control for the number of multinational branches, the total size of assets and the revenue of the firm. I also allow for country- and sector-specific intercepts. My main specification is a

[1]The terms "dependent" and "explanatory" are used only for descriptive convenience, since in my theory firms simultaneously decide on the amount of investment and on the extent of tax avoidance.

Bayesian multilevel model with three sets of random intercepts that allow for different country-level and sector-level baseline effects, while also achieving efficiency by pooling information across units. I also confirm that my results remain the same if I use the conventional fixed effects estimates.

Because neither the concentration of fixed assets nor the presence of a tax-haven branch is randomly assigned, my statistical procedure relies on the "selection on observables" assumption. To alleviate threats to validity, I explore a set of alternative explanations. First, I look into the issue of reverse causality: higher tax avoidance leads to lower fixed-assets concentration because of higher after-tax profits. I control for global after-tax profits and find that my results still hold. Second, if each country sets different tax rates for different sectors and those tax rates correlate with tax avoidance and with fixed-assets concentrations, this might bias my results. To alleviate this concern, I control for country-sector fixed effects, allowing the identification to be driven by within country-sector variation. Third, because the presence of a tax-haven branch might be a noisy measure, I confirm that my results hold if I use a tax-to-revenue ratio as a proxy for tax compliance.

Still, the omitted variable bias (OVB) can pose a challenge, so I perform a sensitivity analysis to quantify how large the OVB should be in my regression to nullify the results. Using the procedure proposed by Cinelli and Hazlett [2018] (which follows from the OVB formula for linear regression), I find that the effect of an omitted variable needed to nullify my results should be from six to ten times larger than the effect of the most consequential covariate in my specification. Thus, although it is impossible to rule out the OVB completely, my results are relatively stable with respect to omitted variables.

My findings challenge some existing theories on the political economy of economic development. Besley and Persson's influential "pillars of prosperity" theory (Besley and Persson [2011]) asserts that successful states simultaneously develop "legal capacity" (protection of property rights and political constraints) and fiscal capacity. In this view, these two capacities reinforce one another. My theory implies

a more complex relationship between fiscal capacity and property rights protection: Because firms pay taxes to increase the opportunity costs of expropriation, perfect property rights protection actually *harms* fiscal capacity. While in many countries fiscal capacity correlates with democracy, economic development, and property rights protection, it is not clear that this positive relationship is dynamically sustained. In fact, my theory offers an argument about why, at some stages of economic development, an optimal level of expropriation risk might be larger than zero.

One of the examples where expropriation threats have ostensibly helped improve tax compliance of firms is Russia's "Yukos Affair." In 2003, the richest person in Russia, oil tycoon Mikhail Khodorkovsky, was arrested and sentenced to eight years in prison, while his main asset, an oil refinery in a Siberian town of Nefteyugansk, was expropriated and merged with state-owned oil company Rosneft. These set of events triggered nearly universal condemnation from independent media, Putin's liberal opposition, and the West. While few observers denied that Khodorkovsky's firm Yukos had used clever tax optimization schemes and had exploited weaknesses in the tax code and imperfect tax enforcement, so too had other private firms that did not face punishment (Shleifer and Treisman [2001]). The arrest and sentencing of Khodorkovsky were widely regarded as a selective application of justice. Nevertheless, after Khodorkovsky's arrest, other private oil firms changed their tax behavior rapidly. Figure A1 in the Appendix shows that before the "Yukos Affair," the effective tax rate was around 20 percent, but in 2003 it jumped up to 29 percent (which cannot be explained by changes in oil prices or Russia's tax law) and remained high until oil prices collapsed in 2009. While the treatment of Yukos was unfair and violated international law, it did contribute to the buildup of Russia's fiscal capacity (Taylor [2011]).

My research contributes to several strands of literature. Most importantly, it contributes to a study of fiscal capacity. In this area, Li [2006] argued that authoritarian states often offered tax incentives to firms to compensate for the weak rule of law.

Gehlbach [2008] showed that post-communist governments in Russia tended to actively promote businesses that were easier to tax. Fairfield [2015] demonstrated, on a selection of case studies from Latin America, that businesses can use their infrastructural power (a threat to divest) to shape tax policy in their favor. Hollenbach and Silva [2018], using data from Brazilian municipalities, showed that municipalities with higher levels of inequality collected less revenue in property taxes. Theoretically, many of these contributions follow the state capture framework that asserts that economic elites are able, in one form or another, to shape redistributive outcomes in their favor (see, for example, Acemoglu et al. [2011] for an elaborate version of this theoretical argument).

My approach differs from many of these contributions. Instead of focusing on tax rates or the total amount of collected taxes, I look at tax compliance at firm level. This approach allows for developing and testing more nuanced theories and offers richer data. In this sense, the closest research to mine is a recent paper by Chen and Hollenbach [2018]. This paper uses a data set of 500,000 Chinese firms and shows that firms with higher capital mobility face higher tax rates and are less likely to use tax exemptions. They suggest an explanation: Firms that relocate more often have less time to acquire the political connections required for a tax exemption. The authors conclude that more research is needed on firm-level tax compliance.

The topic of firm-level tax compliance spans several disciplines in the social sciences. From an economic perspective, Durnev and Guriev [2011] argue that oil firms are more likely to engage in non-transparent financial accounting when oil prices are high. Beck et al. [2014] show that firms in countries with better information-sharing infrastructure are less likely to evade taxes. From an accounting perspective, Dyreng et al. [2008] demonstrate that there is significant variation among firms' tax compliance, Khan et al. [2016] argue that institutional ownership leads to more tax avoidance, and Huizinga and Laeven [2008] show that international tax differentials play an important role in profit-shifting.

My approach is also relevant to the classical literature on obsolescing bargain

(Jenkins [1986], Kobrin [1987]) which argues that once a multinational firm makes some tangible investment in a country, the government of that country acquires more bargaining power. My approach is consistent with the view that a certain bargaining – either overt or tacit – happens between the firm and the host government. Unlike the theorists on obsolescing bargain, I do not assume that it is the governments goal to get a larger share of gains from the firm. The goal of the government can be any: It can attack a certain firm for ideological reasons or send a signal to domestic audiences (as in the model by Acemoglu et al. [2013]). My approach does not assume either benevolence or malicious intent on the government's side. The only assertion I make is that, whatever the goals of the government, the opportunity cost of the expropriation goes up if the firms pay more taxes.

My argument has important policy implications. Scholars and policy makers often view the protection of the property rights of international investors as one of the most important goals of international law. If a government increases taxes or introduces a new regulation, a multinational firm, in many cases, can sue the state and win a generous compensation (Franck [2007], Schultz and Dupont [2014], Wellhausen [2016]). Poor and middle-income states are often bullied into signing tax treaties or investment treaties that limit the set of actions that they can implement on foreign firms. A huge part of international diplomacy is devoted to the protection of broadly defined property rights (Wellhausen [2014]). If my argument is correct, then the additional protection that companies enjoy from courts, treaties, and diplomatic efforts might have the undesirable side effect of harming the tax capacity of host nations. The international effort to protect the property rights of international investors – however laudable it might be – should be balanced by the equally enthusiastic embracing of tax compliance.

2. Background: How Firms Avoid Taxes and Why It Matters for Governance

Countries differ in the tax rates they offer to firms. In 2018, the United Kingdom taxes firms at 19 percent, Spain at 25 percent, and the Bahamas at 0 percent. A zero tax rate is also offered by Anguilla, Bahrain, Bermuda, the Cayman Islands, Guernsey, and others. Such nations, to whom corporations pay zero or near zero, are commonly referred to as tax havens.[2][3] According to recent research, they play an increasingly important role in the global economy (Hampton and Abbott [1999], Shaxson [2011], Palan et al. [2013], Zucman [2015], Murphy [2017]).

These jurisdictions serve several purposes. First, they lower the transactions costs of international operations. In many cases, when firms decide on implementing a merger or an acquisition, they choose to do it using a neutral jurisdiction. The second function is providing secrecy to the financial holdings and operations of individuals and firms. These jurisdictions are usually reluctant to cooperate with other governments in revealing the true beneficiaries of the entities registered there and the true owners of bank accounts.[4]

Most importantly for the analysis of this paper, tax havens serve as vehicles for the minimization of tax burden on multinational corporations. If a firm earns revenue in a high-tax jurisdiction (for example, the UK) but also has an affiliate in a low-tax jurisdiction (for example, the Bahamas), then in many cases, it becomes possible to "pay" the Bahamas' tax rate of 0 percent on the revenue that has been

[2]Importantly, some countries that have a non-zero statutory tax rate offer zero or near-zero tax rates to foreign firms, while domestic firms are taxed at a higher rate. Examples include Ireland and Switzerland.

[3]Literature uses several terms for such countries, and none of those terms are ideal: tax havens (though they are not always used for tax avoidance), offshore jurisdictions, or just offshores (though not all of them are located offshore, most notably, landlocked Switzerland), offshore financial centers (though many of them do not have a financial sector, besides registering companies and selling synthetic citizenship). I will use the term "tax havens" since it is most the common one in non-technical literature; tax avoidance is the focus of this paper.

[4]Recently, G20 and OECD states attempted to enact new regulations designed to limit the secrecy provided by such jurisdictions. However, as Johannesen and Zucman [2014] demonstrate, those measures had limited success, only causing the reallocation of accounts into the least compliant jurisdictions.

earned in the UK. This can be achieved by several methods that often are not illegal, but occupy a gray area between tax evasion and full tax compliance (Palan et al. [2013]).

These methods include arranging a corporate multinational structure in such a way that an affiliate in a high-tax jurisdiction customarily pays huge sums of money to an affiliate in a low-tax jurisdiction. From the point of view of a high-tax jurisdiction, these are usually legitimate expenses that decrease the profit of a firm. But from the point of view of economic substance, these transactions leave the money within a multinational firm, and the only difference such transactions make is the adjustment of tax liability. Because this process reduces the taxable base in high-tax jurisdictions, it is sometimes referred to as "base erosion" (Dharmapala [2014]).

One of the most widely studied cases of tax avoidance through income-shifting is the set of techniques used by Starbucks UK (Hodge [2016]). First, Starbucks UK paid royalties for using the brand (6 percent of turnover) in another European Starbucks branch located in Netherlands (a low-tax jurisdiction).[5] Second, Starbucks UK had to buy expensive coffee beans from a Starbucks branch located in Switzerland (also a low-tax jurisdiction).[6] Third, Starbucks UK had to borrow from the Amsterdam branch and pay high interest rates. These payments had been conveniently consuming Starbucks UK's profit margin for many years. Such techniques – royalty payments, buying intermediate goods and services from low-tax affiliates, and paying high interest rates on intra-firm loans – are common in firms that engage in income-shifting (Palan et al. [2013]).

From the perspective of Pareto efficiency, this practice is neutral since the incomes are essentially redistributed from the budgets of high-tax jurisdictions to the shareholders of the firms. Nevertheless, this practice is often considered norma-

[5]Starbucks' European headquarters had been located in Amsterdam, though the president of Starbucks Europe had been running the operations from London. In 2013, the headquarters were officially relocated to London.

[6]"Buying from Switzerland," as Hodge [2016] explains, means just transferring money to a Swiss branch, since "no coffee bean ever reaches Switzerland."

tively undesirable for several reasons. First, it exacerbates inequality since only relatively rich firms can afford to create effective profit-shifting processes (Krautheim and Schmidt-Eisenlohr [2011]). Secondly, it damages the tax capacity of the host jurisdictions. This activity is even more damaging for poor and middle-income countries because their budgets are more reliant on the corporate taxes (Crivelli et al. [2015]).

The extent of profit-shifting is hard to measure, but several estimates are available in the literature. Baker [2005] estimated the yearly loss of revenue in the developing countries due to the mispricing of intra-firm transactions at a minimum of 200 billion dollars. According to valuations published by Christian Aid (Aid [2008]), the loss of revenue in the developing countries was 160 billion dollars.[7] Other estimates available in the literature show similar figures. Tax haven usage has been increasing almost uninterruptedly since the late 1990s (Zucman [2014]).

3. Protection Racket Approach: Theoretical Argument

This section offers a more detailed description of the theory of tax compliance proposed in this paper. The basic intuition is that the firms that pay more tax are more valuable to the government because they provide it with much-needed cash. Therefore, the government should treat these firms more carefully and be less willing to damage them. Thus, firms that feel more vulnerable because of the nature of their operation – such as relying on just one country for much of their income – must be more tax-compliant. In Section A.2 in the Appendix, I offer a straightforward game-theoretic formalization of the argument.[8] Here, I present an intuitive explanation of my theory.

The game consists of two players: the government and the host country. The firm decides how concentrated it wants to make its assets (what share to invest in

[7] A useful reference point to have is 1 trillion dollars: This was the total GDP of all Sub-Saharan Africa (in current US dollars) in 2008.

[8] The goal of the formal theory here is not to produce a methodological contribution but to introduce transparency into the assumptions.

the country) and how much tax to pay. The government has a nominal tax rate it offers to the firm but is unable to enforce legally. The firm decides on the level of its tax compliance (i.e., how much of the tax rate to actually pay): It can pay zero (full tax avoidance), the full tax rate (full tax compliance), or anything in between.

Once the investment is made, the government can expropriate the invested assets. If it decides to expropriate, it receives the value of the asset, potentially discounted because it might not be able to manage the asset as efficiently as the firm. It also might receive some political benefits from the expropriation: For example, an extreme left-wing or a populist government might win some support from its voters. A corrupt government that survives on patron client exchanges might win the loyalty of its cronies by redistributing the asset through the patron - client networks. It is also possible that the government will get in trouble if it expropriates the asset because some voters or elite members might not approve, or the international community might impose penalties on the government. The extent of the political benefits or losses from the expropriation is unknown to the firm when it decides whether to invest and whether to comply with taxation.

If the government decides to not expropriate the asset, it receives the taxes that the firm has decided to pay. If the firm has decided to pay zero, the government receives nothing. Therefore, the higher the tax compliance of the firm, the higher the opportunity costs of expropriation.

If the government decides to expropriate the asset, the firm loses the asset, and it might suffer additional damage proportional to the lost asset. This additional loss can be thought of as the impact of the domino effect on the rest of the firm's supply chain, which is larger the more vertically integrated the firm is. Firms that operate in more vertically integrated sectors suffer additional damage because their assets are scattered across the supply chain. For example, if an agricultural firm has 900 million dollars worth of land in different countries and the government expropriates 300 million worth of it, the firm still has 600 million worth of a productive asset. But if a firm has 300 million worth of land, 300 million worth of food-processing fa-

cilities, and 300 million worth of retail outlets that sell the food, the expropriation of any of these assets would severely damage the firm's capacity to generate profits beyond the value of the expropriated assets.

If the asset is not expropriated, the firm gets the asset back and also earns whatever remains after it has paid taxes.

It is easy to show that, under these assumptions, three implications should hold (formal proof is relegated to the Section A.2 in the Appendix).

Hypothesis 1. *Tax compliance is inversely related to the concentration of assets.*

Intuitively, if the firm decides to put more assets into a country, it should also decide to pay more taxes and engage in less tax avoidance. The result is straightforward: If the firm chooses to have more concentrated assets, it has more to lose if the government decides to expropriate the asset.

Hypothesis 2. *Tax compliance should be higher if a firm operates in a vertically integrated sector.*

The logic here is the same: The costlier expropriation is for the firm, the more motivated the firm is to be tax-compliant, since tax payments increase the opportunity costs of expropriation.

As the formal theory presented in Section A.2 in the Appendix shows, this framework can generate many other predictions: Specifically, one can derive correlations and higher-order interactions involving tax compliance, the tax rate, the quality of the government, and the country-specific rate of return. However, for the purpose of this paper, I focus mainly on Hypothesis 1 (Sections 5 and 6), also report the results of exploring Hypotheses 2 (Section 7), and leave explorations of other implications for further research.

4. Data and Descriptive Statistics

My main dependent variable is income shifting – the utilization of low tax jurisdictions to ease the burden of taxation in host countries. For this, I used a relatively coarse measure: an indicator variable for whether a firm has a branch in a jurisdiction commonly described as a "tax haven." This measure is not ideal because not all branches in tax havens serve the purpose of avoiding taxation: Some of them are used because those jurisdictions provide smaller transactions costs for cross-border deals and for other reasons. Nevertheless, it has been documented in the literature that the firms with affiliates in tax havens usually have higher profits and lower tax burdens (see, for example, Janský and Prats [2015]). As a robustness check, I used tax-to-revenue ratio as a dependent variable in an additional set of specifications (see Table 4)

To find out which firms have affiliates in tax havens, I used the proprietary database Orbis compiled by Bureau van Dijk. The Orbis database has financial information on public and private companies around the world. It has information on 54 million companies from North and Central America, 38 million from South America, 7 million from Africa, 43 million from Western Europe, 50 million from Eastern Europe and Asia, and 24 million from Australia.

Most importantly for the purpose of this research, Orbis contains a a global ultimate owner (GUO) in most firms. A GUO is an individual, government, or firm that is at the beginning of the hierarchy of ownership and has at least 51 percent stock of its subsidiary. For example, the database has information on the US technology firm Apple Inc. and 134 of its subsidiaries. Such firms as Apple Operations Europe (registered in Ireland), Apple France (registered in France), and Apple Italy (registered in Italy) have the US firm Apple Inc. as their GUO. This allows them to observe the existence of branches in "tax havens" for a large sample of firms in many countries. I used the dummy for the existence of such branches as the main

dependent variable in this study.[9]

I collected the dataset as follows. First, I compiled a list of jurisdictions classified as "tax havens" following a definition supplied by Palan et al. [2013]. I included all the jurisdictions that had been in the OECD Uncooperative Tax Havens list and augmented them with several jurisdictions that have not been in that list but are frequently mentioned in online lists of "world's best tax havens" compiled by journalists and consultants on tax optimization. As a result, I ended up with a list of 73 jurisdictions (reproduced in full in Section A.4 in the Appendix).

Then, for all other countries (countries that are not in the list of tax havens), I downloaded 1,000 firms that have the largest revenues in 2014 if a country had more than 1,000 firms in the database. I downloaded information in all those firms in a country that had fewer than one thousand firms. I used revenue as as a selection criteria because I wanted to compile a list of the largest potential taxpayers. As a result, I got 83,836 firms. I then downloaded the revenue, industry (Orbis uses a four digit indicator for industry), total assets, and fixed assets of each firm.

Then I removed all government-owned firms from the dataset (since my theory applies to private firms only) and collapsed the firms according to their GUO. I then calculated the total assets (some of the assets of the subsidiary firms), total revenue (sum of the revenues of all the subsidiary firms), and total fixed assets (sum of the fixed assets of all the subsidiary firms)of every GUO.

The most relevant measure for Hypothesis 1 is the concentration of fixed assets (intuitively, a share of the total easily expropriated assets, such as plants, real estate, land, and equipment, that the firm has in a country or in a small set of countries). To measure the asset concentration, I calculated the Herfindahl-Hirschman index (HHI):

$$FixedHHI_k = \sum_{i=1}^{N} r_{ik}^2. \tag{1}$$

[9] I used GUO as the unit of analysis because the profit shifting behavior is, in most cases, determined by the headquarters of the multinational firm (Dischinger et al. [2014]).

Here, $FixedHHI_k$ is a concentration of fixed assets (as measured by HHI) of GUO k and r_{ik} is a share of fixed assets of GUO k that comes from country i, and N is the number of countries in my dataset. Only fixed assets located in non-tax-haven countries are included in the procedure.

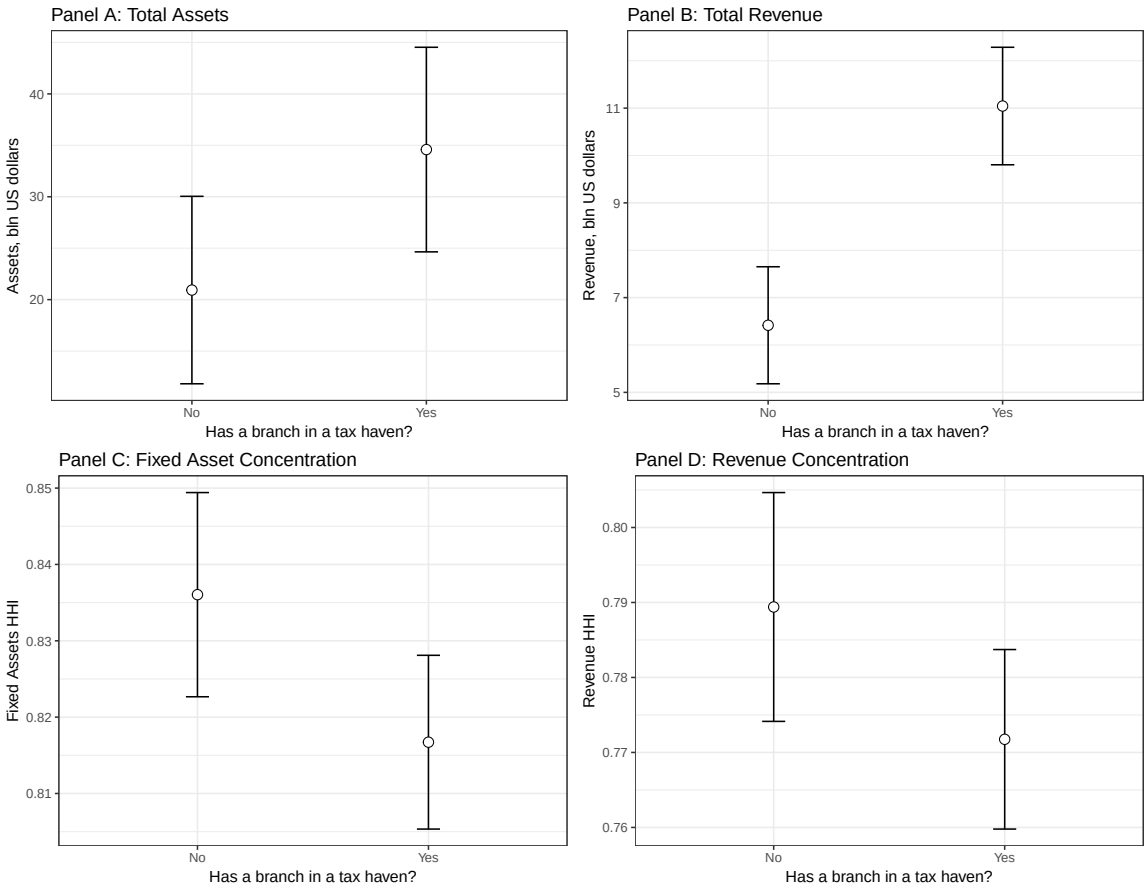
The HHI is widely used in studies to measure market concentration (Rhoades [1993]), ethnic homogeneity of countries (Mauro [1998],Roeder [2001]), and party competition (Laakso and Taagepera [1979]). It has a straightforward interpretation: If we randomly pick 2 dollars of fixed assets of a GUO, the HHI will show the probability of those 2 dollars are in the same country. For example, if a firm operates in one country only, then its HHI equals 1 (the maximum possible level of concentration). If the firm relies on two countries in equal proportion, then the HHI is equal to $\frac{1}{4}+\frac{1}{4}=\frac{1}{2}$. I expect the GUOs where firms are concentrated (whose HHI is higher) to be more vulnerable to government intervention and thus be less likely to use tax havens. In addition to concentration measured by fixed assets, I also calculate the HHI index measured by the sources of revenue.

To make sure that my inference is not driven by the difference between big and small firms, I removed from the dataset firms that have more than 70 billion dollars' revenue and those that have more than 26 international branches (which have more than 95 percent propensity to have an affiliate in a tax haven). After these procedures, I ended up with 6,985 firms. The descriptive characteristics of my sample can be found in Table A.1 in the Appendix. Also, Table A.2 shows the distribution of countries in my sample (the number of firms per country is roughly proportional to the economic development of a country), and Table A.3 in the Appendix shows the number of firms per economic sector.

Figure 1 shows the descriptive comparison of multinational firms that have known affiliates in tax havens with multinational firms that do not have known affiliates in tax havens (for a more informative comparison, I excluded for this figure firms that operated in one non-tax-haven country only).

Panels A and B of Figure 1 show, unsurprisingly, that firms that have affiliates

Figure 1: Firms that Have/Have Not Affiliates in Tax Havens



Note: Averages and 95-percent uncertainty intervals of certain characteristics of multinational firms in the sample. On a vertical axis, Panel A shows total assets in billions, Panel B shows total yearly revenue in billions, Panel C shows the HH index calculated using fixed assets, and Panel D shows the HH index calculated using revenue.

in tax havens are, on average, richer in terms of total assets and revenues even after the richest firms have been trimmed from the dataset. On average, firms that use tax havens boast around 35 billion dollars in assets, while those that do not use tax havens have only around 20 billion dollars in assets. Data on the revenues of firms show a similar pattern: the average revenue of a firm that does not use tax havens is around 6 billion dollars, while that of one that uses tax havens is 11 billion dollars.

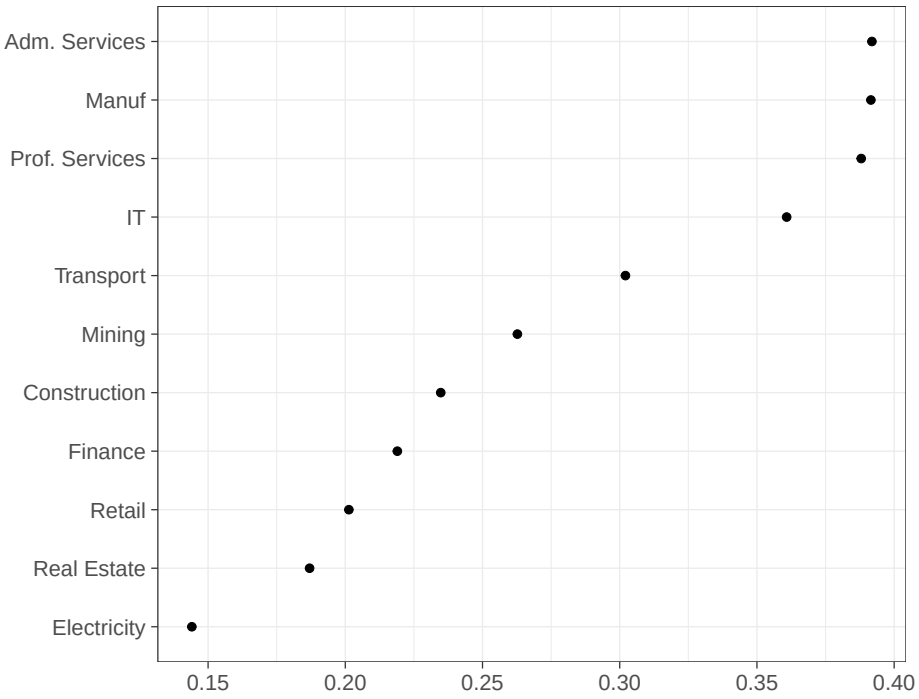
Panels C and D of Figure 1 show the differences in fixed asset concentrations and revenue concentrations among multinational firms. For both measures, we see patterns that are consistent with my theory: both fixed asset concentrations and revenue concentrations are larger among firms that do not use tax havens.

The rates of tax haven usage are also different among economic sectors. Figure 2 shows the shares of firms that use tax havens in different sectors. We see that the largest proportion of tax having usage is observed among firms that operate in the administrative services and manufacturing sectors, while the lowest rate of tax haven usage is observed among the real estate and electricity sectors.

The rates of tax haven utilization also vary across countries. Figure 3 shows the differences in the proportion of firms using tax havens in different nations. The largest proportion of tax- haven-using firms is observed in Guatemala, Dominican Republic, United Kingdom, and Kuwait. Latin American nations (such as Chile, Argentina, and Brazil) demonstrate relatively low rates of tax haven utilization, while Western European countries, the United States, Russia, and India have relatively high rates of tax haven utilization. [10]

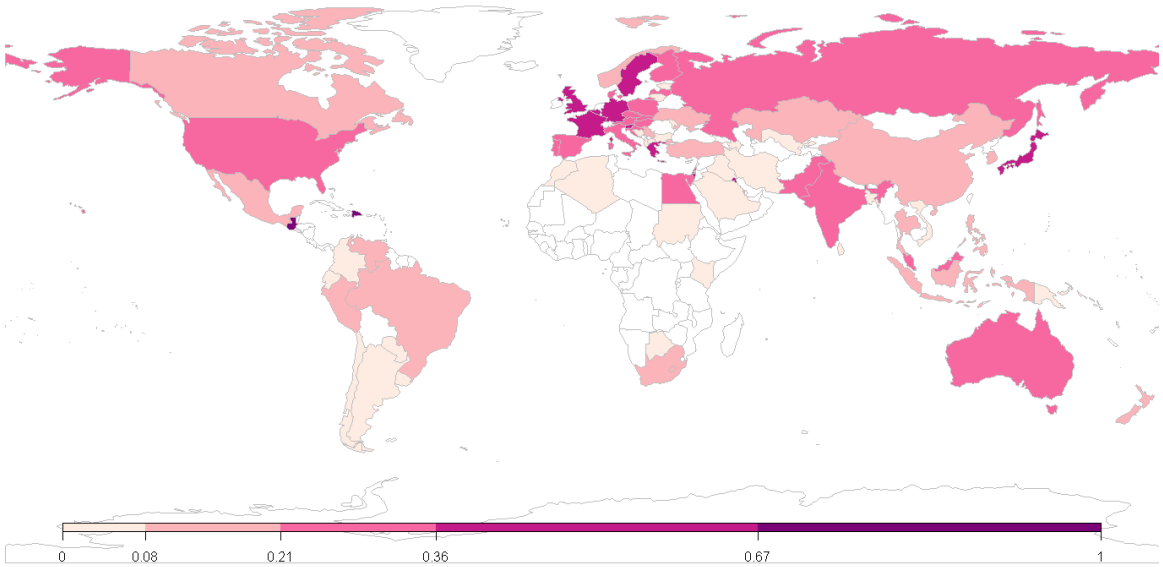
[10] Figure A2 in the Appendix shows some of the country-level correlates of income shifting.

Figure 2: Tax Haven Utilization Across Sectors



Note: Each line represents a share for firms that are connected to an affiliate in a tax haven among the firms of that primary sector. Data come from Orbis and the author's calculations.

Figure 3: Tax Haven Utilization Across Countries



Note: The map shows, for every country, the share of firms with the affiliates in the tax haven among the firms from that country. The thicker the color, the larger is the share. Data come from Orbis and the author's calculations.

5. Statistical Evidence

Multilevel Specification

The nature of the theory outlined earlier as well as the nature of the data invite the usage of Bayesian multilevel models for statistical analysis. While, by definition, multinational firms operate in more than one country, they still have "nationalities" (Wellhausen [2014]) and, in many ways, their behavior their "home" country. Also, it is reasonable to expect firms in different industries to behave differently. For these reasons, I use two non-overlapping levels in my statistical specification ("home" countries and industries), allowing different baseline values of tax compliance for each level.

Because the main dependent variable is binary, I start with a simple Bernoulli distribution, assuming that each observation $Haven_i$ (which equals 1 if a firm i uses tax havens, and 0 otherwise) is a draw from a Bernoulli distribution with an underlying (unobserved) probability p_i:

$$Haven_i \sim Bernoulli(p_i). \tag{2}$$

Probability p_i is an inverse logit function from a linear component that combines effects of firm-level covariates and three types of intercepts: global intercept (α^1), industry-specific ($\delta_{m[i]}$, where m is an index of an industry), and country-specific intercepts ($\gamma_{j[i]}$, where j indexes a country):

$$p_i = logit^{-1}(\alpha^1 + \beta_2^1 * FixedHHI_i + X_i'\gamma + \delta_{m[i]} + \theta_{j[i]}). \tag{3}$$

Here, $FixedHHi_i$ is the concentration of fixed assets of firm i, and X_i is a set of firm-level covariates: log total assets, number of branches, and log total revenue.

All the unobserved quantities are given non-informative (improper) priors so that only the data, and not prior expectations about their values, influence their

posterior distributions.

Equations 2 - 3 describe my preferred specification. However, to demonstrate that the specific choice of covariates does not drive my results, I fit several models where an increasingly rich set of covariates is added in a consecutive way. I estimated the posterior distribution of the coefficients using MCMC sampling in Stan (Team [2014]). For every coefficients, I ran four chains with 10 000 iterations, the first 5000 being treated as burn-in iterations and discarded from the calculation of the posterior density of the coefficients and other summaries.

Results

The results of the estimation are shown in Table 1. Models 1-4 present increasingly rich multilevel specifications. In Model I, I do not include any predictor, the logit coefficient on HHI Assets (concentration fixed assets) is -0.66, which corresponds to the odds ratio of 0.51. All the variables are standardized. Once I control for the number of log revenue and number of (non-tax-haven) branches, the main coefficient goes down in magnitude to -0.27 and remains the same with an addition of other controls.

Table 1: Main Results: Asset Concentration and Income Shifting

	Model 1	Model 2	Model 3	Model 4
HHI Assets	-0.66	-0.27	-0.27	-0.27
	$[-0.72;\ -0.59]$	$[-0.34;\ -0.19]$	$[-0.35;\ -0.20]$	$[-0.35;\ -0.19]$
Revenue		0.27	0.13	0.17
		[0.21; 0.33]	[0.05; 0.22]	[0.06; 0.27]
Branches		0.75	0.76	0.57
		[0.63; 0.90]	[0.63; 0.88]	[0.45; 0.70]
Assets			0.19	0.47
			[0.10; 0.27]	[0.36; 0.58]
Country Intercept				✓
Industry Intercept				✓
N.	6985	6985	6985	6985
n_{eff}/n	0.47	0.42	0.39	0.36

Note: Models 1-4 present different specifications to test the connection between asset concentration in tax haven usage. The unit of observation is a firm. The dependent variable is an indicator for having an affiliate in a tax haven. None of the standard convergence diagnostics indicate non-convergence.

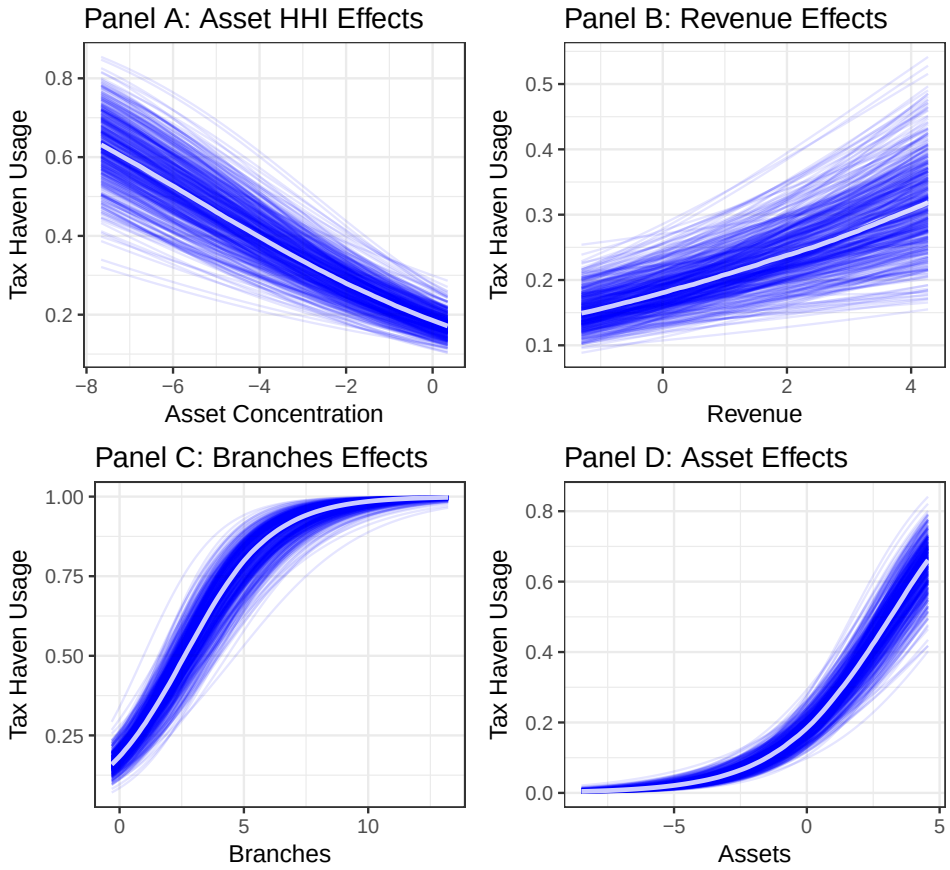
The logit coefficient of - 0.27 implies an odds ratio of 0.76. Given the baseline value of tax haven utilization of 0.28, the increase in one standard deviation of concentration of assets implies the decrease in the propensity to use a tax haven by 5.7 percentage points. Section A.7 in the Appendix reports the results of Bayesian hypothesis tests for the coefficients: For the substantive coefficients the hypothesis of the coefficient being sufficiently close to zero and substantively negligible (between -0.05 and 0.05 on logit scale) is rejected.

As far as other coefficients are concerned, firms that are larger (in terms of total revenue, total assets, number of branches) are more likely to have affiliates in the tax havens. This is understandable since establishing a multinational corporate structure that involves the affiliates in tax havens entails certain fixed costs and marginal costs (as De Simone et al. [2016] point out, "Efficient transfer pricing strategies can be expensive to put in place").

Figure 4 shows the marginal effects of asset concentration (Panel A) and other firm-level explanatory variables (Panels B - D). Once the asset concentration goes up from the smallest possible value to the largest possible one, the propensity to utilize

tax havens goes down from sixty percent on average to around twenty percent on average. Other explanatory variables (revenue, assets, and number of branches) are positively associated with tax haven utilization.[11]

Figure 4: Marginal Effects of Firm-Level Predictors



Note: The panels represent the effects for firm-level marginal effects (calculated using sampling from posterior distribution) from Model 4 in Table 1.

The substantively meaningful negative association between the asset concentration and the tax haven usage is consistent with my theory (in particular – with Hypothesis 1). But it can also be consistent with other theories. The next section considers some of those alternative explanations.

[11] Figure A3 in the Appendix presents plots of posterior distributions for all substantive coefficients. Figure A4 in the Appendix shows the estimated intercepts and their uncertainty intervals for all country-level intercepts. Section A.8 in the Appendix presents various Bayesian goodness-of-fit estimates of the main specification.

6. Alternative Explanations

The previous section established an empirical result: negative association between the tax haven utilization and fixed assets concentration. I interpret this result as an evidence in support of my theory: tax revenue from a particular firm increases an opportunity cost of expropriating that firm, and a strategic reaction of the firm is to pay more taxes if it chooses to have more concentrated fixed assets. However, this conditional correlation can also be potentially consistent with other theories. In this section, I consider several of such alternative explanations.

Tax Avoidance Helps Firms Diversify Their Assets and Sources of Revenue

Previously, I argued that once a firm chooses to have more concentrated assets, it should pay more taxes to try to shield itself from expropriation. This logic yields a negative correlation between concentration and tax avoidance. Another explanation is possible: firms that evade taxes get higher de-facto after-tax profits, and that revenue allows them to expand to diversify their assets and their sources of revenue.

This explanation is plausible. Indeed, firms that use tax havens have higher profits and pay less taxes (which is the primary reason for using tax havens). Unfortunately, limitations in the data prevent me from using a dynamic specification.[12] Nevertheless, because this mechanism operates through total after-tax profits (which are later used for expansion and diversification), the effect should be alleviated once after-tax profits are controlled for.

Figure 5 shows the causal diagram that illustrates the logic of this argument. Tax avoidance leads to higher global after-tax profit, and the higher global after-tax profits allow the firms to diversify into more countries thus leading to lower concentration. The diagram also shows how one might test this explanation em-

[12] While there are several years of firm-level financial data in the Orbis database I cannot reliably establish in which year a firm created an affiliate in a tax haven, since the transparency of tax havens was undergoing rapid changes at the same time.

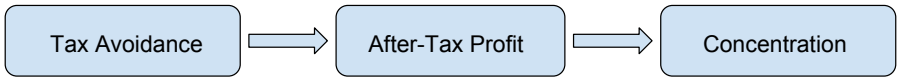
Table 2: Alternative Explanation I: Tax Avoidance Drives Concentration

	HHI Assets	HHI Revenue	HHI Assets	HHI Revenue
Profit	-0.07	-0.05	0.02	0.05
	$[-0.09;\ -0.04]$	$[-0.08;\ -0.03]$	$[-0.00;\ 0.05]$	$[0.02;\ 0.07]$
Haven	-0.68	-0.73	-0.28	-0.27
	$[-0.74;\ -0.63]$	$[-0.78;\ -0.68]$	$[-0.32;\ -0.23]$	$[-0.32;\ -0.22]$
Controls	✓	✓	✓	✓
Country Effects			✓	✓
Industry Effects			✓	✓
N	5738	5738	5738	5738
R^2	0.11	0.24	0.10	0.24
n_{eff}/n	0.67	0.66	0.27	0.26

Note: All models control for firm-level covariates: log total assets, log total revenue, and the number of (non-tax-haven) branches. Intercept is not shown. *Profit* is total log after-tax profit of a firm. *Haven* is a an indicator for whether a firm has a connection to tax haven. For every estimate shown, I put 95-percent HDI intervals in parenthesis. n_eff/n is convergence diagnostic that for every specification shows the smallest ratio of effective sample size to nominal sample size across all estimated coefficients (the ratio less than 0.1 indicates non-convergence). Models in columns (2) and (3) include random intercepts for industries and for countries.

pirically: because the effect of tax avoidance is fully mediated by after-tax profits, the concentration and tax avoidance should be independent once conditioned on after-tax profits.

Figure 5: Alternative Explanation I: Tax Avoidance Drives Concentration



Note: Causal diagram that shows a plausible causal connection between tax avoidance and concentration

Table 2 shows the regressions with concentration as a left-hand-side variable and tax avoidance (usage of tax havens) and global after-tax profit (including profits booked in tax havens, where available).[13]

Table 2 demonstrates that controlling for after-tax profits does not nullify the association between concentration and the usage of tax havens. Columns (1) and

[13]These specifications reverse the "order" of variables. Concentration is a "dependent" variable, and an indicator for tax haven usage is an "explanatory" variable. To evaluate conditional independence, the order of the variable is irrelevant. I kept the concentration as the dependent variable to match the causal explanation in Figure 5.

(2) demonstrate that once after-tax profits are added in a bivariate regression for the concentration of assets or concentration of revenue, the estimated association with tax haven usage remains negative and large in magnitude (on average, having an affiliate in tax haven is associated with a 0.7 standard deviation decrease in the concentrations of both assets and revenue). Columns (3) and (4) show a controlled multilevel specification with log assets, log revenue, number of branches, and random intercepts for industries in countries. The estimated coefficients of tax haven usage are still negative and nontrivial in magnitudes (usage of tax havens is associated with 0.3 smaller concentrations).

These estimates demonstrate that the causal model in Figure 5 might be wrong since there is still an association between the concentration of assets/revenue and tax haven utilization even when after-tax profits are controlled for.

Statutory Tax Burden as Omitted Variable

Another alternative explanation is that both concentration and tax haven usage depend on tax rates. If tax rates are high, then firms have an incentive to relocate parts of production and revenue-generating activities. In this case, tax rates lead to lower concentrations and to tax haven usage.

There are several ways one might test if the previously presented results are driven by tax rates. First, one might control for the country-level tax rate explicitly. Second, because tax rates might be measured with errors, one might control for country fixed effects (which include tax rates because those are fixed on the level of a country). Third, because tax rates might differ within a country across sectors, one might include country-sector fixed effects, so that the variation to be explained by the concentration comes from the country-industry cell.

Table 3 presents the results of the estimations. Here, I fit linear probability model since the main rationale for using fixed effects for identification is to control by factors that are common for the group and contribute to risk difference in an "additive and constant" way (Angrist and Pischke [2008]).

Table 3: Alternative Explanation II: Tax Burden as Omitted Variable

	Baseline	Model 1	Model 2	Model 3
HHI Assets	-0.12	-0.07	-0.06	-0.07
	$[-0.13;\ -0.11]$	$[-0.08;\ -0.06]$	$[-0.08;\ -0.05]$	$[-0.08;\ -0.05]$
Tax		-0.04		
		$[-0.05;\ -0.03]$		
GDP		0.00		
		$[-0.01;\ 0.01]$		
Polity2		0.04		
		$[0.03;\ 0.05]$		
Controls	✓	✓	✓	✓
Industry FE			✓	
Country FE			✓	
Country-Ind FE				✓
N	6985	6966	6985	6985

Note: The dependent variable in all specifications is an indicator for tax haven utilization. Column (1) shows the baseline specifications with firm-level controls but no country-or-industry covariates. Model 1 shows the results once country-level controls are added. Model 2 shows the results once country-level fixed effects and industry-level fixed effects are added, and Model 3 shows the results once country-industry fixed effects are added. The specification is fit using OLS with robust standard errors; 95-percent confidence interval is shown as an uncertainty estimate for all the coefficients. The regressions are estimated using `estimatr` package in `R`.

Table 3 presents the results. In all the specifications, one standard deviation increase in the concentration of assets or the concentration of revenue is associated with a four percentage point decrease in the probability of a firm using a tax haven. Therefore, the concentration of revenue/assets is associated with tax haven usage even when statutory tax rates and other country, sector, and potentially country-sector factors are accounted for.

Tax Havens Are Used to Lower Transactions Costs, not to Avoid Taxes

While tax havens are used primarily for tax avoidance, this is not their only usage. Tax havens are also used to reduce transactions costs in cross-border mergers and acquisitions. It often preferable for all the parties to a deal to structure the deal in a "neutral" jurisdiction (Palan et al. [2013]). In this case, the negative association between asset concentration and tax haven utilization might reflect their usefulness in reducing administrative burdens and not necessarily in reducing the tax burden.

Table 4: Taxes as Dependent Variable

	Model 1	Model 2
HHI Assets	0.04	
	[0.01; 0.08]	
HHI Revenue		0.04
		[0.01; 0.08]
Controls	✓	✓
	[−0.03; 0.04]	[−0.03; 0.05]
Country Intercept	✓	✓
Industry Intercept	✓	✓
N	5201	5201
n_{eff}/n	0.33	0.35

Note: The dependent variable is firm-level tax to revenue ratio. Intercept is not shown. In both specification, controls for log assets, log revenue, and a number of (non-tax-haven) branches are added. Model 1 shows concentration of assets as the main explanatory variable, and Model 2 shows the concentration of assets as main dependent variable. Both specifications include industry-level and country-level random intercepts. Uncertainty estimates are 95-percent HDI intervals. The models are estimated using `brms` package in `R` (Bürkner [2017]).

In this section, instead of using connections to a tax haven as an indicator of tax avoidance, I use the actual amount of taxes paid divided by firms' revenues.[14] Table 4 presents the results.[15]

I find that, consistently with the previous results, asset concentration is positively correlated with tax compliance.

Highly Concentrated Firms Use Domestic Havens

Certainly, avoiding taxes through cross-border income shifting is not the only way firms avoid taxes. Many use *domestic* tax havens. It is also customary for countries in different parts of the world to offer location-based tax incentives to attract foreign capital. Firms thus can substitute cross-border tax avoidance by domestic tax optimization. The availability of domestic havens in certain countries can also lead firms to allocate more of their activities into that country thus increasing the concentration of assets.

[14] I do not divide by pre-tax profits because the profits can be manipulated by tax planning

[15] The number of observations is limited because, unfortunately, Orbis database data does not have the amount of taxes paid for many firms.

While this is a plausible concern, it can be alleviated with the results of Table 4, where the actual amount of taxes is used as a dependent variable.

Political Connections and Bribes as Omitted Variables

It has been documented that firms that are more politically connected receive more benefits from the government (Fisman [2001], Szakonyi [2015]). Firms can try to reduce their tax burden by acquiring political connections. Because the role of political connections and bribes is well-documented, my argument in this paper focuses on a different factor of tax compliance – infrastructural vulnerability. Nevertheless, firms' skill at acquiring political connections can play two roles. First, it can be an effect *modifier*: if a firm has a lot of activity in a country, it has incentives to acquire political connections to try to protect itself from any adverse government action and possibly reduce the tax burden. This mechanism would lead to attenuation bias in my main specifications since higher concentration in this explanation leads to higher level of political connections, and political connections in turn lead to lower tax burden and thus less urgency for tax planning.

Political connections can also serve as *a confounder*. If a firm is skilled in acquiring political connections in corrupt environments, then it will expand into more countries (this leads to lower concentration), but it also might use bribes to make host governments tolerate its tax avoidance. In this case, the negative relationship between the concentration of activity and tax avoidance would be spurious.

I do not observe firm-level political connection. However, if a firm has a comparative advantage in acquiring political connections in corrupt environments, then this skill can be measured by the average corruption index (weighted by assets in different countries). Table 5 presents estimations with this variable. I estimate a linear probability model for better interpretability of the interaction terms and the fixed effects. Column *Baseline* shows the baseline coefficient without the average corruption, but with the standard set of controls (log total assets, log total revenue, number of international branches, industry fixed effects, country fixed ef-

Table 5: Asset Concentration and Corruption Complementarity

	Baseline	Model 1	Model 2
HHI Assets	-0.06	-0.06	-0.07
	$[-0.08;\ -0.05]$	$[-0.08;\ -0.05]$	$[-0.09;\ -0.06]$
Corruption		0.01	-0.00
		$[-0.05;\ 0.06]$	$[-0.06;\ 0.05]$
HHI Assets $\times$ Corruption			-0.02
			$[-0.04;\ -0.01]$
Controls	✓	✓	✓
Country FE	✓	✓	✓
Industry FE	✓	✓	✓
R^2	0.27	0.27	0.27
N	5637	5637	5637

Note: The dependent variable is an indicator for tax haven usage of a firm. *Corruption* is 1 -asset-weighted standardized TI Corruption Perception Index. All specifications contain controls for log total assets, log total revenue and number of (non-tax-haven) branches. Uncertainty estimates are 95-percent confidence intervals calculated using robust standard errors. The regressions are estimated using `estimatr` package in `R`.

fects). Model 1 in Table 5 adds a firm-level average corruption index (measured as Corruption Perception Index from Transparency International) to test the confounding, and Model 2 adds an interaction term between corruption and asset concentration to test the effect modification.[16]

First, I found that adding those controls did little to the estimate of the main coefficient. The average estimate of the coefficient on corruption index is zero, and the coefficient on the interaction term is negative. These results are consistent with the effect modification and not consistent with the confounding.

7. Additional Prediction: Tax Compliance and Vertical Integration

This section offers data exploration inspired by Hypotheses 2 from Section 3: Firms that operate in more vertically integrated sectors are more tax compliant because the expropriation of a part of their supply chain produces more additional damage than the expropriation from a non-integrated firm.

[16]For better interpretability, I use 1-TI index, so that the larger values signify more corruption.

To test Hypothesis 2, I need a measure of sector-level vertical integration.[17] For this, I used a share of added value produced within the sector from the US Bureau of Economic Analysis.

To quantify the effect of sector-level vertical integration, I estimated the specification defined by Equations 2 - 3, but added an additional step that decomposed the sector-level random intercept (δ) from Equation 3:

$$\delta_{m[i]} = \eta + \rho VI_{m[i]} + \xi_m, \tag{4}$$

$$\xi_m \sim N(0, \sigma^2_{ind}). \tag{5}$$

Here, $VI_{m[i]}$ is a measure of vertical integration of sector m of firm i, and ξ_m is randomly distributed error term.

Table 6 presents the results of the estimation. Model 1 shows the minimal specification with concentration of assets, vertical integration, and industry effects. Model 2 adds firm-level control variables (branches, log revenue, and log assets) and country-level random intercepts.

I find that, as predicted by the model, vertical integration is negatively associated with the propensity to have an affiliate in a tax haven. In the fully controlled specification, the average coefficient is -0.59, which implies that, given the baseline tax haven utilization of 29 percent, one standard deviation in vertical integration is associated with 12 percentage points decrease in the propensity to have an affiliate in a tax haven. Section A.7 in the Appendix reports the results of Bayesian hypothesis tests for the coefficients: For the substantive coefficients the hypothesis of the coefficient being sufficiently close to zero and substantively negligible (between -0.05 and 0.05 on logit scale) is rejected.

[17] I use sector-level vertical integration instead of firm-level vertical integration due to the data availability and because the economic research on vertical integration shows that in many cases the economic determinants of vertical integration are fixed on the level of economic sector (Acemoglu et al. [2009].

Table 6: Tax Compliance and Vertical Integration

	Model 1	Model 2
HHI Asset	-0.66	-0.27
	$[-0.72;\ -0.59]$	$[-0.36;\ -0.19]$
Vert. Integration	-0.45	-0.59
	$[-0.88;\ -0.01]$	$[-1.06;\ -0.13]$
Controls		✓
Industry Effects	✓	✓
Country Effects		✓
N	6172	6172
n_{eff}/n	0.73	0.68

Note: The dependent variable is indicator for tax haven utilization. Vertical Integration is the sector-level proportion of value-added produced within a sector. Uncertainty estimates are 95 percent HDI intervals. Intercept is not shown. All variables are standardized. The models are estimated using `brms` package in `R` (Bürkner [2017]).

8. Sensitivity Analysis

Every statistical analysis of a social phenomena, including the one presented earlier, can suffer from misspecification. In my analysis, I considered the robustness of my results to different controls, including country-level and sector-level intercepts (modeled as either random effects or fixed effects), but the danger of unobserved covariates influencing both the main explanatory variable and the outcome cannot be avoided.

One of the ways to mitigate this concern is to investigate the *sensitivity* of the result to unobserved confounding. In this section, I considered such an analysis, following the procedure outlined in Cinelli and Hazlett [2018], which allows making conclusions about how large the potential omitted variable bias must be to nullify the results obtained in a regression.

Cinelli and Hazlett [2018] offer the following decomposition of omitted variable bias:

$$|\hat{bias}| = se(\hat{\alpha}) \sqrt{\frac{R^2_{Y\sim Z|X,D} R^2_{D\sim Z|X}}{1 - R^2_{D\sim Z|X}} (df)} \tag{6}$$

Here, $se(\hat{\alpha_{res}})$ is the standard error of the main coefficient of interest ($\hat{\alpha}$), Y is the

outcome of interest, D is the main explanatory variable, X is a vector of covariates, Z is the omitted variable, and df is degrees of freedom of the regression.

Intuitively, the absolute value of the bias is proportional to the connection of the outcome to the omitted variable (measured by the partial R^2) multiplied by the connection of the main explanatory variable to the omitted variable (also measured by corresponding partial R^2).

To use this formula, one needs a way to calibrate the expectations of the potential partial R^2's, involving the unobserved parameter. One way to do this is to assume (conservatively) that the connection of the unobserved variable to the outcome is unlikely to exceed the maximum observed connection of the *observed* variables to the outcome. In other words, we assume that the strongest possible predictor of the outcome is already in the regression.

In the case of my analysis, the strongest possible predictor of having an affiliate in a tax haven is the number of non-tax-haven branches of a firm. Applying this procedure yields the conclusion that the hypothetical bias *needed* to nullify my results is six to ten times larger than the estimate of the maximum bound of the potential bias. Thus, though my results can be vulnerable to misspecification and omitted variable, the procedure shows that the results are relatively stable (see Section A.9 in the Appendix for the detailed description of the steps of the procedure).

Another procedure for evaluating the sensitivity to unobserved confounding has been suggested in Oster [2017] and widely used across social sciences. The intuition behind it is relatively simple: If the main coefficient remains stable when the control variables are added, but R^2 goes up, then the main coefficient is declared relatively insensitive to unobserved confounding (under the assumption that the observed covariates constitute a representative sample for all the potential covariates). The procedure, outlined by Oster [2017] under a different set of assumptions than Cinelli and Hazlett [2018], calculates a value (δ) for how large the effect of unobserved confounder must be to nullify the effect of the main explanatory variable. I found that the unobserved confounder should be at least 1.4 times stronger than

any of my main variables to nullify my main coefficients (thus, $\delta = 1.4$). A "rule of thumb" suggested by Oster [2017] is that the values of δ smaller than 1 might indicate sensitivity to unobserved confounding.

9. Conclusions and Policy Implications

Corporate tax avoidance is one of the most important challenges to the fiscal capacity of states, especially in the poor and middle income parts of the world. Firms often establish affiliates in low-tax jurisdictions and use various accounting schemes to move their profits into those jurisdictions, thus robbing their host countries of tax revenue. As states are not able to achieve any of their goals (defense, contract enforcement, redistribution, public goods provision) without tax revenue, this problem is extremely consequential to the quality of governance.

In this paper, I aimed at proposing a theory of corporate income shifting. Inspired by Charles Tilly's view of the state as a quintessential protection racket, I hypothesized that the firms that are more vulnerable to value-reducing activity by the government will be *less* likely to engage in income shifting and more likely to demonstrate tax compliance. Using firm-level data that allow measuring connections to tax havens, I found that firms that have a higher concentration of assets and operate in more vertically integrated sectors are less likely to have an affiliate in a tax haven.

There are several important caveats to this analysis. My dependent variable is almost certainly measured with an error. Some firms hide their tax haven affiliate better than others, while other firms do not need to set up a tax haven affiliate because they are perfectly happy with the deductions they get from their host governments. Also not included in the analysis are within-country offshore zones. I attempted to mitigate those concerns by looking not only at the connection to tax havens but also at tax to revenue ratio as another measure of tax compliance and found results that are consistent with the results of my main specification.

My argument has important policy implications. In conflicts between firms and states, international community often sides with firms. Governments can be labeled "predatory," or "grabbing". For many years, the goal of the international system was to facilitate international investments, especially in the form of foreign direct investment, to protect property rights, and to encourage "pro-business" climate. My findings imply that, the same policies, agreements, and moral climate that promote "pro-business" attitudes might end up hurting the fiscal capacity and economic development of the host nations. A more balanced approach that embraces the value of tax compliance at least as enthusiastically as it now embraces the value of broadly defined property rights is needed.

REFERENCES

Daron Acemoglu, Simon Johnson, and Todd Mitton. Determinants of vertical integration: financial development and contracting costs. *The journal of finance*, 64(3): 1251–1290, 2009.

Daron Acemoglu, Davide Ticchi, and Andrea Vindigni. Emergence and persistence of inefficient states. *Journal of the European Economic Association*, 9(2):177–208, 2011.

Daron Acemoglu, Georgy Egorov, and Konstantin Sonin. A political theory of populism. *The Quarterly Journal of Economics*, 128(2):771–805, 2013.

Daron Acemoglu, Jacob Moscona, and James A Robinson. State capacity and american technology: evidence from the nineteenth century. *American Economic Review*, 106(5):61–67, 2016.

Christian Aid. Death and taxes: the true toll of tax dodging. *London, Christian Aid*, 2008.

Joshua D Angrist and Jörn-Steffen Pischke. *Mostly harmless econometrics: An empiricist's companion*. Princeton university press, 2008.

Raymond W Baker. *Capitalism's Achilles heel: Dirty money and how to renew the free-market system*. John Wiley & Sons, 2005.

Thorsten Beck, Chen Lin, and Yue Ma. Why do firms evade taxes? the role of information sharing and financial sector outreach. *The Journal of Finance*, 69(2): 763–817, 2014.

Timothy Besley and Torsten Persson. *Pillars of prosperity: The political economics of development clusters*. Princeton University Press, 2011.

Paul-Christian Bürkner. brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, 80(1):1–28, 2017. doi: 10.18637/jss.v080.i01.

Ling Chen and Florian M Hollenbach. Capital mobility & taxation in autocracies–evidence from china. *Working Paper*, 2018.

Carlos Cinelli and Chad Hazlett. Making sense of sensitivity: Extending omitted variable bias. 2018.

Ernesto Crivelli, Ruud A De Mooij, and Mr Michael Keen. *Base erosion, profit shifting and developing countries*. Number 15-118. International Monetary Fund, 2015.

Lisa De Simone, Kenneth J Klassen, and Jeri K Seidman. Unprofitable affiliates and income shifting behavior. *The Accounting Review*, 92(3):113–136, 2016.

Dhammika Dharmapala. What do we know about base erosion and profit shifting? a review of the empirical literature. *Fiscal Studies*, 35(4):421–448, 2014.

Mark Dincecco and Mauricio Prado. Warfare, fiscal capacity, and performance. *Journal of Economic Growth*, 17(3):171–203, 2012.

Matthias Dischinger, Bodo Knoll, and Nadine Riedel. The role of headquarters in multinational profit shifting strategies. *International Tax and Public Finance*, 21(2): 248–271, 2014.

Art Durnev and Sergei Guriev. Expropriation risk and firm growth: A corporate transparency channel. *Working Paper*, 2011.

Scott D Dyreng, Michelle Hanlon, and Edward L Maydew. Long-run corporate tax avoidance. *the accounting review*, 83(1):61–82, 2008.

Tasha Fairfield. *Private wealth and public revenue*. Cambridge University Press, 2015.

Raymond Fisman. Estimating the value of political connections. *American economic review*, 91(4):1095–1102, 2001.

Jessica Fortin. Is there a necessary condition for democracy? the role of state capacity in postcommunist countries. *Comparative political studies*, 45(7):903–930, 2012.

Susan D Franck. Empirically evaluating claims about investment treaty arbitration. *NCL Rev.*, 86:1, 2007.

Scott Gehlbach. *Representation through Taxation: Revenue, Politics, and Development in Postcommunist States*. Cambridge University Press, 2008.

Jessica Gottlieb and Florian M Hollenbach. Fiscal capacity as a moderator of the taxation-accountability hypothesis. 2018.

Mark Hampton and Jason Abbott. *Offshore finance centers and tax havens: The rise of global capital*. Purdue University Press, 1999.

Margaret Hodge. Called to account: How corporate bad behaviour and government waste combine to cost us millions. *Little, Brown Book Group*, 2016.

Florian M Hollenbach and Thiago Silva. Fiscal capacity and inequality: Evidence from brazilian municipalities. 2018.

Harry Huizinga and Luc Laeven. International profit shifting within multinationals: A multi-country perspective. *Journal of Public Economics*, 92(5-6):1164–1182, 2008.

Petr Janskỳ and Alex Prats. International profit-shifting out of developing countries and the role of tax havens. *Development policy review*, 33(3):271–292, 2015.

Barbara Jenkins. Reexamining the obsolescing bargain: a study of canada's national energy program. *International Organization*, 40(1):139–165, 1986.

Niels Johannesen and Gabriel Zucman. The end of bank secrecy? an evaluation of the g20 tax haven crackdown. *American Economic Journal: Economic Policy*, 6(1): 65–91, 2014.

Leslie Johns and Rachel L Wellhausen. Under one roof: Supply chains and the protection of foreign investment. *American Political Science Review*, 110(1):31–51, 2016.

Mozaffar Khan, Suraj Srinivasan, and Liang Tan. Institutional ownership and corporate tax avoidance: New evidence. *The Accounting Review*, 92(2):101–122, 2016.

Stephen J Kobrin. Testing the bargaining hypothesis in the manufacturing sector in developing countries. *International organization*, 41(4):609–638, 1987.

Sebastian Krautheim and Tim Schmidt-Eisenlohr. Heterogeneous firms,profit shiftingfdi and international tax competition. *Journal of Public Economics*, 95(1-2):122–133, 2011.

John K Kruschke. Rejecting or accepting parameter values in bayesian estimation. *Advances in Methods and Practices in Psychological Science*, 1(2):270–280, 2018.

Markku Laakso and Rein Taagepera. effective number of parties: a measure with application to west europe. *Comparative political studies*, 12(1):3–27, 1979.

Quan Li. Democracy, autocracy, and tax incentives to foreign direct investors: A cross-national analysis. *The Journal of Politics*, 68(1):62–74, 2006.

Daniel Lüdecke. *sjstats: Statistical Functions for Regression Models (Version 0.17.0)*, 2018. URL `https://CRAN.R-project.org/package=sjstats`.

Paolo Mauro. Corruption and the composition of government expenditure. *Journal of Public economics*, 69(2):263–279, 1998.

Richard Murphy. *Dirty secrets: How tax havens destroy the economy*. Verso Books, 2017.

Emily Oster. Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, pages 1–18, 2017.

Ronen Palan, Richard Murphy, and Christian Chavagneux. *Tax havens: How globalization really works*. Cornell University Press, 2013.

Stephen A Rhoades. The herfindahl-hirschman index. *Fed. Res. Bull.*, 79:188, 1993.

Philip G Roeder. Ethnolinguistic fractionalization (elf) indices, 1961 and 1985, 2001.

Kenneth Scheve and David Stasavage. The conscription of wealth: mass warfare and the demand for progressive taxation. *International organization*, 64(4):529–561, 2010.

Thomas Schultz and Cédric Dupont. Investment arbitration: promoting the rule of law or over-empowering investors? a quantitative empirical study. *European Journal of International Law*, 25(4):1147–1168, 2014.

Nicholas Shaxson. *Treasure islands: uncovering the damage of offshore banking and tax havens*. St. Martin's Press, 2011.

Andrei Shleifer and Daniel Treisman. *Without a map: Political tactics and economic reform in Russia*. MIT Press, 2001.

David Szakonyi. Elected public office and private benefit: Firm-level returns from businesspeople becoming politicians in russia. Technical report, Mimeo. Retrieved on 16 October, 2015.

Brian D Taylor. *State building in Putins Russia: Policing and coercion after communism*. Cambridge University Press, 2011.

Stan Development Team. Stan: A c++ library for probability and sampling, version 2.2, 2014.

Charles Tilly. War making and state making as organized crime. *Violence: A reader*, pages 35–60, 1985.

Thomas R Tørsløv, Ludvig S Wier, and Gabriel Zucman. The missing profits of nations. Technical report, National Bureau of Economic Research, 2018.

Aki Vehtari, Andrew Gelman, and Jonah Gabry. Practical bayesian model evaluation using leave-one-out cross-validation and waic. *Statistics and Computing*, 27 (5):1413–1432, 2017.

Rachel L Wellhausen. *The Shield of Nationality: When Governments Break Contracts with Foreign Firms*. Cambridge University Press, 2014.

Rachel L Wellhausen. Recent trends in investor–state dispute settlement. *Journal of International Dispute Settlement*, 7(1):117–135, 2016.

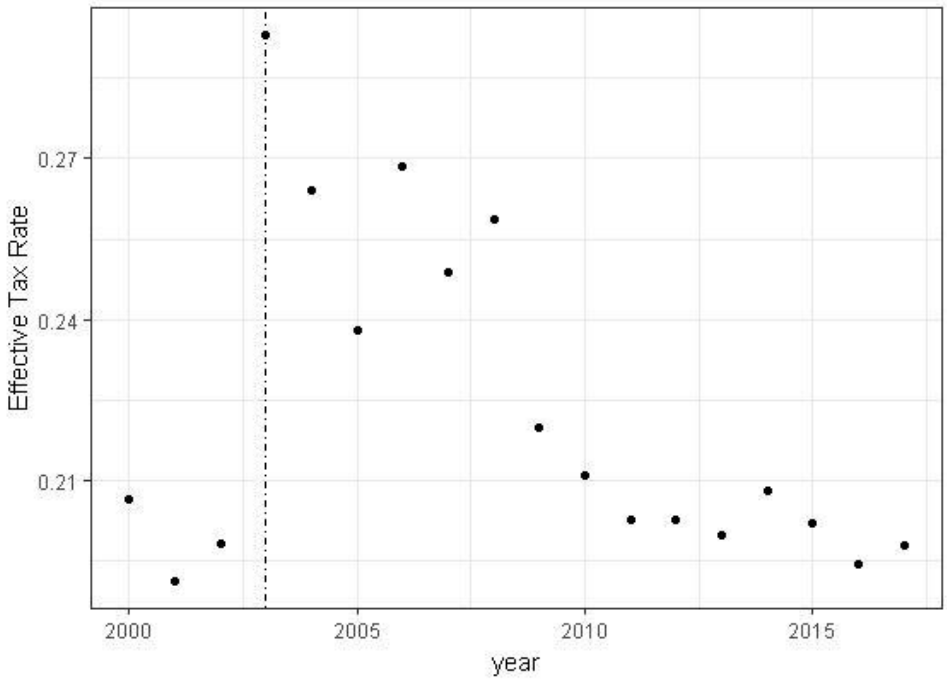
Gabriel Zucman. Taxing across borders: Tracking personal wealth and corporate profits. *Journal of Economic Perspectives*, 28(4):121–48, 2014.

Gabriel Zucman. *The hidden wealth of nations: The scourge of tax havens*. University of Chicago Press, 2015.

APPENDIX

A.1. Russian Private Oil Firms Before and After Yukos Affair

Figure A1: Median Effective Tax Rate



Note: A point represents a median effective tax rate for the subsample of non-government owned oil-producing Russian firms. Data come from WRDS/COMPUSTAT database

A.2. Formal Model of Investment and Tax Compliance

The game consists of two players: a firm and a government of a host country. On the first step, the firm decides what proportion s of their total assets A to put into the host host country, and what proportion to allocate elsewhere. The profit the firm receives in the host country is $p = rsA$, where $0 < r < 1$ and the rate of return elsewhere is normalized (without loss of generality) to zero. In the host country, the firm also faces a proportional tax rate t such that $0 < t < 1$ and decides about tax compliance $\kappa \in [0, 1]$, a share of the tax rate the firm actually pays. I assume, for simplicity, that the government does not have means to enforce the tax rate t.[18]

On second step, the government chooses an action (a). It decides whether to expropriate a firm ($a = 1$) or not ($a = 0$). If government chooses to expropriate a firm, it gets a utility:

$$U_g[a = 1] = E + msA,$$

Here, E is a political benefit from expropriation. For example, if a government is nationalistic or extremely left wing, then expropriating a foreign firm might boost

[18]This assumption is not restrictive. The enforcement capacity can be modeled as a non-zero low boundary for κ, and it will not change the results.

the government's popularity. If government is corrupt and survives by patron - client exchanges, then E can be interpreted as a redistribution to the government's supporters. The theory is intentionally agnostic about where E comes from. E can also be negative if a government might face a backlash from domestic and international audiences. The government also acquires and asset from firms, suffering some $(1-m)$ loss of the expropriated, $m \in [0,1]$.[19]

If a firm is not expropriated, then the government receives the utility:

$$U_g[a=0] = t\kappa rsA$$

This is just the tax payment the firm chooses to pay.

After the government decides on its action, the firm receives its utility. If the firm is expropriated, then the firm's utility is:

$$U_f[a=1] = -(1+d)sA$$

If the government chooses expropriation then the firm loses an asset and gets nothing in return. It also suffers from additional loss dsA, where d is the coefficient of proportionality of that additional loss to the initial investment. One way to interpret d is to view it as a measure of vertical integration of the firm. This additional damage is captured by parameter d.

If the government chooses not to expropriate the firm, then the firm gets its profit minus the effective tax:

$$U_f[a=0] = (1-\kappa t)rsA$$

I follow Johns and Wellhausen [2016] in assuming that the firm does not observe E directly, but instead has a uniformly distributed belief about E:

$$E \sim U[-k,k]$$

Here, $2k$ is the length of the interval, so that k can be interpreted as a measure of uncertainty about the goals of the government and the risks that the firm faces.

This completes the definition of the model. We are interested in the equilibrium relationship between tax compliance κ and the share of investment a as well as the dependence between κ and other parameters of the model.

To solve this game using subgame-perfect Nash equilibrium, we first need to notice, that the firm is expropriated if:

$$E + msA \geq \kappa trsA$$

Without loss of generality, we set $A=1$.

Because, from the point of view of the firm, E is a random variable, then, conditional on the choices of the firm, the probability of not being expropriated is:

$$P[a=0] = \frac{\kappa trs - ms + k}{2k}$$

[19] It is possible for m to be 1 if the government is as competent as the firm, or 0 if the government is completely incompetent, or anything in between.

Before the firm makes any decisions, its expected utility is:

$$EU(s,\kappa) = P[a=0] \times (1-\kappa t)(rs+s) + P[a=1] \times (-ds) - s$$

Maximizing the expected utility wrt κ and s, we get the following expression for κ (given that $k \neq 0$ and $rst \neq 0$):

$$\kappa = \frac{ds - k + ms + rs + s}{2rst}$$

This expression is increasing in s (Hypothesis 1), increasing in d (Hypothesis 2).

A.3. Characteristics of the Sample

Table A.1: Descriptive Statistics for Firm-Level Financial Data

	Variable	Mean	SD
1	Assets, billion USD	17.33	92.06
2	Revenue, billion US	5.47	11.69
3	Tax Haven Use, share	0.29	0.45
4	Fixed Assets HHI	0.96	0.11
5	Revenue HHI	0.95	0.12

Note: The main descriptive variables for the 6,985 observations in my sample.

Table A.2: Number of Firms per Country

Country	Number of Firms	Country	Number of Firms
JPN	754	USA	746
CHN	518	KOR	384
DEU	362	GBR	351
ITA	338	FRA	323
IND	230	CAN	205
ESP	190	RUS	190
COL	176	SWE	170
AUS	169	DNK	156
NOR	153	MYS	128
AUT	126	BEL	124
BRA	103	FIN	96
THA	90	ZAF	90
TUR	87	POL	57
ISR	47	PHL	41
IDN	40	GRC	39
MEX	38	CZE	33
PRT	30	CHL	29
IRN	29	SAU	28
VNM	27	HUN	25
NZL	24	UKR	24
PER	23	DZA	21
KAZ	18	PAK	17
ARG	15	LTU	13
VEN	12	ROM	11
SVK	8	SVN	8
BWA	7	HRV	7
URY	7	MAR	5
SRB	5	EGY	4
BGD	3	BIH	3
ECU	3	KWT	3
LVA	3	UZB	3
DOM	2	EST	2
GTM	2	KEN	2
LKA	2	MKD	2
ALB	1	AZE	1
BGR	1	IRQ	1
MDA	1	PNG	1
PRY	1	SDN	1
SLV	1		

Table A.3: Number of Firms per Economic Sector

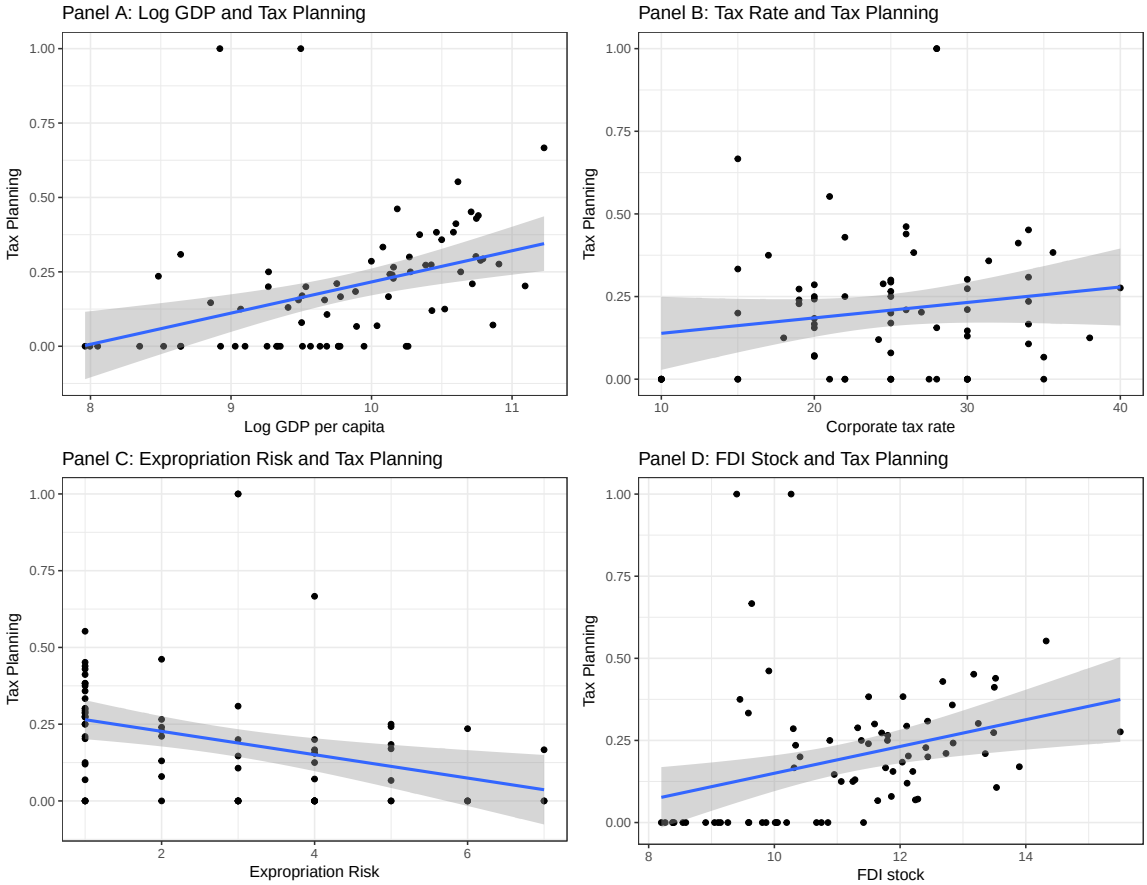
	Sector	Number of firms per sector
1	Manufacturing	2028
2	Finance	1297
3	Retail	1222
4	Prof. Services	433
5	Electricity	347
6	IT	291
7	Transport	288
8	Mining	236
9	Construction	230
10	Adm. Services	148
11	Real Estate	123
12	Health	62
13	Accomodation	53
14	Agriculture	52
15	Arts	47
16	Pub. Adm	47
17	Other Services	45
18	Water	33
19	Education	8

Andorra	Anguilla
Antigua and Barbuda	Aruba
Bahamas	Bahrain
Barbados	Belize
Bermuda	Bolivia
Brunei Darussalam	Cape Verde
Cayman Islands	Comoros
Costa Rica	Curaao
Cyprus	Djibouti
Dominica	Falkland Islands
Fiji	Gambia
Gibraltar	Grenada
Guernsey	Guyana
Honduras	Hong Kong
Ireland	Isle Of Man
Jamaica	Jersey
Jordan	Kiribati
Lebanon	Liberia
Liechtenstein	Luxembourg
Macao	Maldives
Malta	Marshall Islands
Mauritius	Monaco
Netherlands	Oman
Palau	Panama
Pitcairn Islands	Qatar
Saint Helena	Saint Kitts and Nevis
Saint Lucia	Saint Vincent and the Grenadines
Samoa	San Marino
Sao Tome and Principe	Seychelles
Singapore	Sint Maarten
Solomon Islands	Swaziland
Switzerland	Taiwan
Tonga	Trinidad and Tobago
Turks and Caicos Islands	Tuvalu
United Arab Emirates	Vanuatu
Vatican City State	Virgin Islands
Yemen	

A.5. Country-level Determinants of Income-Shifting

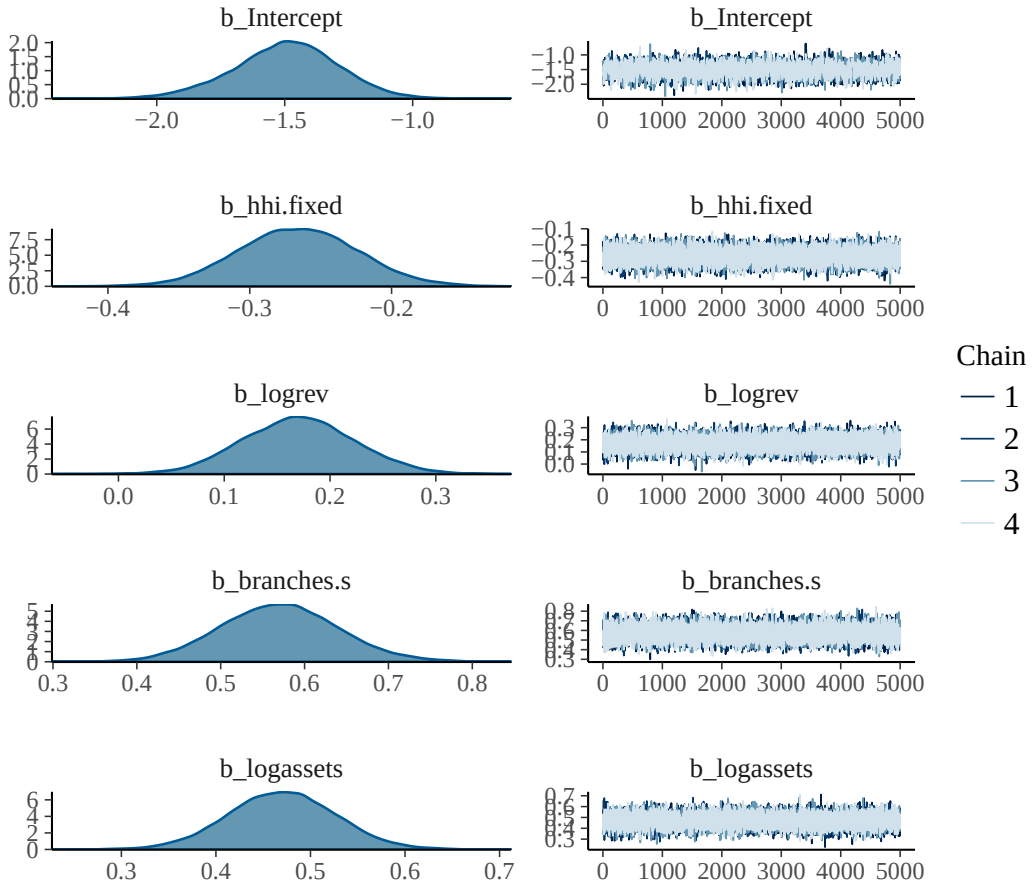
Figure A2 shows several correlations that might help systematize the data on the country-level prevalence of income shifting. We see that more number of firms engage in income shifting in countries that are richer (Panel A), have higher taxes (Panel B), have *lower* expropriation risk (according a Belgian consulting firm Credendo) (Panel C), and have higher stock of FDI (Panel D). All these results are consistent with the conventional wisdom on tax avoidance, except for the results in Panel C. According to the prevailing institutional theory, we would expect more profit shifting to the offshore jurisdictions from the firms that operate in places with less secure property rights. Nevertheless, we observe the opposite: The firms who operate in less secure countries are less likely to have affiliates in tax havens. This correlation is consistent with the protection racket theory presented earlier.

Figure A2: Country-Level Covariates for Tax Haven Utilization



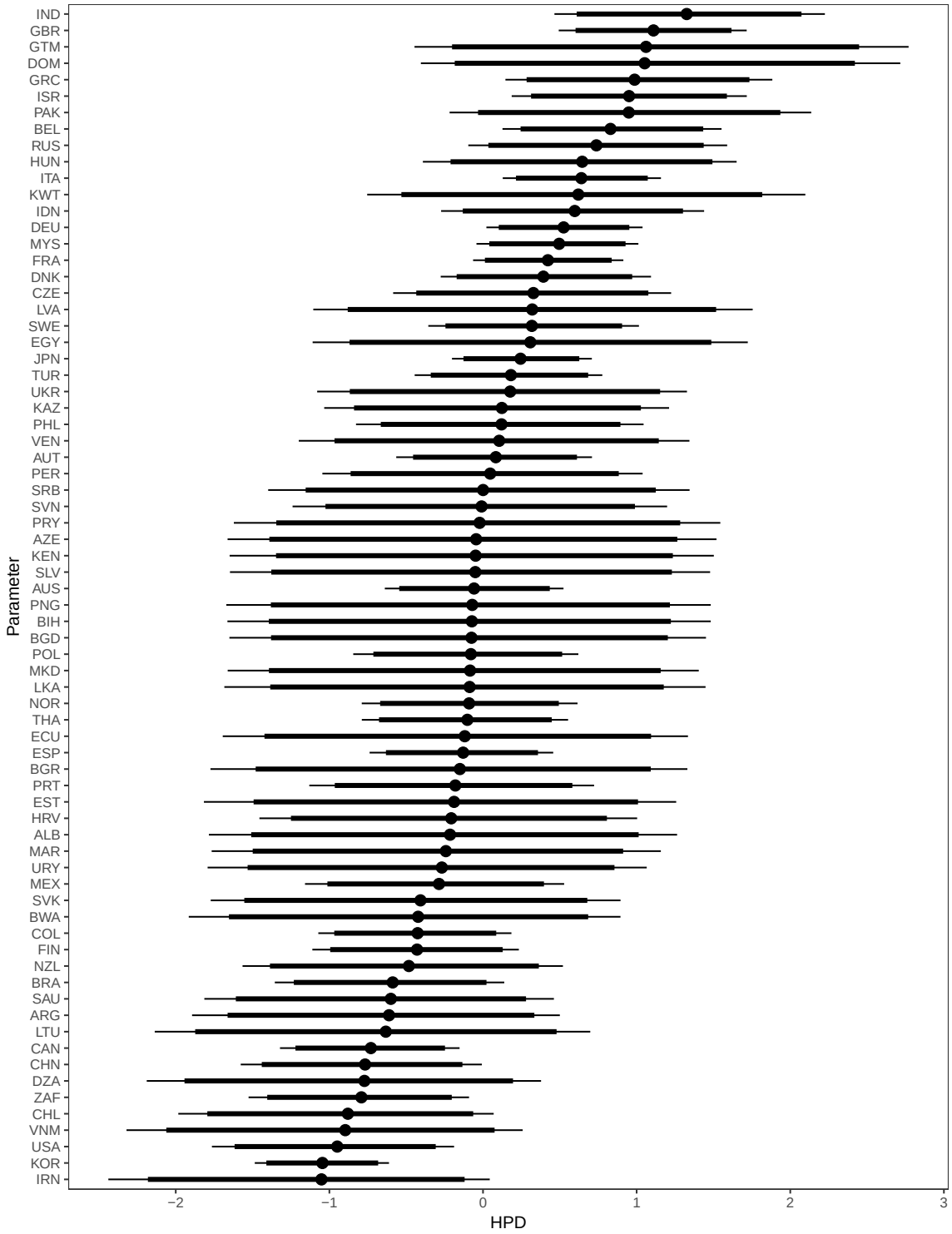
Note: Vertical axis shows the value Tax Planning variable (share of the largest firms utilizing tax havens): Share of firms with affiliates in tax havens operating in a particular country. On a horizontal axis, Panel A shows log GDP per capita, Panel B shows corporate tax rate, Panel C shows expropriation risk, and Panel D shows FDI stock.

Figure A3: Posterior Distributions and MCMC Chains



Note: Posterior plots for the estimates of the main coefficients in Table 1. *b_Intercept* is an estimates of the intercept, *b_hhi.fixed* is an estimate of the coefficient on asset concentration, *b_logrev* is an estimate of the coefficient on log revenue, *b_branches.s* is an estimate of the coefficients on the number of branches, *b_logassets* is an estimate on the coefficient on log assets

Figure A4: Estimates of Country-Level Random Effects for Tax Haven Utilizations



Note: Random effect estimates from Model 4 in Table 1

A.7. *Bayesian Hypothesis Tests*

In Table 1, I do not perform hypotheses tests since in the Bayesian framework, the null hypothesis of "no association" is trivially rejected: because the coefficient of interest is represented by a continuous random variable, any single value (including 0) has a probability of 0.

It is still possible to perform hypotheses tests, but one needs to specify a *range* of values that are as good as 0 for all practical purposes (this region is called ROPE – "region of practical equivalence").[20] Such tests are more conservative than the traditional tests in the Neyman-Pearson framework (we want to reject a range of values, not just one value of zero) and arguably more substantively meaningful since in the traditional framework one might declare "a finding" with enough data even if the size of the effect is too small to be substantively important. In the Bayesian framework, we first define what counts as a negligible effect, and then evaluate what proportion of the posterior distribution (or HDI) lies inside/outside the ROPE.

A substantive judgment is needed to specify a ROPE in any applied context. Kruschke [2018] suggests a ROPE of $[-0.05; 0.05]$ on a logit scale for models with binary dependent variable. Intuitively, if one standard deviation of the explanatory variable is associated with less than one percentage point of a probability of the tax-haven connection (given the baseline value of 30 percent), the effect size should be deemed too small to matter substantively. Figure A5 shows a proportion of HDI inside the ROPE for the substantive coefficients in Model 4 from Table 1 (my preferred specification). We see that all the HDIs lie outside the recommended ROPE. So, if we are to make a binary decision about the parameter values, we can reject the hypothesis that the effect of the concentration of assets is "zero".

Figure A6 shows the same procedure for Model 2 in Table 6 that relates the sector-level vertical integration to the the tax-haven connection. Similarly to the previous example, the HDI for the vertical integration coefficient lies outside the ROPE.

A.8. *Goodness of Fit*

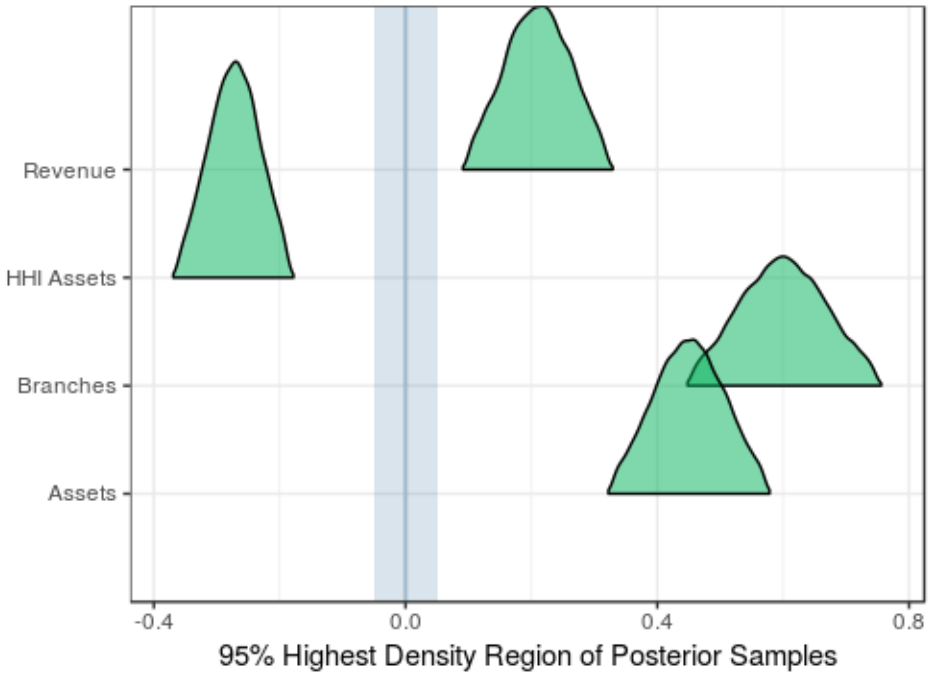
How well does the statistical specifications explain the variation in my data? Because I have the binary outcome data, the standard measure of goodness-of-fit in a regression framework might not be applicable since the fitted values are interpreted as probabilities, while the outcomes are probabilistic realizations of a Bernoulli random process parametrized by those probabilities. Values of R^2 calculated out of those quantities would be very low, but those low values would not imply low explanatory power of the model.

Figure A7 presents the histogram of actual outcome data (in red), and the histogram of predicted probabilities from Model 4 in Table 1 (in blue). As one might see, the distribution of predicted probabilities peak around low values, which is consistent with most of the firms in my dataset that do not have a known connection with tax havens.

One can also see a significant proportion of probability density lying above the value of 0.5, indicating that there is still a proportion of firms that the model would

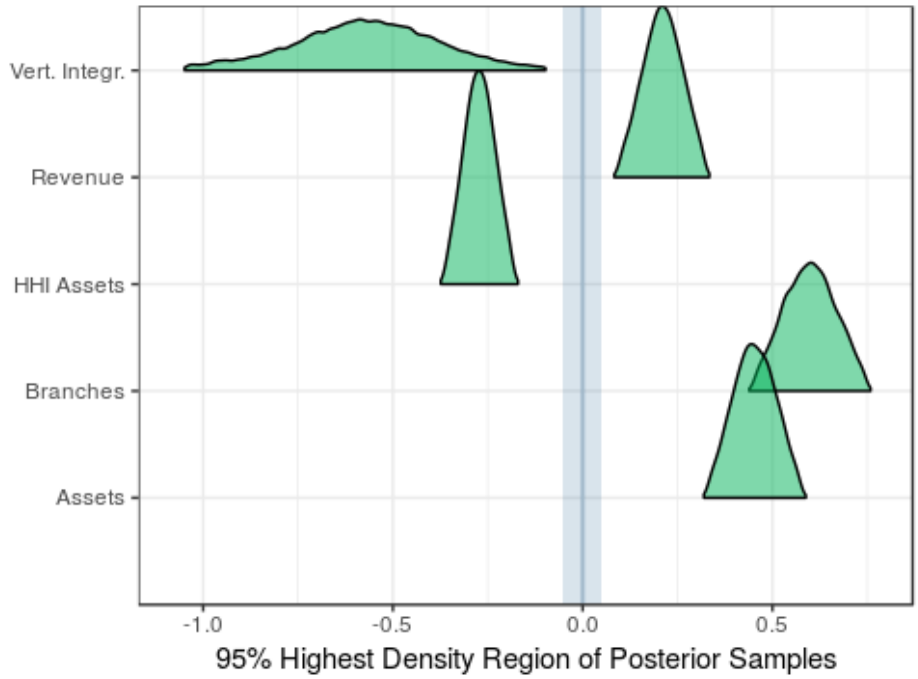
[20] See Kruschke [2018] for technical details.

Figure A5: Equivalence Tests for Table 1



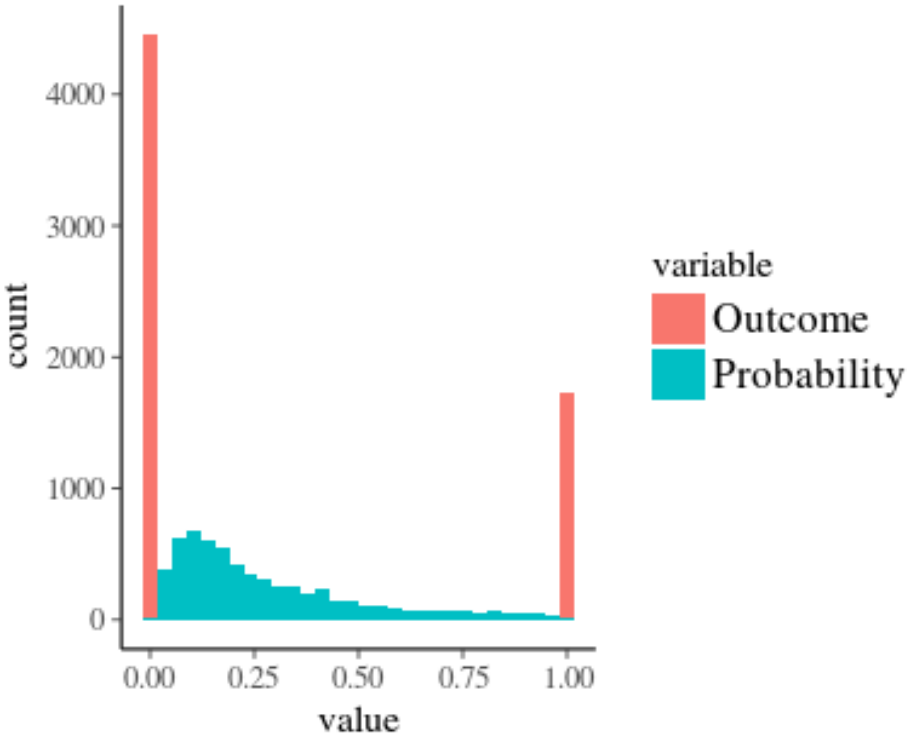
Note: The density plots represent posterior distributions for the coefficients in Model 4, Table 1. Shaded area represents a region of practical equivalence: an interval of logit values between -0.05 and 0.05. The tests are implemented using package `sjstats` in `R` (Lüdecke [2018]).

Figure A6: Equivalence Tests for Table 6



Note: The density plots represent posterior distributions for the coefficients in Model 2, Table 6. Shaded area represents a region of practical equivalence: an interval of logit values between -0.05 and 0.05.The tests are implemented using package `sjstats` in `R`. The tests are implemented using package `sjstats` in `R` (Lüdecke [2018]).

Figure A7: Outcomes and Posterior Probabilities



Note: A histogram of the outcomes (red) and predicted probabilities (blue). Outcomes are taken from the data (1 if a firm has an affiliate in a tax haven, 0 if a firm does not have an affiliate in a tax haven). Predicted probabilities are produced by Model 4 in Table 1.

classify as having an affiliate in a tax haven.

To calculate the explanatory power of the model, I used an in-sample predictive accuracy: I used a fitted value for the p_i and then predicted that those firms that have fitted value of p_i larger than 0.5 do have an affiliate in a tax haven, while the firms whose predicted p_i is less or equal to 0.5 do not. As a result, I can compare how well my model explains the outcome and compare its performance against the relevant benchmarks: A random prediction (treating the outcomes as equally likely), or assigning to all the variables as a prediction, the most frequent outcome.[21]

I calculated the predictive accuracy as $Accuracy = \frac{number\_of\_correctly\_predicted}{N}$, a ratio of correctly predicted to the total number of cases. This calculation yields a value of 0.78. This value is larger than potential accuracies from random assignment of the outcomes and from the assignment of most frequent cases. A more general approach to assess the explanatory power of the model is to create an ROC curve that would, for each of the potential value of a cutoff (not just the value 0.5), calculate the true positive rate (number of firms correctly predicted to have tax haven affiliate divided by the total number of firms that have tax haven affiliates), and 1- false positive rate (number of firms incorrectly predicted to have tax haven affiliate divided by the number of firms that do not have a tax haven affiliate). This curve will show the quality of the explanatory model in explaining the outcomes for every possible fitted value of p_i used as a threshold for classification. The resulting ROC curve is shown in Figure A8. The curve shows that for all nontrivial values of the cutoff, the classifier outperforms the random assignment benchmark.

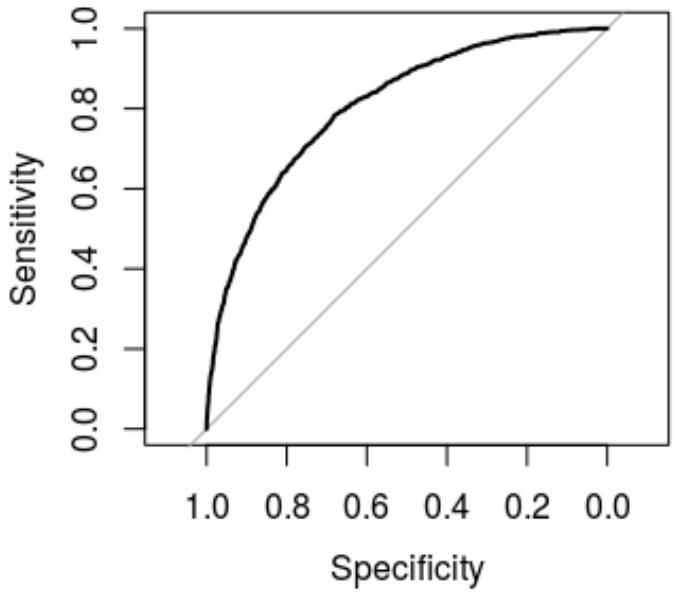
Another important question that can be asked is what is the role of the proxies in vulnerability in the explanatory power of the model? One way to answer this question is to use an information criteria based on leave-one-out (LOO) cross-validation (Vehtari et al. [2017]). To investigate whether the proxies for vulnerability can improve the predictive power of the model, I considered two models. In the first one, the baseline model, I removed all the proxies for the vulnerability and kept only the control variables: Log revenue, log assets, FDI stock, sector-level effects and country-level effects. In the second model, I added back the proxies for vulnerability: Concentration of fixed assets, concentration of fixed assets, sector-level vertical integration, and country-level expropriation risk. For both the models, I calculated the value of LOO information Criteria (LOOIC) and its standard deviation.

Figure A9 shows the result. Vertical axis shows the values of LOOIC and its 95-percent uncertainty intervals. While the values itself are not interpretable, the smaller values indicate better fit. As one can see from Figure A9 the full model is displaying a considerably better fit than the baseline model (the model without the proxies for vulnerability).

In sum, this subsection demonstrated that the Bayesian multilevel specification used for the analysis lead to the classification that fits the data better than the appropriate benchmarks, and, as implied by the protection racket theory of tax compliance developed earlier in this paper, the measures of the firm's vulnerability lead to better predictions of firm-level utilization of tax havens.

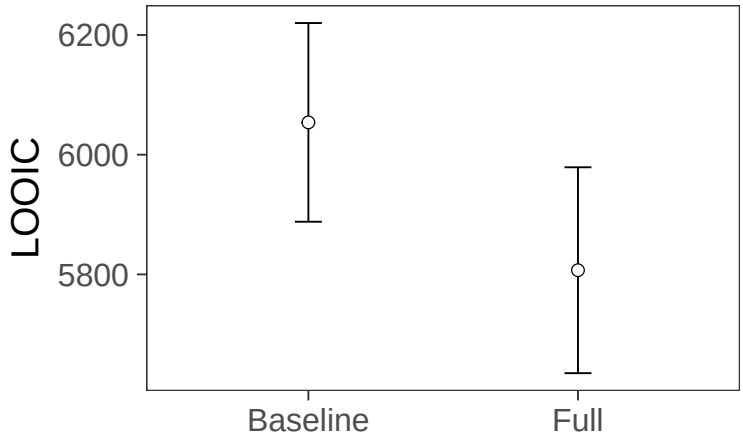
[21] This procedure would yield an accuracy of 72 percent.

Figure A8: ROC Curve for the Classifier



Note: The ROC curve is based on the fitted values from Model 4 of Table 1. Sensitivity, depicted on the vertical axis is the true positive rate. Specificity, depicted on the horizontal axis is 1-false positive rate. The black curve shows the characteristics of the classifier based on Model 4 of Table 1. The gray line shows the characteristics of the random classifier. The area between the gray line and the black curve can be used for the assessment of the quality of the classifier

Figure A9: Leave-One-Out Criteria



Note: Leave-one-out cross-validation information criteria (LOOIC). Baseline model (the model with all the covariates except for the proxies for vulnerability) is shown on the left. Full model (baseline + proxies for the vulnerability) is shown on the right. Smaller values indicate better fit.

A.9. Procedure for Sensitivity Analysis

This section outlines the procedure for the sensitivity analysis. I perform the following steps.

1. Standardize all the variables, including the outcome.
2. Estimate an OLS regression of indicator of having an affiliate in a tax haven on all the variables from Model 4 in Table 1, except for random intercepts. Save the residuals (r_y)
3. Estimate an OLS regression of indicator of number of branches on all the variables from Model 4 in Table 1, except for random intercepts. Save the residuals (r_x)
4. Regress r_y on r_x. Save R^2. In this procedure, I get an R^2 of 0.007.[22]
5. Use the this number as an estimate of both $R^2_{Y\sim Z|X,D}$ and $R^2_{D\sim Z|X}$. For $se(\hat{\alpha})$ use the largest of the standard errors on any of the vulnerability proxies (concentration of assets and vertical integration).[23] I use the value 0.022 (the standard error of the coefficient on the concentration of assets). The value of df is 6147.
6. After plugging in the numbers into the Cinelli & Hazlett formula, I get the maximum value of bias 0.011
7. To nullify the results the bias should be at least as large as the value of the coefficient of interest. In the OLS regression with standardized value of the outcome, the estimate of the coefficient on concentration of assets is -0.11, the estimate of the coefficient on the concentration of revenue is -0.15, the coefficient on vertical integration is -.20. Thus the values of the bias needed to nullify the results exceed the estimated bound of the potential bias at least ten times, or if the closest to zero estimate of the bound of confidence interval is considered, at least six times.

[22] I do not include fixed effects into the regression to recover partial R^2, the actual value of partial R^2 is likely to be biased upwards. The resulting bound for bias would also be biased upwards because of this reason and because the omission of fixed effects drives up the degrees of freedom.

[23] A more accurate way it calculate the bounds for the bias separately for all the coefficient of interest. Here, for simplicity, I implement a more conservative procedure: calculate the maximum possible bound and use it for all the coefficients.