NBER PE – Tax Evasion and Inequality*

Annette Alstadsæter (Norwegian University of Life Sciences)
Niels Johannesen (University of Copenhagen)
Gabriel Zucman (UC Berkeley and NBER)

January 17, 2017

Abstract

How does tax evasion vary with wealth in developed countries, and why? We address this question using randomized audit data and recent leaks from financial institutions in offshore tax havens matched to population-wide registry data in Norway, Sweden, and Denmark. Randomized audit data show that the probability to deliberately evade taxes is very low up to the 95th percentile but then rises with wealth up to the top 1%. Leaks from tax havens show that evasion rises steeply within the top 1%, a phenomenon that random audits fail to capture. Overall, we estimate that the fraction of the population that evades taxes rises from 1%-2% in the bottom 95% of the wealth distribution to about 40% in the top 0.05%. A simple model of the supply of tax evasion services can explain why evasion rises steeply with wealth at the top. We use our findings to construct revised estimates of wealth inequality in Scandinavia. Taking tax evasion into account increases the rise in inequality seen in tax data markedly. We also find that tax avoidance rises sharply with wealth, and that after voluntarily reducing tax evasion—by using tax amnesties—tax evaders start avoiding more. Taken together, these findings highlight the need to move beyond tax data to capture income and wealth at the top, even in countries where tax compliance is generally high. They also suggest that collecting extra tax revenue from the wealthy would require curbing the supply of evasion services (e.g., by making it more costly and risky for banks to serve tax evaders) and simultaneously closing loopholes to prevent substitution from evasion to avoidance.

*Annette Alstadsæter: annette.alstadsater@nmbu.no; Niels Johannesen: niels.johannesen@econ.ku.dk; Gabriel Zucman: zucman@berkeley.edu. We thank Torhild Henriksen, Paul Gunnar Larssen, Petter Lundberg, Audun Solli, Gard Thomassen, Susanna Wanander, Søren Pedersen and UiO Services for Sensitive Data for tireless and extremely valuable assistance. Johannesen acknowledges financial support from the Danish Council for Independent Research.

1 Introduction

There is considerable interest among the public in issues of tax evasion. A key bone of contention in the public debate is who evades taxes. In developed countries, the widespread reporting of income by third parties limits the ability to evade. Some believe that whatever evasion nonetheless occurs is concentrated at the top-end of the distribution and primarily benefits the wealthiest among us. Others stress that poorer individuals are also likely to evade taxes, highlighting fraud to refundable tax credits or tax evasion by self-employed individuals and small businesses. Beyond shedding light on this debate, more information on how tax evasion is distributed would be valuable for two reasons. First, it would enable tax authorities—which usually face tight budget constraints—better target their enforcement effort. Second, scholars are increasingly relying on tax data to study questions related to inequality, mobility, or the distributional effects of tax and transfer policies. A key concern raised by the use of tax returns to measure inequality, and indeed one of the main reasons why tax data have for a long time been viewed with skepticism, is that they only provide information about income or wealth reported to the tax authority—not true economic income and wealth. The tax rates that wealthy households face, the tax evasion technologies that they can use, and the enforcement mechanisms that governments implement all have changed dramatically over time, potentially biasing time series and cross-country comparisons of inequality substantially.

Despite interest in this question, we do not have a clear view of how tax evasion varies across the wealth distribution. There are two reasons for this. First, and most important, the main data source used to study tax evasion—stratified random audits—has important limitations. Random audits are a powerful way to capture relatively simple forms of tax evasion, e.g., by self-employed individuals and small businesses, and have been fruitfully used to do so (e.g., Kleven et al., 2011). But it is widely acknowledged that they miss the most sophisticated forms of evasion, in particular wealth concealment through shell corporations and offshore holdings, the detection of which would require much more resources than available to tax authorities for their random audit programs. Second, few countries have administrative data on wealth; the literature therefore focuses on how evasion varies with taxable income. While a useful indicator, taxable income can be quite far from the actual capacity to pay taxes; this might especially be the case for wealthy tax evaders who in addition to evading taxes may reduce taxable income through various legal means.

This paper casts new light on the distribution of tax evasion by analyzing micro-data leaked from a large offshore financial institution—HSBC Private Bank Switzerland—and combining

them with randomized audit data and population-wide administrative wealth records. In 2007, an employee of HSBC Switzerland extracted the complete internal records—including the names and assets—of the more than 30,000 clients of this Swiss bank, a large fraction of whom were evading taxes. This list was subsequently obtained by the French tax authority who shared it with foreign administrations. Thanks to a cooperation with Scandinavian tax authorities, we are able to match the names on the HSBC list to individual tax data in three countries, Norway, Sweden, and Denmark. To our knowledge, it is the first time that leaked micro-data matched to tax returns are used to study tax evasion. These data allow us to provide the first evidence on how offshore tax evasion is distributed, enabling us to fill what is perhaps the key gap of random audit studies. While other forms of evasion may go undetected in audits—e.g., some fringe benefits not subject to third-party reporting, or in-kind exchanges among professionals—offshore evasion is the one that has received the greatest amount of attention in the public debate in recent years. And as we shall see, including it leads to a very different picture than the one painted by audit data.

The leaked HSBC file has four key strengths. First, the leak is not the result of specific enforcement effort by tax authorities targeted at HSBC; it can be seen as a random event. Second, it involves a major—and, the available evidence suggests, representative—player in the offshore wealth management industry, a fact we use to make inference about the overall prevalence of offshore tax evasion in the entire population. At the time of the leak, HSBC Private Bank Switzerland, the Swiss subsidiary of the banking giant HSBC, managed $118 billion in wealth, about 4% of all the foreign wealth held in Switzerland; it had clients from all over the world; and as we show below, the wealth it managed was similarly distributed across countries as that managed by the overall Swiss banking industry. Third, HSBC Switzerland recorded the name of the beneficial owners of the wealth it managed, even when this wealth was held—as is frequently the case—through intertwined shell companies incorporated in Panama and similar offshore havens. This allows us to link wealth to its actual owners. Fourth, while owning bank accounts in Switzerland or offshore companies in Panama is not illegal per se, the leaked file matched to tax returns offers a clear-cut way to identify illegal tax evasion: taxpayers who reported their HSBC accounts were not evading; those who did not were. In practice, about 95% of all the individuals on the HSBC list that we can match to a tax return did not report their Swiss bank account.

Because Norway, Sweden and Denmark have or used to have a wealth tax, they collect detailed individual information on wealth; we match the HSBC leaked file and random audit

data to these administrative wealth records to study how the prevalence of tax evasion varies with wealth. The analysis yields a number of spectacular findings. First, as reported in Section III, we find that in randomized audit data, the probability to deliberately evade taxes is very low up to the 95th percentile of the wealth distribution but then rises with wealth to about 5% in the top 1%. The prevalence of tax evasion then appears to remain stable (at around 5%) within the top 1%. Thus, rates of detected tax evasion are generally low; consistent with the analysis of Kleven et al. (2011), they are almost zero for income that is third-party reported, and only become significant for self-reported income (e.g., from self-employment), which is more common at the top.

Second, and in sharp contrast with what random audit data show, we find that the probability to appear in the HSBC list is almost zero up to the 99th percentile but then rises very steeply within the top one percent (Section IV). For the top group we consider—the top 0.05% of the wealth distribution—we find that about 0.7% of taxpayers were evading taxes through HSBC Switzerland in 2006. Because this bank had about 2% of the offshore wealth management market at the time of the leak, in our benchmark scenario we estimate that about 35% of the wealthiest Scandinavians had unreported offshore accounts—considerably more than the rates of evasion implied by random audits. We stress at the outset that there uncertainties on the exact fraction of the offshore tax evasion market HSBC Private Bank Switzerland captured, and therefore our estimates of evasion rates at the top come with a margin of error. We construct lower and upper bounds and find we can bound the evasion rate for the top 0.05% at 20% from below—four times higher than in random audit studies. Evidence from other leaks—the Panama Papers—and from tax evaders who voluntarily declare assets previously hidden in offshore tax havens all paint the similar picture: offshore tax evasion is extremely concentrated and rises steeply with wealth up to the very top. Combining random audits with leaks, in our benchmark scenario we estimate that the fraction of the population that evades taxes rises from 1%-2% in the bottom 95% of the wealth distribution to about 40% in the top 0.05%.

In Section V, we consider potential explanations for this finding. Existing models focus on the rational behavior of a tax evader under uncertainty (Allingham and Sandmo, 1972), what can be seen as the demand side for tax evasion services. Evasion is high when the probability to be detected is low or when penalties are low, and the effect of tax rates is nil or ambiguous. These models do not provide a direct explanation for the sharp gradient in evasion with wealth. We argue that to understand this gradient, it is necessary to consider the supply side of tax evasion services. We introduce such a model. Swiss banks decide on the number of clients they

serve by weighing the benefits (fees proportional to the hidden wealth they manage) against the cost (the probability to be caught facilitating tax evasion—which rises with the number of clients served—and having to pay a fine). We derive and calibrate a closed-form expression for the fraction of the population served when wealth is Pareto distributed. The higher inequality, the lower the quantities of evasion services supplied—when inequality is very high, as is the case for wealth, it is optimal for banks to only supply tax evasion services to the top 1% or smaller groups.

We discuss two types of implications of our results. First we analyze how accounting for offshore wealth affects measured wealth inequality (Section VI). We illustrate this with the case of Norway where high quality, long-run time series of top wealth shares exist. Because offshore wealth is extremely concentrated—virtually all of it belongs to the top 1%, and about 80% belongs to the top 0.1%—taking it into account lifts top wealth shares significantly. By our estimate about 40% of the Norwegian households in the top 0.1% hid wealth abroad in the mid-2000s and these tax evaders hid on average about 50% of their assets. Thus the top 0.1% hid 20% of its total wealth, and including this previously unmeasured wealth increases the top 0.1% wealth share from 8% to 10%. For the top 0.01%—the wealthiest 300 or so Norwegian households—taking offshore wealth into account increases reported wealth by more than 30%. Our results highlight the need to move beyond tax records to capture income and wealth at the very top, even in countries where tax compliance is generally high. They also suggest that tax data may significantly under-estimate the rise of wealth concentration over the last four decades, as the world was less globalized in the 1970s, it was harder to move assets across borders, and offshore tax havens played a less important role. Because most Latin American, and many Asian and European economies own much more wealth offshore than Norway, the results found in Norway are likely to be lower bound for most of the world's countries.

Last, we consider the implications of our results for public finances in Section VII. Should tax evasion become impossible, would wealthy individuals pay significantly more taxes? The answer largely depends on how substitutable illegal tax evasion and legal tax avoidance are. To address this question, we analyze the behavior of a large sample of Norwegians who voluntarily disclosed previously hidden wealth in the context of a tax amnesty. In an event study design, we find that after voluntarily reducing tax evasion, tax evaders start avoiding more. Taken together, our findings on inequality and tax avoidance highlight the need to move beyond tax data to capture income and wealth at the top, even in countries where tax compliance is generally high. They also suggest that collecting extra tax revenue from the wealthy would require curbing the supply

of evasion services (e.g., by making it more costly and risky for banks to serve tax evaders) and simultaneously closing loopholes to prevent substitution from evasion to avoidance.

2 Related Literature

2.1 Tax Evasion in Random Audit Data

Our paper first contributes to the literature on tax evasion. A large part of this literature is based on randomized audits. In the US, the key data source is the Taxpayer Compliance Measurement Program (TCMP), now known as the National Research Program, a thorough, periodic audit of stratified sample of tax returns. Audit studies such as the TCMP consistently find large rates of tax evasion for self-employment and small business income, for which the absence of third-party reporting makes tax evasion easy. For instance, the IRS (2012) estimates that half of the U.S. tax gap owes to unpaid self-employment Social Security/Medicare taxes plus unpaid taxes on small business profits. Also based on stratified random audits, Kleven et al. (2011) find that 44.9% of Danish self-employed evade taxes. These findings are confirmed by a large literature on tax evasion by the self-employed (Pissarides and Weber, 1989; Feldman and Slemrod, 2007; Artavanis et al., 2015) which typically finds that the true income of self-employed is on average about 1.5 times their reported income. Because tax evasion by self-employed individuals and small businesses is thought to be reasonably well captured by audits, estimates for this form of evasion are built into the national accounts: the evaded income—as estimated using random audits—is included in official GDP statistics.

A number of studies have studied the distributional implications of tax evasion. Bishop, Formby, and Lambert (2000) and Johns and Slemrod (2010) use randomized audit micro-data to study how detected tax evasion affects income inequality in the United States. Nygård, Slemrod and Thoresen (2016) study how unreported goods and services transactions affect income inequality in Norway. Our main contribution compared to these studies is that we provide evidence on a form of tax evasion (offshore tax evasion) that is not captured adequately in audit data and that turns out to play a critical role; we are also able to analyze how evasion varies with wealth rather than taxable income.

2.2 Offshore Wealth and Offshore Tax Evasion

It is widely acknowledged that audits data do not properly capture the most sophisticated forms of tax evasion (e.g., Slemrod, 2007; Kleven et al. 2011). Tax authorities simply do not have the

resources necessary to investigate such cases in their random audit programs. Audits consist of line-by-line information about what the taxpayer reported and what the examiner concluded was correct. Examiners can check that taxpayers duly report the capital income earned through domestic financial institutions, because these institutions automatically and truthfully report data to the tax authority, but they cannot check that they duly report income earned through offshore financial institutions, because they cannot audit all the banks from all over the world.

As a result, while tax evasion by small businesses is a known unknown, offshore evasion is largely unknown. The wealth hidden offshore escapes published measures of national wealth, and the income it generates is not included in national income (contrary to evaded self-employment income), a problem long recognized by statisticians (see, e.g., Bertaut et al 2007; ECB, 2001). Because audit data fail to capture evasion on offshore capital income, and this income is likely to be very concentrated, audits are unlikely to paint a realistic picture of tax evasion by the wealthy.

There have been attempts at estimating the macroeconomic amount of wealth held in offshore tax havens. Zucman (2013) estimates that the equivalent of 10% of the world's GDP is held in tax havens globally. Pellegrini et al. (2016) have a similar estimate. According to the US Treasury International Capital statistics, offshore centers manage close to 10% of all the listed equities of US companies, against less than 1% in the early 1940s (Zucman, 2014). Two US Senate (2008, 2014) reports found that up to 2008, 85 to 95 percent of US-owned accounts at UBS and Credit Suisse in Switzerland were undeclared. Roussille (2015) finds that more than 80% of the wealth held by Europeans in Switzerland is undeclared in 2013, and more than 90% was before 2010. Until recently, however, there was no micro data to study who owns the wealth hidden offshore.

2.3 Long-Run Trends in Inequality

Over the last fifteen years, there has been renewed interest for the long-run evolution of the distribution of income and wealth. Following the pioneering work of Kuznets (1953) and Atkinson and Harrison (1978), a number of studies have used tax data to construct top income and wealth shares for many countries.[1] A key concern raised by the use of tax returns to measure inequality, and indeed one of the main reasons why tax data have for a long time been viewed with skepticism, is tax evasion.

[1]See, e.g., Piketty 2001, Piketty and Saez 2003, Saez and Zucman 2015; for surveys see Atkinson et al. 2011, Alvaredo et al. 2013. Top income and wealth share series are collected in the World Wealth and Income Database (Alvaredo et al. 2015), `http://wid.world`.

The lack of data about tax evasion by the very wealthy is particularly problematic for inequality studies because most of the cross-country and historical variations in inequality come from the very top of the distribution. In the United States, the share of wealth owned by the top 0.1% (with more then $20 million in net wealth today) has increased from about 7% in in the late 1970s to 22% today, while the wealth shares of the groups just below has remained mostly flat (Saez and Zucman 2016). Similarly, the decline in income concentration seen in France and most developed countries during the first half of the twentieth century owes to the fall of the "rentiers," the top 1%, and not the next 9% (Piketty 2003).

Tax evasion can bias cross-country and time-series comparisons of top income and wealth shares in multiple directions. Countries with high rates of capital taxation (such as Norway) may hide more wealth than countries with low rates on capital (such as Japan and the US), leading to artificially low measured inequality in Europe; but this incentive effect may be offset by differences in social norms and overall level of compliance. Since 2013, many tax havens have promised to automatically exchange bank information with foreign tax authorities, This recent policy change may have led to more reporting and caused measured inequality to increase spuriously.[2] The literature on inequality has discussed the potential problems raised by tax evasion (e.g., Piketty 2001; Atkinson, Piketty, Saez, 2011), but until recently there was little data that would allow to systematically quantify it. Larudee (2016) investigates the extent to which capital flight to Switzerland can explain the decline in the French top 1% income share between the two world wars. Roine and Waldenström (2009) is the only paper that focuses on the distributional implications of hidden wealth for the recent period. They use an indirect method—residual flows in the balance of payments and financial accounts—to estimate the amount of wealth hidden by Swedish residents, and assume that this wealth primarily belongs to the top. Under this assumption, the share of wealth owned by the top one percent rises from about 20 percent in the 2000s to over 30 percent.

3 How Tax Evasion Varies with Wealth in Random Audits

We start our analysis by estimating how the prevalence of detected tax evasion varies with wealth, before moving to the distribution of—traditionally undetected—offshore tax evasion.

[2]However, because foreign banks cannot easily be audited by foreign governments, and because some of them may still find it profitable to sell tax evasion services, it is unclear how truthful this form of reporting will turn out to be.

3.1 Danish Random Audit Data

Our baseline measure of detected tax evasion comes from high quality, stratified random audits conducted by the Danish Tax Authority (SKAT). The first wave of this program, for tax year 2006, was studied by Kleven et al. (2011); we refer to them for a description of the Danish institutional context. In this paper, we analyze the three subsequent waves of audits, which were conducted for tax years 2008, 2010, and 2012. The random audit program works as follows. Once every two years, SKAT randomly selects a sample of self-employed individuals and a sample of individuals who are not self-employed (including, e.g., wage-earners and retirees). The sampling rate is higher for the self-employed (who are relatively more numerous at the top of the distribution), and in both groups taxpayers with complex tax returns are over-sampled, ensuring that sufficiently many wealthy individuals end up being audited. Our final sample pools tax years 2008, 2010, and 2012 and includes 18,985 randomly audited taxpayers (6,223 self-employed and 12,762 non self-employed), of which 663 turned out to be in the top 1% of the wealth distribution.[3]

As explained in Kleven et al. (2011), the audits are comprehensive and examine every line item on the tax return using several verification procedures and sources: data available internally at SKAT; information provided by employers, banks, credit card companies, and other financial institutions; and supporting documentation requested to the tax-payers themselves. This procedure enables to detect a wide range of errors, from mistakes in the claiming of deductions (e.g., for alimony or commuting expenses) to mis-reporting of income that is not declared by a third party (e.g., taxable fringe benefits). In addition, the tax authority compares disposable reported income to estimates of consumption and the change in wealth over the tax year—drawing on high quality administrative wealth records—to attempt to detect unreported labor market activity. SKAT improved its audit technique after 2006—the first year it conducted the random audit program—and now detects more errors. While mistakes were found for 10.7% of all individuals audited in 2006 (Kleven et al., Table 2 col. 2 line 2), the error rate climbed to 12.5% in 2010 and 2012.[4] Despite the amount of resources devoted to them, the random audits do not capture all forms of evasion, however. They entirely miss, in particular, offshore evasion. In the rare cases when the auditor might suspect such type of evasion, the case is transferred to

[3]Overall, 0.15% of the entire adult Danish population is randomly audited each year in the context of this program. The empirical sampling rate is 3.5 times higher for taxpayers in the top 1% of the distribution (0.53%); see Appendix Table S.1.

[4]In addition to an improvement in the conduct of audits, part of this increase may reflect a real decline in compliance between 2006 and 2010.

a specialized unit within SKAT that has the skills to conduct a specific investigation. Whatever is found at the end of this long process is not included in the result of the random audit study, as this would delay the publication of the results for too long.

The Danish random audit data we use have two unique strengths. First, they contain an indication of whether mistakes found on the tax return were deliberate or not. At the end of each audit, SKAT rates the compliance of the audited taxpayer on a scale from 0 to 6. Taxpayers with ratings between 0 and 2 are considered to be deliberate tax evaders, whereas those with ratings between 3 and 6 are considered tax compliant, although some may make involuntary errors (see SKAT, 2016). The distinction between deliberate tax evasion and unintentional error is a critical one, as wealthy taxpayers with complex tax returns and income that is not all third-party-reported may be more likely to make unintentional mistakes and yet no more likely to be deliberately breaking the law. To our knowledge the literature so far has not been able to separate deliberate from unintentional mistakes. When focusing on deliberate evaders, the fraction of individuals caught evading taxes falls from 12.5% to 1.5% in 2012.

Second, the Danish audit data can be matched to population-wide administrative data that contain information on wealth. This enables us to compute the rank of each audited taxpayer in the distribution of Danish household wealth.

3.2 Danish Wealth Data

We compute the distribution of Danish wealth following standardized, international guidelines codified in Alvaredo et al. (2016) in order to make the results comparable across countries—and in particular comparable to Sweden and Norway studied below. We distribute 100% of the macroeconomic amount of household wealth at market value recorded in Denmark's national accounts. This total includes all the non-financial and financial assets that belong to Danish resident households, minus debts. It includes in particular all funded pension wealth and excludes consumer durables and valuables. In 2012, household wealth in Denmark amounted to 379% of national income; average household wealth per adult was US$ 242,000 (using market exchange rates to convert Danish kronor to dollars), a level comparable to that of the United States (US$ 234,000).

Our starting point to compute the wealth distribution is the detailed individual wealth records maintained by Danish authorities, a legacy of the period (before 1997) when Denmark had a wealth tax. Although these data are generally of very high quality, they suffer from a number of limitations. First, they do not capture forms of wealth that are not reported by

third parties, most importantly unlisted corporate equities and non-corporate business assets. Second, they disregard funded pension assets, which are large in Denmark—43.5% of all Danish household wealth in 2012, cf. Appendix Table D.1c. We deal with these issues as follows. First, we impute pension wealth as in Saez and Zucman (2016) by assuming that 40% belongs to wage-earners and 60% belongs to retirees; pension wealth allocated to wage-earners is then distributed proportionally to wage income and pension wealth allocated to retirees proportionally to the income derived from pension funds. This imputation introduces a small noise at the micro-level which is irrelevant for our purposes. Second, we impute unlisted equities by capitalizing dividends earned from unlisted corporations. The capitalization rate is the ratio of the market value of unlisted firms (as recorded in the national accounts) by the flow of unlisted dividends reported on tax returns. Third, we impute non-corporate business assets by capitalizing business income in the same way. After these corrections we capture almost 100% of the total household wealth reported in the national accounts; we fill in the residual gap by scaling up multiplicatively registered assets, separately for each asset class—listed equities, fixed-income claims, housing assets, and debts.

Our unit of analysis is the tax unit, defined as either a single person aged 20 or above or a married couple. We disregard children below 20 who own negligible wealth. Summing the wealth of married spouses addresses the fact that jointly-owned assets are allocated to only one individual in the registry. Another possibility would be to split the wealth of married spouses 50/50; we prefer to conduct all our analysis at the tax unit level, as spouses are typically liable for the taxes owed by their partner.

3.3 Tax Evasion and Wealth in Random Audits

We start the analysis in Figure 1 by showing how the prevalence of detected tax evasion varies with wealth. We restrict our definition of tax evasion to tax units caught deliberately evading taxes, i.e., where the audit has given rise to an adjustment to taxable income and the error is evaluated as being deliberate or based on an improbable interpretation of the law. We match tax evaders to their rank in the wealth distribution in the tax year of the audit, and we pool the three waves of audits (for tax years 2008, 2010, and 2012). Each observation is weighted by the corresponding audit weight. There are three main findings.

First, rates of (detected) deliberate evasion are low across the board. On average, only 1.35% of all audited Danes were found deliberately evading taxes. Even where evasion is the highest, the rate of detected deliberate tax evasion does not exceed 5%. Second, evasion appears to be

rising around the 95th percentile of the wealth distribution. The evasion rate is below 2% for all groups below the 95th percentile—which corresponds to US$ 1.3 million in net wealth in 2012. Past this threshold, it starts rising, to about 5%. Third, evasion rates appear to stabilize (to around 5%) within the top 1%. Evasion rates are also slightly higher than average in the bottom decile of the wealth distribution, which includes tax units with large debts and potentially mismeasured assets (hence artificially low net wealth). Consistent with the analysis of Kleven et al. (2011), we find that rates of tax evasion are almost zero for income that is third-party reported, and only become significant for self-reported income (e.g., from self-employment), which is more common at the top.

4 How Tax Evasion Varies with Wealth in HSBC Leaks

In this Section we introduce the main dataset we use to supplement random audit studies: micro-data leaked from HSBC Private Bank Switzerland.

4.1 HSBC Private Bank Switzerland Leak

In 2006 and 2007 a systems engineer, Herve Falciani, extracted the list of HSBC Private Bank's clients and of their holdings. Falciani turned the data over to the French government in 2008 who shared it with foreign administrations. The newspaper *Le Monde* obtained a version of the tax authority data and shared it with the International Consortium of Investigate Journalists. ICIJ assembled a global team of journalists and in January 2015 published the results of its investigation, called "Swiss leaks."

The documents leaked by Falciani include the complete internal records—including the names and assets—of the more than 30,000 clients of this Swiss bank in 2006-2007, a large fraction of whom were evading taxes. We are able to match the names on the HSBC list to individual tax data in three countries, Norway, Sweden, and Denmark. In Denmark and Norway, the data we use are from the version of the HSBC list shared by the French tax authority with foreign administrations. In Sweden, the data are from the list obtained by the ICIJ and affiliated journalists. To our knowledge, it is the first time that leaked micro-data matched to tax returns are used to study tax evasion. These data allow us to provide the first evidence on how offshore tax evasion is distributed, enabling us to fill what is perhaps the key gap of random audit studies.

The leaked HSBC file has four key strengths. First, the leak is not the result of specific enforcement effort by tax authorities targeted at HSBC; it can be seen as a random event.

Second, it involves a major—and, the available evidence suggests, representative—player in the offshore wealth management industry, a fact we use to make inference about the overall prevalence of offshore tax evasion in the entire population. At the time of the leak, HSBC Private Bank Switzerland, the Swiss subsidiary of the banking giant HSBC, managed $118 billion in wealth, about 4% of all the foreign wealth held in Switzerland; it had clients from all over the world; and as we show below, the wealth it managed was similarly distributed across countries as that managed by the overall Swiss banking industry. Third, HSBC Switzerland recorded the name of the beneficial owners of the wealth it managed, even when this wealth was held—as is frequently the case—through intertwined shell companies incorporated in Panama and similar offshore havens. This allows us to link wealth to its actual owners. Fourth, while owning bank accounts in Switzerland or offshore companies in Panama is not illegal per se, the leaked file matched to tax returns offers a clear-cut way to identify illegal tax evasion: taxpayers who reported their HSBC accounts were not evading; those who did not were. In practice, about 95% of all the individuals on the HSBC list that we can match to a tax return did not report their Swiss bank account.

4.2 Who Hides Wealth at HSBC?

We match the names on the HSBC list to population-wide registries in Norway, Denmark, and Sweden. We construct the full distribution of household wealth in Norway and Sweden in the exact same way as we do in Denmark. That is, following standardized, international guidelines codified in Alvaredo et al. (2016) we distribute 100% of the macroeconomic amount of household wealth at market value recorded in the national accounts. To do so, we start from the high quality administrative wealth data maintained by Scandinavian administrations and we supplement these data by imputing pension wealth, business wealth (by capitalizing business income), and unlisted shares (by capitalizing dividends). Table 1 shows the distribution of wealth in Norway, Sweden, and Denmark (excluding offshore wealth). The three countries have very similar levels of inequality, comparable to those of France (Garbinti, Goupille and Piketty) and way lower than US inequality levels (Saez and Zucman, 2016).

Next, we aggregate Denmark, Norway and Sweden into a single Scandinavian aggregate as follows. We collapse each country's population-wide data into small bins (of as few as 10 tax units at the top), compute average, minimum, and maximum wealth in each bin, and append the three collapsed dataset. This makes it possible to study the distribution of wealth and tax evasion in Scandinavia as a whole. The individual country distributions turn out to all be very

similar to the overall Scandinavian distribution, as reported in the Appendix.

Figure 2 shows that in sharp contrast to what random audit data show, the probability to evade taxes rises very sharply with wealth at the top of the distribution. The probability to have an undeclared account at HSBC Switzerland is almost zero up to the 99th percentile. It then rises very steeply within the top one percent. For the top group we consider—the top 0.05% of the wealth distribution—we find that about 0.7% of taxpayers were evading taxes through HSBC Switzerland in 2006. While in random audit studies, households in the top 0.05% are as likely to be found evading taxes than households in the bottom of the top 1%, in the leaked data they are 7 times more likely to conduct offshore tax evasion.

4.3 How Representative is HSBC Switzerland?

There are two potential issues when making inference from the leaked HSBC data about the overall prevalence of offshore tax evasion in the Scandinavian population. First, it is possible that HSBC Switzerland catered to only particularly wealthy individuals and that offshore tax evasion is generally less skewed than what is shown in Figure 2. Second, there are uncertainties on what fraction of the offshore wealth management market HSBC Switzerland captured. We discuss these two issues in turn.

First, a body of evidence suggests that the probability to use offshore financial institutions—beyond HSBC Switzerland—to conceal wealth rises sharply with wealth up to the very top of the distribution. The additional evidence comes from two sources: a micro dataset on who voluntary declares assets previously hidden in offshore tax havens in the context of a tax amnesty in Norway, and another leak—this time from the Panama firm Mossack Fonseca.

The Norwegian tax authority offers taxpayers who have evaded taxes the option to voluntary disclose wealth hidden offshore. This amnesty program has been in place since 1950, but received renewed attention in 2007, when as part of a divorce settlement the son-in-law of the mayor of Oslo revealed that his step-father had undeclared bank accounts in Switzerland. We have access to all disclosures made since 2006. Just like for the HSBC leak, we match voluntary disclosers to detailed tax return and socio-economic characteristics, including rank in the full distribution of wealth at market value. The relatively large sample size (1,428 individuals, the equivalent to close to 100,000 individuals in the United States scaling for population size) ensures that this source provides useful information on the owners of hidden wealth. As reported in the Online Appendix, the probability to disclose previously hidden assets rises very sharply with wealth exactly like in the HSBC leak. One caveat here is that from a purely logical perspective, it

is possible that only the wealthiest tax evaders self-selected into disclosing previously hidden assets. In the end, however, it is striking to see that the prevalence of offshore evasion looks strikingly the same in the HSBC leak—where thre is no self-selection problem—and the amnesty dataset.

Another piece of evidence that offshore evasion rates rises steeply with wealth comes from the Panama Papers. In the Spring of 2016, the ICIJ published leaked documents containing the names and addresses of the owners of shell companies created by Mossack Fonseca, mostly in Panama and in the British Virgin Islands. The probability to appear in the Panama Papers again sharply rises with wealth at the top, as shown in the Appendix. One caveat here is that we cannot prove that all Norwegians whose name appears in the Panama Papers were evading taxes, in contrast to the HSBC leak where we can link names to tax returns, and in contrast to the amnesty data that definition only include tax evaders. There are legal uses of shell companies, and we do not have information on whether the shell companies were used for tax evasion purposes. At a minimum, however, these results further confirm the notion that the use of tax havens steeply rises with wealth.

We now turn to the more important question of what share of the offshore wealth management market HSBC Switzerland captured at the time of the leak. First, Table 2 shows that HSBC Private Bank Switzerland managed about 4% of all the wealth managed by Swiss banks, US$ 118.4 billion out of US$ 2,912 billion. The US$ 118.4 billion figure is the official value published by HSBC. The amount of foreign wealth managed by all Swiss banks (US$ 2,912 billion) is from the official statistics published by the Swiss central bank, which reports on the stock of wealth managed by Swiss banks on behalf of foreigners on a monthly basis (see Zucman, 2013, 2014, 2015 for a detailed analysis of this unique, high quality dataset). The Swiss central bank also provides an annual country-by-country breakdown of this total, which enables us to estimate the amont of wealth managed by Swiss banks on behalf of Danish, Swedish, and Norwegian residents. As reported in Table 2, we find that HSBC Private Bank Switzerland also managed about 4% of all the Swiss wealth belonging to Scandinavians. This result suggest that HSBC was not catering to Scandinavian households in particular.

More generally, the distribution of the wealth managed by HSBC Switzerland is remarkably similar to that of the wealth managed by all Swiss banks. A country-by-country breakdown of the wealth managed by HSBC Switzerland in 2006-2007 based on an exhaustive exploitation of the file leaked by Falciani is published by the ICIJ at `http://projects.icij.org/swiss-leaks/`. Figure 3 compares the country distribution of the foreign wealth managed by

HSBC Switzerland to the country distribution of all the foreign wealth managed by all Swiss banks. Some countries are slightly over-represented in the HSBC leak, most notably Venezuela, the United States, and Brazil. This can be explained as follows. In 1999, HSBC Switzerland merged with the Republic National Bank of New York and Safra Republic Holdings SA, two private wealth management banks with a strong customer base in the United States and Brazil respectively. In addition, according to the ICIJ, the biggest account at HSBC Switzerland was a US$ 11.9 billion account registered in the name of Venezuela's National Treasurer. Despite these idiosyncrasies, the share of HSBC's wealth owned by each country i (s_i^H) is overall tightly linked to its share of all Swiss banks' wealth s_i. A regression of s_i^H on s_i has a slope of 1.00 and R^2 of 0.80.

In sum, there is a body of evidence suggesting that at the time of the leak, HSBC Private Bank Switzerland was representative of the overall offshore wealth management industry, both in terms of how wealthy its clients were, and in terms of where its clients came from.

4.4 Aggregating Random Audits and Leaks

In order to infer the overall prevalence of offshore tax evasion in the Scandinavian population from the HSBC leak, we proceed as follows. First, we start from Zucman's (2013) estimate of the amount of wealth held in all the world's offshore tax havens at the end of 2006—US$ 6.3, or about twice the amount held in Swiss banks alone. This estimate implies that HSBC Switzerland managed about 2% of all the world's offshore wealth—and 2% of the estimated wealth held offshore by Scandinavian countries. In our benchmark scenario, we therefore multiply the number of HSBC customers matched to tax return by 50 in order to obtain an estimate of the overall rate of offshore evasion. As Figure 4 shows, this implies that about 35% of the wealthiest Scandinavians had unreported offshore accounts—considerably more than the rates of evasion implied by random audits. Third, since by construction random audits do not include offshore tax evasion, we add up the rates of tax evasion found in random audits with those estimated from the HSBC leak. In our benchmark scenario, this implies that the fraction of households evading taxes rises from 1%-2% in the bottom 95% of the wealth distribution to about 40% in the top 0.05%.

The main source of uncertainty in this computation is what was the exact fraction of the offshore tax evasion market that HSBC Private Bank Switzerland captured at the time of the leak. There is no doubt that HSBC Switzerland managed 4% of all the offshore wealth managed by Swiss banks, as this figure is based on official data. Therefore in our lower bound scenario,

we simply multiply the number of Scandinavian HSBC evaders we can match to a tax return by 25, i.e., we assume that there are no tax evaders with undeclared offshore accounts in tax havens other than Switzerland. As Figure 5 shows, this implies that the evasion rate for the top 0.05% is at least 20%—four times higher than in random audit studies. In our upper bound scenario, we assume that all the HSBC Switzerland accounts connected to Scandinavia and obtained by the tax authorities but that we could not match to a tax return belong to tax evaders and are similarly distributed as the accounts that were matched. If that is true, the prevalence of tax evasion rises to 60% at the very top of the wealth distribution.

5 A Model of Tax Evasion and Inequality

How can we explain the sharp gradient of evasion with wealth that we estimate? In this Section, we argue that a simple model of the supply side of tax evasion services can account for this pattern.

We consider a population of mass one where wealth y is distributed according to some probability density function $f(y)$. Consider first the case of a monopolistic firm that sells tax evasion services; think for instance of a Swiss bank (or a cartel of Swiss banks). The bank charges θ per dollar of wealth that it hides from tax authorities. To simplify the analysis and focus on the main mechanism at play, we assume the bank faces an infinitely elastic demand, hence takes θ as given and optimizes on the number of clients it serves. The assumption of a perfectly elastic demand is irrelevant for our qualitative findings, and any model of the demand side of evasion services can easily be nested into our model.[5] Each of the bank's client has a probability λ to be found evading taxes. When the bank offers tax evasion services to the fraction of the population $s = 1-F(y) = \int_{z \geqslant y} f(z)dz$ that has wealth greater than y, it manages $k(s)$ in wealth hence earns $\theta k(s)$ in revenue. It also has a probability $\lambda(s)$ that at least one of its clients is found evading taxes. When a client is found evading, this reveals that the bank is breaking the law and triggers a fine. The fine $\phi k(s)$ is proportional to the amount of wealth managed. We assume the bank is risk-neutral, hence it maximizes profits

[5] Assuming a less than perfectly elastic demand would reduce the equilibrium fraction of the population that evades taxes, reinforcing the main conclusion of the analysis, but it would make the formulas below slightly more complicated as θ would now depend on s. Assuming infinitely elastic demand means the bank acts in effect as on a competitive market. The advantage of this simple market structure (as compared to assuming there is a continuum of banks that acts competitively) is that we do not need to model how banks split customers, and can instead model the market supply curve directly. More importantly, support for the monopolistic assumption comes from the fact that Swiss banks historically had a cartel agreement, the Convention IV of the Swiss Bankers Association, that strictly regulated fees and prevented non-linear pricing (e.g., banks setting lower prices for wealthier customers); see Zucman (2015).

$$\pi(s) = \theta k(s) - \lambda(s)\phi k(s)$$

An interior optimum for s satisfies the first-order condition $d\pi/ds = 0$, which is $\theta k'(s) = \lambda'(s)\phi k(s) + \lambda(s)\phi k'(s)$.[6] Denote by $\epsilon_\lambda(s) = s\lambda'(s)/\lambda(s)$ the elasticity of the detection probability with respect to s, and by $\epsilon_k(s) = sk'(s)/k(s)$ the elasticity of the amount of hidden wealth managed with respect to s. The first-order condition can be rewritten

$$\theta = \left(\frac{\epsilon_\lambda(s)}{\epsilon_k(s)} + 1\right)\phi\lambda(s)$$

The left-hand side is the marginal revenue of servicing an extra percent of the population, and the right-hand side is the marginal cost. If $\epsilon_\lambda(s) = 0$, i.e., in the simple case where the probability for the bank to be found breaking the law does not depend on the size of its clientele, then the marginal cost is simply $\phi\lambda(s)$, the probability to be detected times the associated fine. But in the general case, the probability to be found breaking the law is endogenous and rises with s, hence the marginal cost is greater than $\phi\lambda(s)$.

When λ is small, $\epsilon_\lambda(s)$ is equal to 1. Furthermore, in the case where wealth is Pareto distributed with a Pareto coefficient $a > 1$, then $\epsilon_k(s)$ has the remarkable property that it does not depend on s either; it is equal to one over the inverted Pareto-Lorenz coefficient $a/(a-1)$. A high a corresponds to a relatively equal distribution of wealth; a low a corresponds to an unequal distribution; when $a \to 1$, inequality tends to infinity. It is well known that income and wealth tend to follow Pareto distributions at the top, and a large literature estimates Pareto and inverted Pareto-Lorenz coefficient over time and across countries (see, e.g., Atkinson, Piketty and Saez 2011, Table 3 and Figures 12, 13, 14 and 15). A typical value of a for the wealth distribution is $a = 1.5$, which corresponds to an inverted Pareto-Lorenz coefficient $a/(a-1)$ equal to 3. When wealth is Pareto distributed, the supply curve of tax evasion services has thus a simple closed-form expression.

Proposition 1. *When wealth is Pareto-distributed, the supply curve of tax evasion services is:*

$$\theta = \left(\frac{a}{a-1} + 1\right)\phi log\left(\frac{1}{1-p}\right)s$$

Where $p = 1 - e^{-\lambda}$ is the probability that the bank is caught when it serves the full population.

[6] Our formulation of the bank's problem assumes that the bank starts by servicing the wealthiest individuals and then moves down the wealth distribution up to the point where the marginal revenue equals the expected marginal cost. Given that each client has the same probability λ to be found evading taxes but that wealthier clients generate more revenue, this is the optimal sequence for the bank to follow. Second-order condition....

Proof. The probability $\lambda(s)$ that at least one client is found evading when the bank serves s clients is $1 - e^{-\lambda s}$, which is equal to λs when λ is small. Hence $\epsilon_\lambda(s) = s\lambda'(s)/\lambda(s) = 1$.

By definition $k(s) = \int_{z \geqslant y} zf(z)dz$ and $s = 1 - F(y) = \int_{z \geqslant y} f(z)dz$, therefore $k'(s) = y$ and

$$\frac{\epsilon_\lambda(s)}{\epsilon_k(s)} = \frac{1}{sk'(s)/k(s)} = \frac{\int_{z \geqslant y} zf(z)dz}{(1 - F(y))y}$$

When y is Pareto distributed above wealth y_{min} with Pareto coefficient a, then the probaility density function is $f(y) = ay_{min}^a/y^{1+a}$, the survival function is $1 - F(y) = (y_{min}/y)^a$, and straightforward integration shows that

$$\int_{z \geqslant y} zf(z)dz = \frac{a}{a-1} y(1 - F(y))$$

Therefore:

$$\frac{\epsilon_\lambda(s)}{\epsilon_k(s)} = \frac{a}{a-1}$$

That is, the Pareto distribution has the unique property that ϵ_k is constant and equal to one over the inverted Pareto-Lorenz coefficient $a/(a-1)$.

Finally, since each client has an instantaneous probability λ to be detected, the probability that at least one person is detected when the bank services the full population $(0, 1)$ is $1 - e^{-\lambda}$. We denote this probability by p, hence $\lambda = log(1/(1-p))$ and

$$\lambda(s) = \lambda s = log\left(\frac{1}{1-p}\right) s$$

□

The supply curve is upward slopping with $s = 1 - F(y)$. The higher the price charged per dollar hidden, the greater the fraction of the population covered (i.e., the lower the wealth y of the marginal client). When the demand for evasion services is infinitely elastic at price θ, the equilibrium fraction of the population that evades taxes is:

$$s = 1 - F(y) = \frac{\theta}{\left(\frac{a}{a-1} + 1\right) \phi log(\frac{1}{1-p})}$$

When the probability to be found p tends to 1, the fraction of the population evading taxes goes to 0. When the fine ϕ tends to infinity, the fraction s also goes to 0. Most importantly, when inequality rises (a falls), then s falls; it tends to zero when inequality becomes infinite ($a \to 1$). In a very unequal world, banks will tend to service only very few clients. Since wealth is very concentrated, this can explain why evasion is almost zero everywhere except at the very

top of the distribution. The intuition for this result is the following. When inequality is high ($a \to 1$), a handful of individuals own the bulk of wealth; they generate a lot of revenue for the bank and are unlikely to be detected. Moving down the distribution would mean reaching a big mass of the population that would generate only relatively little additional revenue but would increase the risk of detection a lot; it is not worth it. Conversely, when inequality is low ($a \to \infty$), the bank needs to move down a lot to hit a significant mass of the population.

As a simple calibration of the model, assume that 10% of the population owns more than \$1 million, a threshold above which wealth is Pareto distributed.[7] Assume the cartel of banks charges an annual fee of $\theta = 1\%$ on the wealth it manages, there is a $p = 10\%$ annual chance that it is caught if it serves all individuals with more than \$1 million,[8] and if caught it has to pay a fine equal to $\phi = 20\%$ of the wealth managed.[9] When inequality is low ($a = 3$), the cartel serves all the clients with more than \$1.7 million in wealth, which corresponds to $s = 19\%$ of the population above \$1 million. When inequality is high ($a = 1.2$), it only serves clients with more than \$9.4 million in wealth, which corresponds to $s = 6.8\%$ of the population above \$1 million.[10] In other words, in the low inequality case the cartel supplies tax evasion services to the top 2% of the population, and in the high inequality case it serves the top 0.6%.

6 Corrected Inequality Statistics Taking Into Account Tax Evasion

In this Section, we analyze how accounting for offshore wealth affects measured wealth inequality.

Figure 7 contrasts the distribution of wealth in Scandinavia (excluding hidden offshore assets) with the distribution of offshore wealth. Offshore wealth is extremely concentrated: about 80% of it belongs to the top 0.1% of the distribution, against 10% for wealth that is not hidden. Because offshore wealth is so concentrated, taking it into account lifts top wealth shares significantly.

We illustrate this with the case of Norway where high quality, long-run time series of top

[7] In the United States, the threshold for the top 10% of the wealth distribution is about \$800,000 today; see Saez and Zucman (2016).

[8] The probability that the bank is actually caught is $1 - (1-p)^s$ which is less than p.

[9] In 2014, Credit Suisse was fined \$2.6 billion for conspiracy to aid and assist U.S. taxpayers in filing false returns. According to a report by the U.S. Senate (2014), Credit Suisse had about \$22,000 U.S. Customers with 12 Billion Swiss Francs in assets in 2006 (about 10 billion dollars in 2006), the vast majority of which were undeclared.

[10] A Pareto coefficient $a = 3$ corresponds to an inverted Pareto-Lorenz coefficient equal to 1.5 and a level of inequality similar to the level of income inequality in the Netherlands today (see Atkinson, Piketty and Saez, 2011, Table 6). A Pareto coefficient $a = 1.2$ corresponds to an inverted Pareto-Lorenz coefficient equal to 6 and a level of inequality greater than global wealth inequality.

wealth shares exist. By our estimate about 40% of the Norwegian households in the top 0.1% hid wealth abroad in the mid-2000s and these tax evaders hid on average about 50% of their assets. Thus the top 0.1% hid 20% of its total wealth, and including this previously unmeasured wealth increases the top 0.1% wealth share from 8% to 10%, as reported in the top panel of Figure 8. This correction erases half of the decline in the top 0.1% wealth share observed in tax data since the 1930s, at a time when offshore wealth was much less important and when, according to the tax returns of the time, the top 0.1% wealth share was 12%. For the top 0.01%—the wealthiest 300 or so Norwegian households—taking offshore wealth into account increases reported wealth by more than 30% (bottom panel of Figure 8). These results highlight the need to move beyond tax records to capture income and wealth at the very top, even in countries where tax compliance is generally high. They also suggests that tax data may significantly under-estimate the rise of wealth concentration over the last four decades, as the world was less globalized in the 1970s, it was harder to move assets across borders, and offshore tax havens played a less important role.

Importantly, because most Latin American, and many Asian and European economies own much more wealth offshore than Norway, the results found in Norway are likely to be lower bound for most of the world's countries. Figure 9 plots the country distribution of the wealth managed by all Swiss banks against the share of each country in world GDP. Countries above the 45 degree line own more wealth in Switzerland than their share of world GDP would imply. Scandinavian countries are well below the 45 degree line, while countries such as Italy, France, Spain, and Argentina are way above. Taking tax evasion seriously in these countries is likely to have significant effects on long-run trends in income and wealth concentration estimated from tax returns.

7 The Interplay Between Tax Avoidance and Evasion

In this section, we consider the implications of our results for public finances. Should tax evasion become impossible, would wealthy individuals pay significantly more taxes? The answer largely depends on how substitutable illegal tax evasion and legal tax avoidance are. To address this question, we analyze the behavior of the large sample of Norwegians who voluntarily disclosed previously hidden wealth in the context of the tax amnesty described in section 4.3 above.

7.1 Voluntary Disclosers

As we have seen, Norway operates a voluntary disclosure program under which tax payers can avoid penalties and criminal sanctions relating to past tax evasion. Tax evaders can benefit

from the program under three conditions. First, the information has to be offered voluntarily, not in connection with investigations by the tax authority. Second, sufficient information has to be provided in order for the tax administration to assess the correct amount of taxes owed. Last, information on the origin of the disclosed income or wealth must be provided.

While the voluntary disclosure program was rarely used in the decades following its inception in 1950, the number of voluntary disclosers surged dramatically around 2007. The surge coincides with two events that might have led tax payers with hidden offshore wealth to perceive an increase in the risk of detection. In a scandal widely covered by the media, the mayor of Oslo had to resign from office in 2007 after his son-in-law revealed that the mayor owned undeclared bank accounts in Switzerland. At the same time, Norway started signing information exchange treaties that allowed the Norwegian tax authorities to obtain bank information from a number of important tax havens in evasion cases.

As Appendix Figure S.13 shows, most disclosures happened in two waves, in 2009 and 2013-14. The former wave coincides with the G20 tax haven crackdown where the most important tax havens like Switzerland, Luxembourg and Singapore agreed to exchange bank information upon request (Johannesen and Zucman 2014) while the latter wave coincides with the commitment by the same tax havens to exchange bank information automatically. Table 3 compares the disclosers to the rest of the adult population. The data refer to the year 2007, prior to the first wave of disclosure. Disclosers are older (59 years old on average), more educated (almost 60% have a tertiary education) and more likely to be male (66%) and foreign (23%) than non-disclosers. Most strikingly, they are also 25 times wealthier than the average Norwegian.

7.2 Method to Estimate the Substitution Between Evasion and Avoidance

To investigate how tax evaders change behavior when they disclose offshore assets, we employ an event study design. With the aim of determining whether the end of evasion sparks an increase in avoidance, we estimate the change in various measures of reported wealth, income and tax liabilities around the time of disclosure. Our estimating sample includes all 1303 disclosers whose cases have not been dropped as well as a sample of non-disclosers serving to establish a counterfactual.

The specification includes individual fixed effects that absorb level differences in the outcomes across tax payers and ensure that the variables of interest are identified only from time variation. Moreover, we divide the sample of disclosers into ten equally sized groups based on their taxable

wealth in 2007 (before the first wave of disclosures), assign non-disclosers to these same wealth groups, and introduce a separate set of time dummies for each group. This allows time trends to vary across tax payers with different wealth levels and ensures that we identify from a comparison of evaders and non-evaders that are similar with respect to their 2007 wealth. We similarly allow time trends to vary across tax payers of different ages (6 age groups), with different levels of 2007 income (10 income groups) and different levels of 2007 equity wealth (10 equity wealth groups). Equity is the most important form of wealth at the very top and the one whose value can vary the most from year to year with stock market booms and busts. Finally, the specification includes a full set of event time dummies that indicate the year relative to disclosure. Since evaders disclosing offshore wealth in the beginning of year t can incorporate the disclosed wealth into the tax return for year $t-1$, we let year $t-2$ be the omitted event time category. The event time dummies, the main variables of interest, measure changes in the outcomes around disclosure over and above changes in the outcomes of similar non-disclosers at the same point in time.

7.3 Results

The top panel of Figure 10 illustrates the regression results where the outcome is wealth as claimed by the tax payer on the first tax return before any corrections made by the tax administration ("claimed taxable wealth"). While the trend is almost the same for disclosers and non-disclosers in the years -5 to -2, there is a sharp divergence between the years -2 and 0 where the taxable wealth claimed by disclosers increases by around 50% relative to that claimed by non-disclosers. This pattern establishes that disclosures are typically associated with a very significant increase in taxable wealth. Interestingly, the top panel of Figure 10 also shows clear signs of convergence between the years 0 and 4: a considerable part of the increase in claimed taxable wealth observed between years -2 and 0 vanishes between years 0 and 4. This result emerges somewhat sharper in the bottom panel of Figure 10 where the outcome is wealth on the final tax return after any corrections made by the tax administration and thus after integrating the wealth disclosed in year 0 in the tax base. The estimates imply that final taxable wealth drops by around 25% in the 4 years following disclosure. These findings are consistent with the notion that avoidance and evasion are substitutes: when evaders stop evading they also change their behavior in other dimensions so as to offset part of the increase in tax liabilities.

To measure the effect of these behavioral changes on government revenue, we estimate the event study model using as outcome the total tax liabilities (including all income taxes and

wealth taxes) on the final tax return after any corrections made by the tax administration—thus after integrating the wealth and income disclosed in year 0 in the tax base. The top panel of Figure 11 shows that the final tax liabilities of disclosers decrease by 5-10% relative to non-disclosers from the pre-disclosure to the post-disclosure years.

What kind of behavioral changes could drive such a sizable decrease in taxable wealth and tax liabilities? A well-known tax planning technique in the Norwegian context is to shift investments from listed equities, which enter the tax base at market value, to unlisted equities, which enter the tax base at the tax value of the underlying business assets. Since a sizable fraction of corporate assets—such as real estate and intellectual property—have a tax value that is only a small fraction of their market value, unlisted equities generally have a tax value that is much lower than listed equities with a similar market value. Concretely, a recent study estimates that the tax rebate to the most liquid unlisted equities is around 70% and exceeded 90% for a set of unlisted equities that were eventually listed (Gobel and Hestdal, 2015).

To gauge the importance of this tax avoidance technique, we estimate the event study model using as outcome the share of taxable wealth held in unlisted equities. The bottom panel of Figure 11 shows that the unlisted equity share of disclosers increases by around 10 percentage points relative to non-disclosers from the pre-disclosure to the post-disclosure years. This increase can potentially explain a significant part of the decrease in taxable wealth as illustrated by the following simple example. Assume that a tax payer with taxable wealth of 100 sells listed equities for 20 and acquires unlisted equities for the same amount. If the tax value of unlisted equities is 50% of the market value, then taxable wealth falls to 90 and the unlisted equity share increases by roughly 10 percentage points. Thus, the observed increase in the unlisted equity share can account for roughly half of decrease in taxable wealth estimated in the bottom panel of Figure 11.

Another potential explanation for the large decrease in taxable wealth and tax payments is inter vivos transfers from disclosers to their children. Such transfers could potentially serve as a tax avoidance technique by shifting wealth from disclosers to family members with lower marginal tax rates. While we find evidence that the taxable wealth of disclosers' children increases around disclosure, we also find that the estimated decrease in disclosers' own taxable wealth and tax liabilities extends to a sample of individuals with no children suggesting that intra-family transfers play a limited role in explaining the results (see Online Appendix).

Overall, the results suggest that tax avoidance rises sharply after voluntarily reducing tax evasion. This finding suggests that collecting extra tax revenue from the wealthy would require

curbing the supply of evasion services (e.g., by making it more costly and risky for banks to serve tax evaders) and simultaneously closing loopholes to prevent substitution from evasion to avoidance.

8 Conclusion

In this paper, we combine micro-data leaked from financial institutions in offshore tax havens with randomized audit data and population-wide registry data to study how tax evasion varies with wealth. Random audit data, the key source used so far to study tax evasion, largely fails to capture tax evasion at the top of the distribution. While the probability to deliberately evade taxes found in random audits is generally low and does not rise with wealth past the 95th percentile, the probability to hide wealth offshore rises very steeply within the top one percent. Combining random audits and leaks, we estimate that the fraction of the population that evades taxes rises from 1%-2% in the bottom 95% of the wealth distribution to about 40% in the top 0.05%.

Because the income and wealth that evades taxes is extremely concentrated, tax evasion turns out to have important implications for the measurement of inequality. In the case of Norway, accounting for tax evasion leads to an increase of 30% of reported wealth at the top of the distribution. Our results highlight the need to move beyond tax records to capture income and wealth at the very top, even in countries where tax compliance is generally high. They also suggest that tax data may significantly under-estimate the rise of wealth concentration over the last four decades, as the world was less globalized in the 1970s, it was harder to move assets across borders, and offshore tax havens played a less important role. Because most Latin American, and many Asian and European economies own much more wealth offshore than Norway, the results found in Norway are likely to be lower bound for most of the world's countries. Fortunately, many countries have access to leaked data similar to those we exploit in this paper. Our analysis could thus be implemented by tax authorities and researchers around the world, including in countries where offshore tax evasion is much more prevalent than in Scandinavia.

References

Alvaredo, Facundo, Anthony Atkinson, Thomas Piketty, Emmanuel Saez, and Gabriel Zucman. 2011–2016. *The World Wealth and Income Database.*

Atkinson, Anthony, Thomas Piketty, and Emmanuel Saez. 2011. "Top Incomes in the Long-Run of History", *Journal of Economic Literature*, 49(1), 3-71.

Johannesen, Niels, and Gabriel Zucman 2014. *The End of Bank Secrecy? An Evaluation of the G20 Tax Haven Crackdown*, American Economic Journal: Economic Policy.

Kuznets, Simon. 1953. *Shares of Upper Income Groups in Income and Savings* (New York: National Bureau of Economic Research).

Larudee, Mehrene 2015. "Did capital go away? Capital flight as an explanation for declining reported wealth inequality during and after World War I", working paper

Piketty, Thomas. 2014. *Capital in the 21st Century.* Cambridge: Harvard University Press.

Piketty, Thomas, and Emmanuel Saez. 2003. "Income Inequality in the United States, 1913-1998," *Quarterly Journal of Economics* 118(1), 1-39.

Piketty, Thomas, and Gabriel Zucman. 2014. "Capital is Back: Wealth-Income Ratios in Rich Countries, 1700-2010," *Quarterly Journal of Economics* 129(3), 1255-1310.

Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman. 2016. "Distributional National Accounts: Methods and Estimates for the United States, 1913-2013," working paper

Roine, Jesper, and Daniel Waldenstrom. 2015. "Long-run Trends in the Distribution of Income and Wealth." In Handbook of Income Distribution. Vol. 2, , ed. Anthony B. Atkinson and Francois Bourguignon, Chapter 7, 469?592. Elsevier/North Holland.

Saez, Emmanuel and Gabriel Zucman. 2016."Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data," *Quarterly Journal of Economics.*

Zucman, Gabriel. 2013. "The Missing Wealth of Nations: Are Europe and the US net Debtors or net Creditors?" *Quarterly Journal of Economics*, 128(3), 1321–1364.

Zucman, Gabriel. 2014. "Tax Evasion on Offshore Profits and Wealth ", *Journal of Economic Perspectives*, 28(4), 121—148.

Zucman, Gabriel. 2016. *The Hidden Wealth of Nations: The Scourge of Tax Havens*, University of Chicago Press.

	Scandinavia	Sweden	Norway	Denmark
Macroeonomic aggregates				
Adult population (thousands)	**14,711**	7,179	3,434	4,097
National income per adult (US$)	**60,977**	49,949	87,119	58,387
Household wealth per adult (US$)	**201,658**	184,225	189,456	242,431
Household wealth / national income	**331%**	369%	217%	415%
Wealth shares (excluding offshore)				
Bottom 50%	**2.9%**	4.7%	-1.9%	4.1%
Middle 40%	**43.5%**	42.0%	52.7%	42.5%
Top 10%	**53.6%**	53.3%	49.2%	53.4%
Top 1%	**21.2%**	21.0%	17.7%	22.0%
Top 0.1%	**10.2%**	10.1%	8.8%	10.0%
Top 0.01%	**4.9%**	5.3%	4.5%	4.3%

Table 1: Income, wealth and inequality in Scandinavia (2006)

Notes: The top panel reports national income and household wealth in Scandinavia (Sweden, Norway, and Denmark) in 2006 and the bottom panel shows the distribution of household wealth excluding unreported offshore wealth. The wealth distribution is estimated using similar methods in the three countries; it captures 100% of the macroeconomic amount of household wealth at market value reported in the top panel (all financial assets including pension wealth and unlisted businesses, minus debt). The unit of observation is the tax unit; a tax unit is either a married couple or a single individual adult (aged above 20). Children (who own negligible wealth) are dropped. Source: Authors' computations.

	World	Scandinavia	Sweden	Norway	Denmark
Wealth held at HSBC Switzerland (2007)					
Total wealth	118,400	980	332	324	324
Number of client accounts	30,412	767	373	65	329
Of which: matched to tax data		*449*	*227*	*46*	*176*
Mean wealth per client account ($ million)	3.9	1.3	0.9	5.0	1.0
Offshore wealth in Switzerland (2007)					
Total ($ million)	2,912,000	19,500	11,100	4,100	4,300
HSBC Switzerland as % of total Switzerland	4.1%	5.0%	3.0%	7.9%	7.5%
Global offshore wealth (2007)					
Total ($ million)	6,335,000	54,400	29,400	16,700	8,300
HSBC Switzerland as % of total offshore wealth	1.9%	1.8%	1.1%	1.9%	3.9%

Table 2: The wealth at HSBC Switzerland (2007)

Notes: The top panel reports estimates of the wealth and number of clients at HSBC Private Bank Switzerland in 2007. The column "world" is the official total published by HSBC in 2015. Total wealth for Scandinavian countries is the total wealth on the HSBC list received by Scandinavian tax authorities, minus the estimated wealth of non-residents. The number of clients removes double-counting among members of the same family. The middle panel shows estimates of the offshore wealth managed by Swiss banks in 2007 constructed from official statistics published by the Swiss central bank. The bottom panel shows our estimate of the wealth hidden in all the world's tax havens (Switzerland plus other tax havens). The estimate for the global amount of offshore wealth $6,335 billion) is from Zucman (2013). The estimates for Scandinavian countries are based on our computations using Swiss national bank data and Bank for International Settlements Statistics; see Appendix A. Source: HSBC, Zucman (2013), and Appendix A.

	All Norwegian residents. Data for year 2006		
	Non-disclosers	Disclosers	Ratio
number of individuals	3,963,952	1,428	
DEMOGRAPHICS			
age	46	58	1.26
male	0.50	0.63	1.26
number of children	2.25	2.26	1.00
foreign born or foreign national	0.13	0.21	1.62
married	0.41	0.57	1.37
INCOME AND WEALTH ($)			
net wealth (at market value)	140,292	4,373,948	31.18
income (incl. capital gains)	46,633	174,857	3.75
# BANK TRANSFERS			
direct transfers from havens	0.01	0.13	26.15
direct transfers to havens	0.01	0.05	6.19
direct transfers from foreign	0.09	0.79	9.14
direct transfers to foreign	0.25	0.80	3.24
indirect transfers from foreign	0.06	1.49	26.81
indirect transfers from havens	0.00	0.16	41.37
indirect transfers to foreign	0.16	3.22	20.42
indirect transfers to havens	0.01	0.26	23.65

Table 3: The voluntary discloser sample: summary statistics

Notes: This table reports summary statistics on the sample of Norwegian tax evaders who declared their previously hidden assets in the Norwegian tax amnesty. Source: Authors' computations

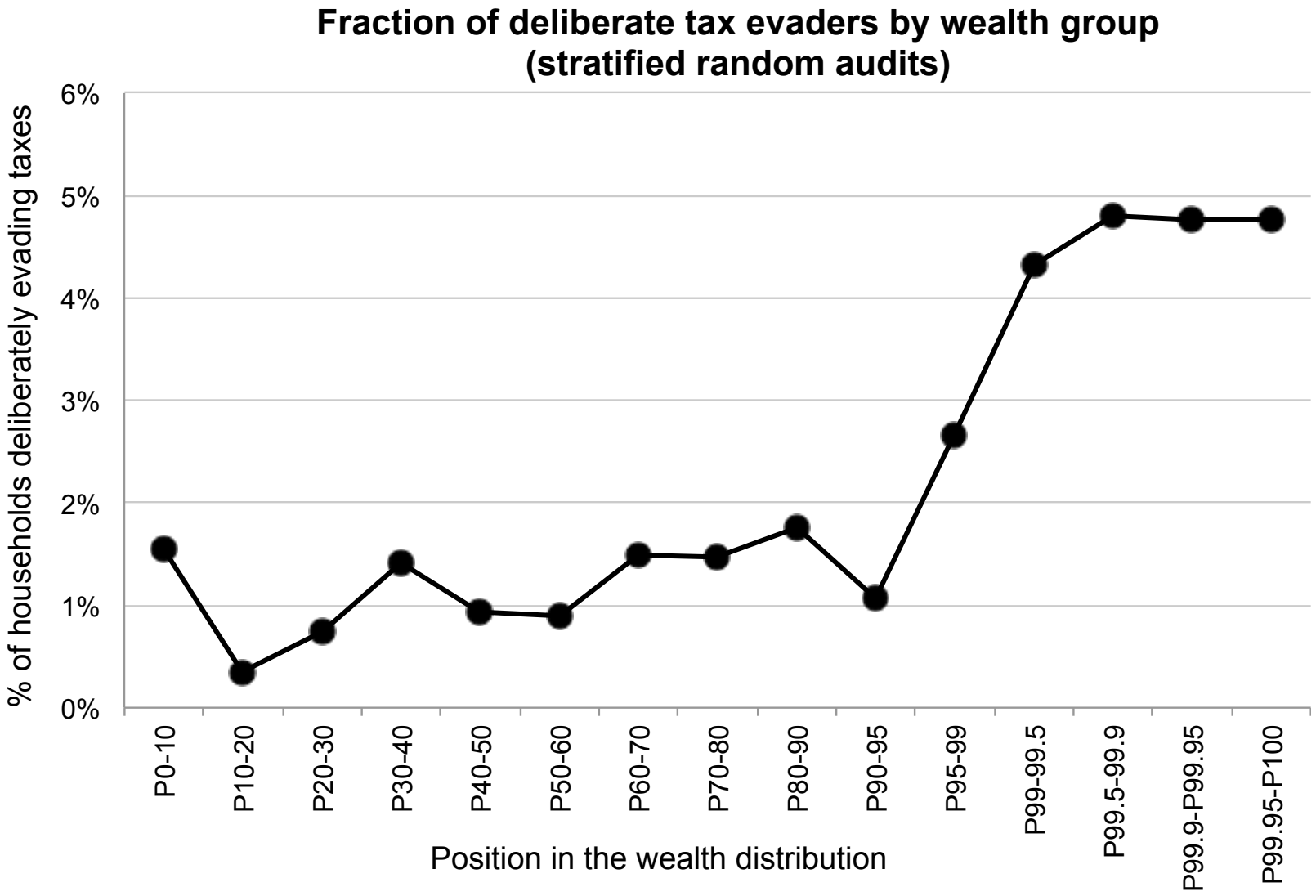


Figure 1: Probability to intentionally evade taxes in random audit data in Denmark by wealth group

Notes: This graph shows the fraction of randomly audited tax units found deliberately evading taxes in Denmark, by wealth bin. The figure pools three waves of random audits, for tax years 2008, 2010, and 2012, and was obtained by matching the sample selected for random audit to population-wide administrative data on wealth. Observations are weighted by audit weights. Deliberate evaders are tax evaders for which the audit has given rise to an adjustment to taxable income and the error is evaluated as being deliberate or based on an improbable interpretation of the law and regulations. Source: Authors' computations.

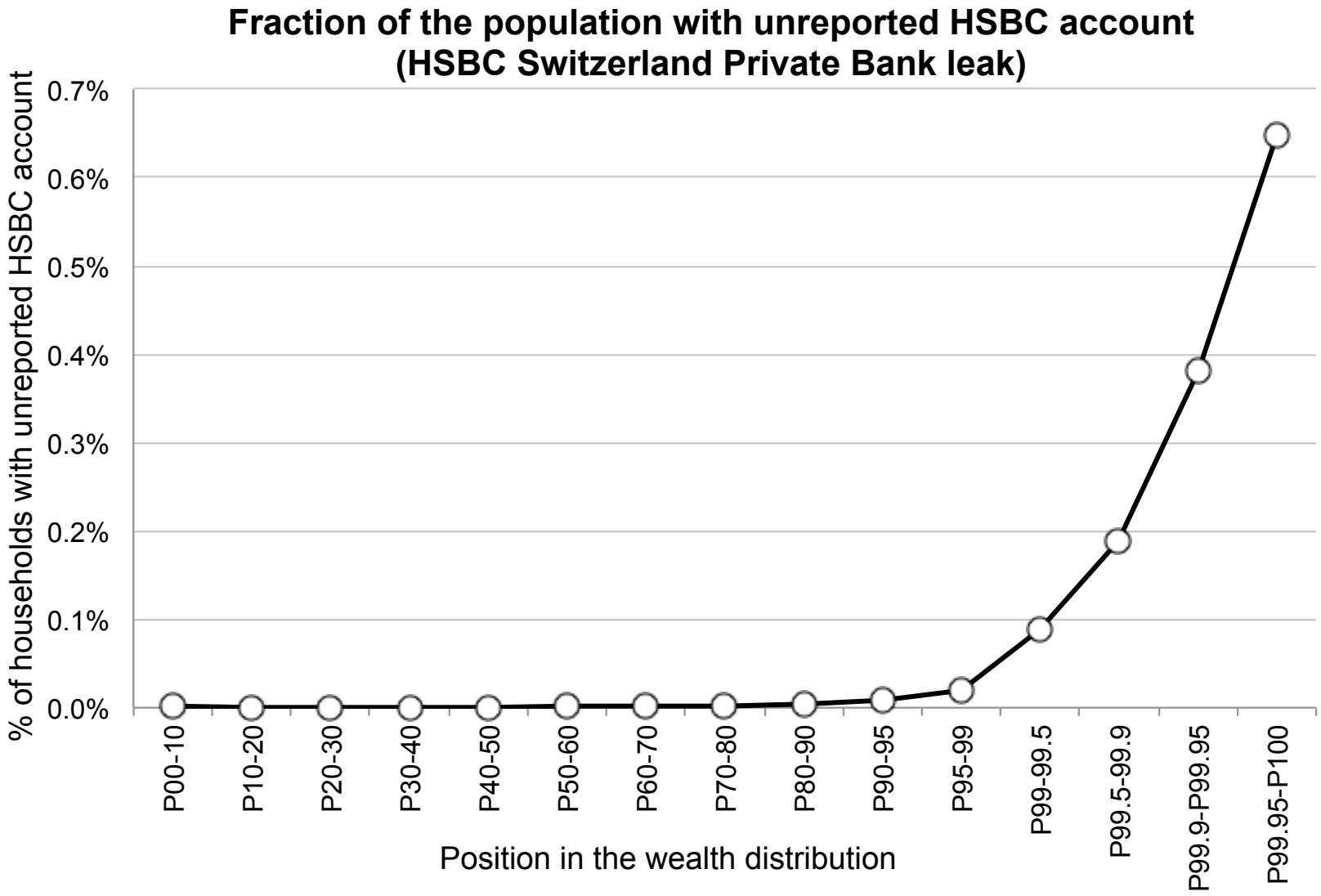
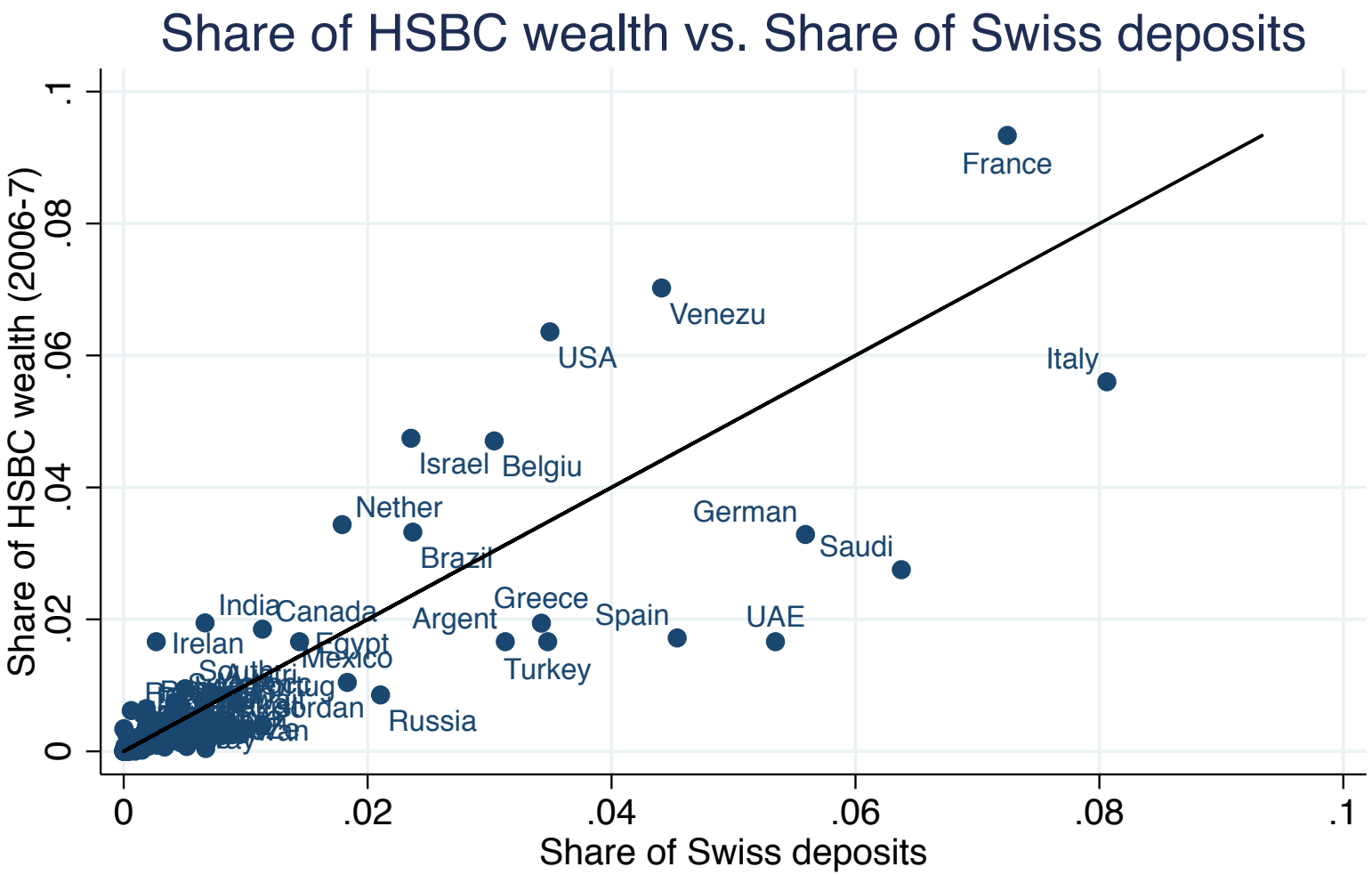


Figure 2: The probability to be found evading taxes in Swiss Leaks

Notes: This figure shows the fraction of households in Scandinavia (Norway plus Sweden plus Denmark) who had an unreported bank account at HSBC Switzerland in 2006, by bins of Scandinavian wealth. In total, 449 households had an account at HSBC Private Bank Switzerland, of which 5% duly reported it and 95% did not. Source: Authors' computations.



Data source: ICIJ and SNB.
Note: In the full sample excluding tax havens (131 countries), a regression of the share of HSBC wealth on the share of Swiss deposits has slope b= 1.00 (se = 0.04) and R-square of 0.80.

Figure 3: Country Distribution of Wealth at HSBC Private Bank vs. Wealth in All Swiss Banks

Notes: This figure shows the country distribution of the wealth managed by HSBC Private Bank Switzerland in 2006-2007 and the country distribution of the foreign wealth managed by all Swiss bank. The black line is the 45 degree line. Source: ICIJ and Swiss National Bank.

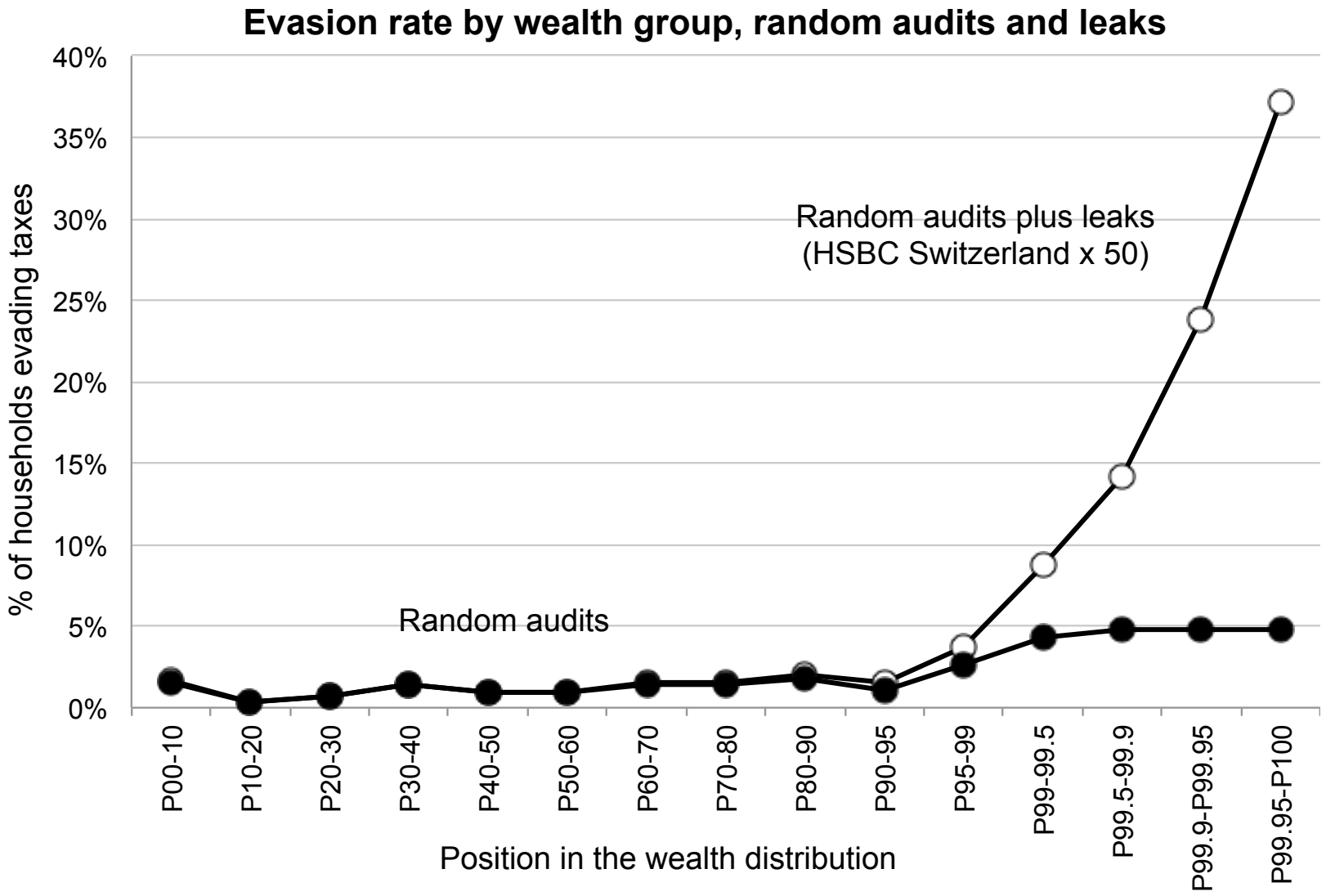


Figure 4: The prevalence of tax evasion across the wealth distribution

Notes: This figure combines random audits with the HSBC Switzerland leaks to estimate the prevalence of tax evasion across the wealth distribution. The number of households with an account at HSBC Switzerland is multiplied by 50 to arrive at an estimate of the prevalence of offshore tax evasion. By design, random audits exclude offshore tax evader. Source: authors' computations.

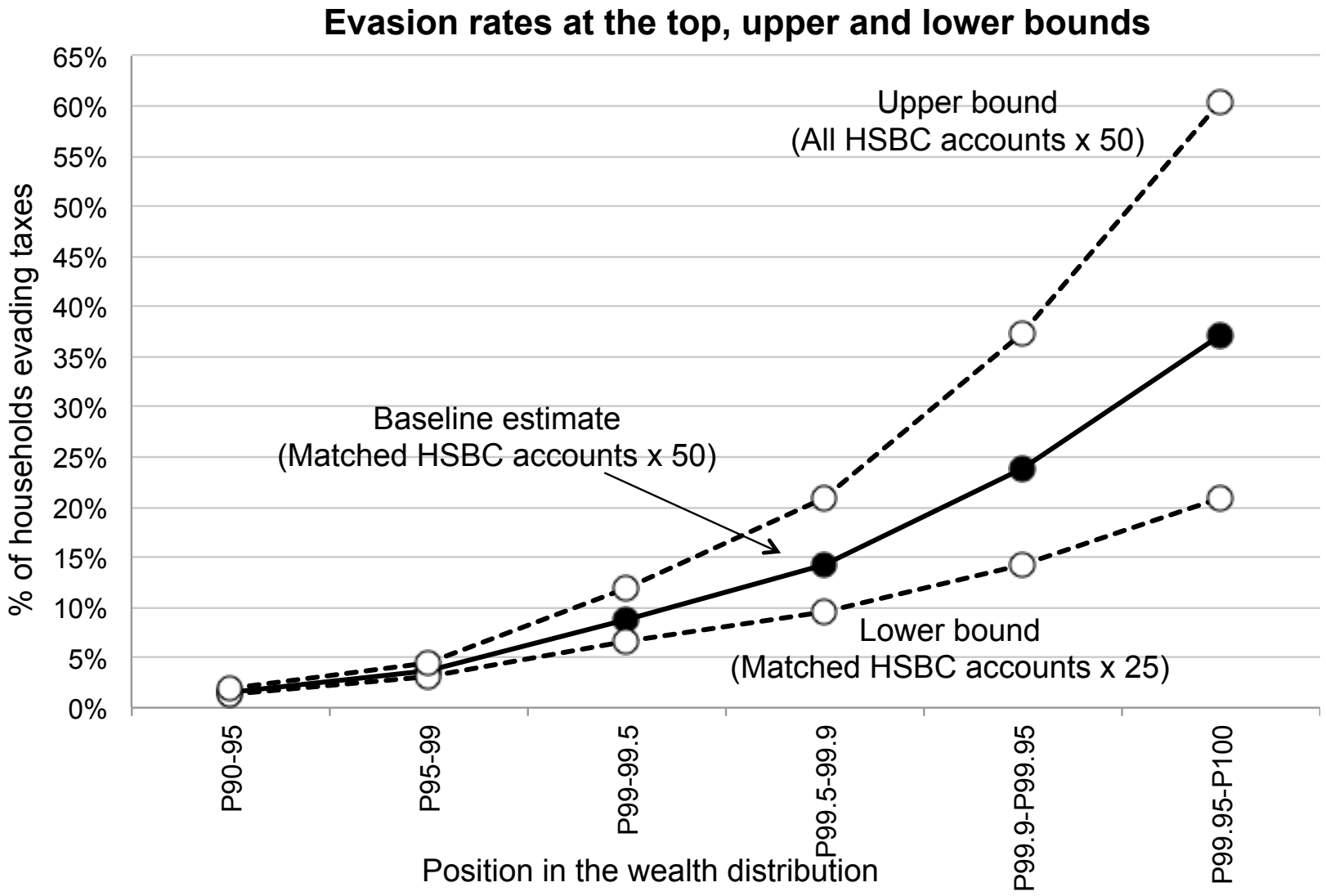


Figure 5: The prevalence of tax evasion across the wealth distribution

Notes: This figure presents lower and upper bounds to the rate of tax evasion in top wealth. The baseline evasion rate is the same as reported in Figure 4. The upper bound assumes that all the HSBC Switzerland accounts connected to Scandinavia and obtained by the tax authorities but that could not be matched to a tax return belong to tax evaders and are similarly distributed as the accounts that were matched. The lower bound assumes that there are no tax evaders with undeclared offshore accounts in tax havens other than Switzerland. Source: authors' computations.

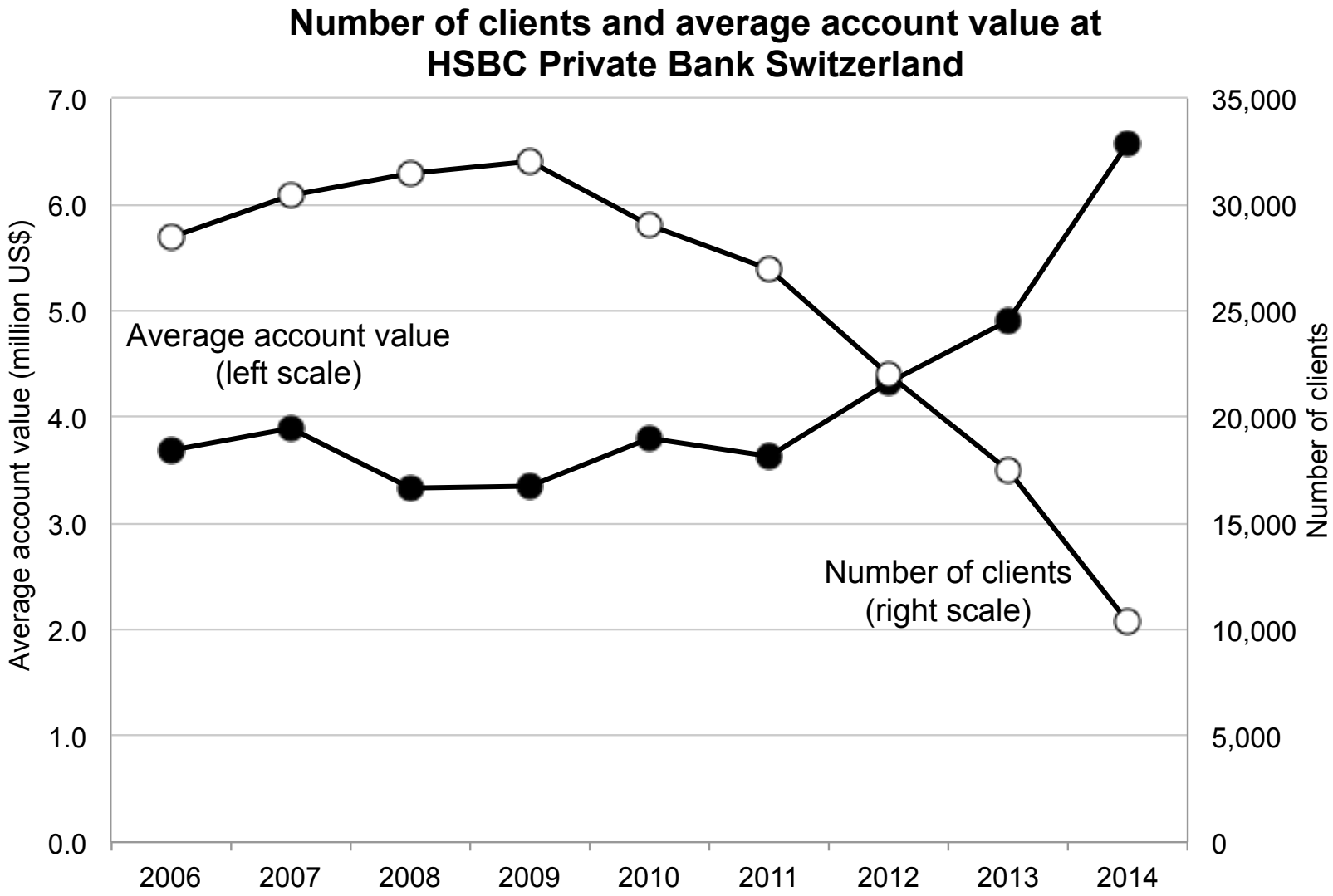


Figure 6: Number of clients and average client account value at HSBC Private Bank Switzerland

Notes: This figure presents the evolution of the total number of clients of HSBC Private Bank Switzerland and the average amount of wealth per client managed by HSBC Private Bank Switzerland. Source: HSBC, `Source:http://www.hsbc.com/~/media/hsbc-com/investorrelationsassets/financial-and-regulatory-reports/gbp-update%20-290115`.

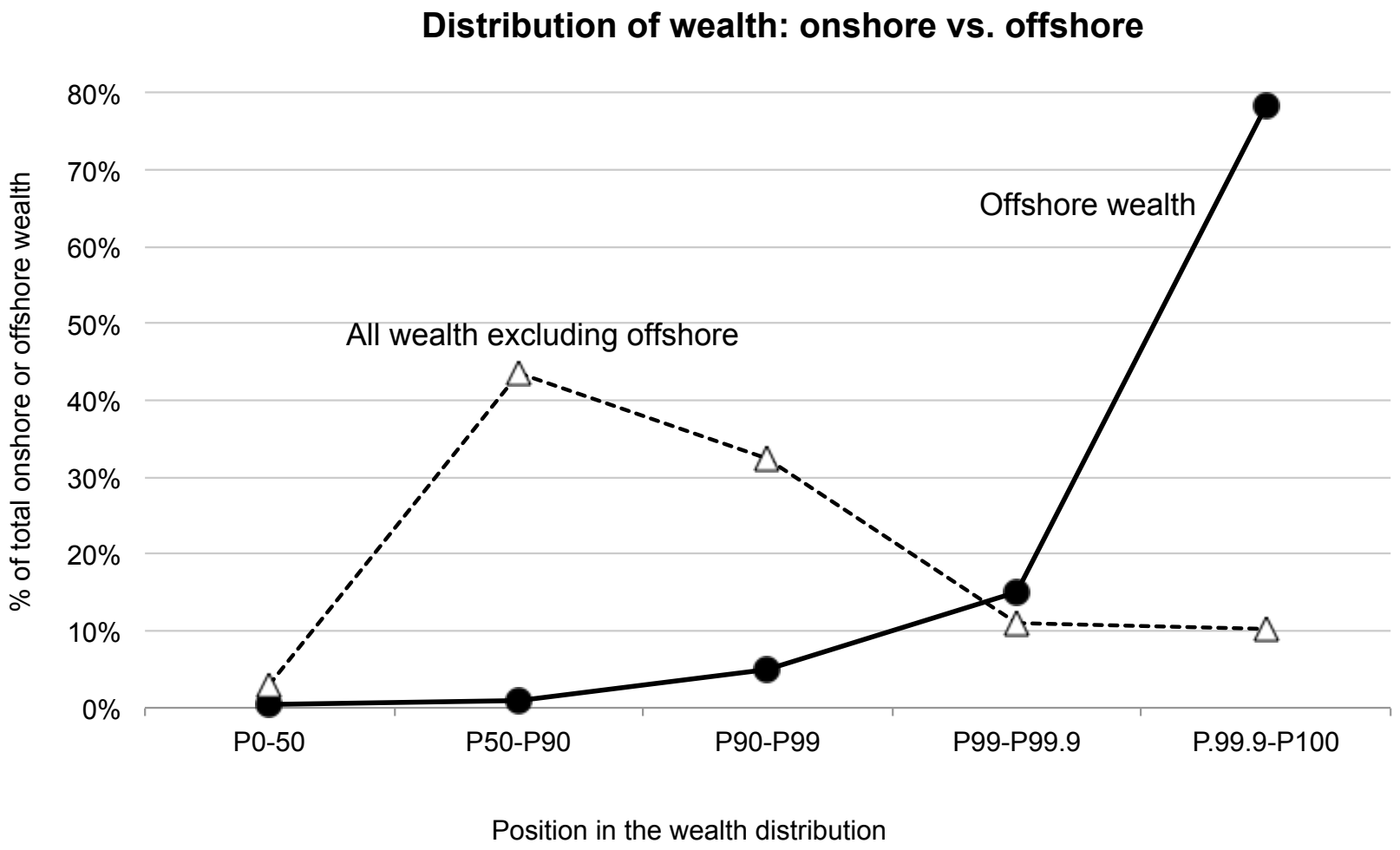
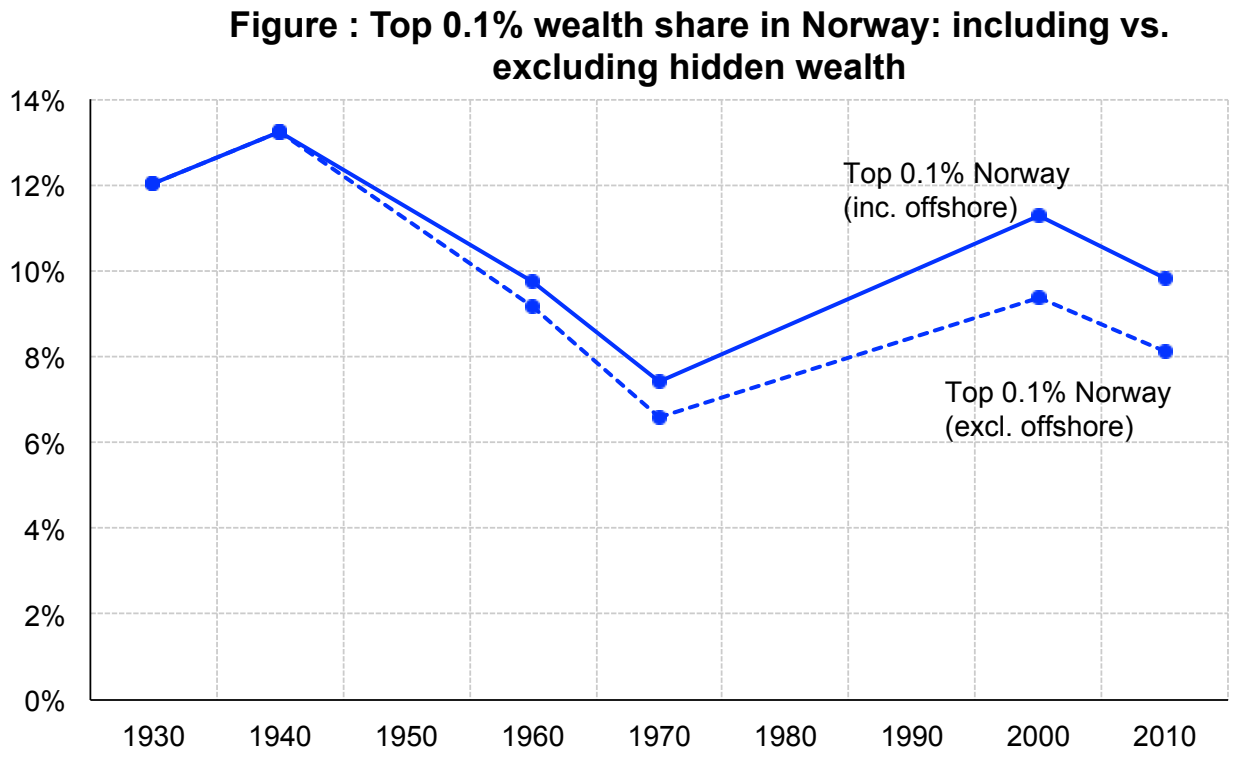


Figure 7: The Distribution of Wealth in Scandinavia: Onshore vs. Offshore

Notes: This figure shows the distribution of wealth in Scandinavia (Norway, Sweden, Denmark) excluding offshore wealth, and our estimate of how offshore wealth is distributed. The distribution of the stock of offshore wealth is based on the distribution of offshore wealth reported by Norwegians in the context of the tax amnesty; wealth held at HSBC Private Bank Switzerland is similarly distributed. Source: authors' computations.



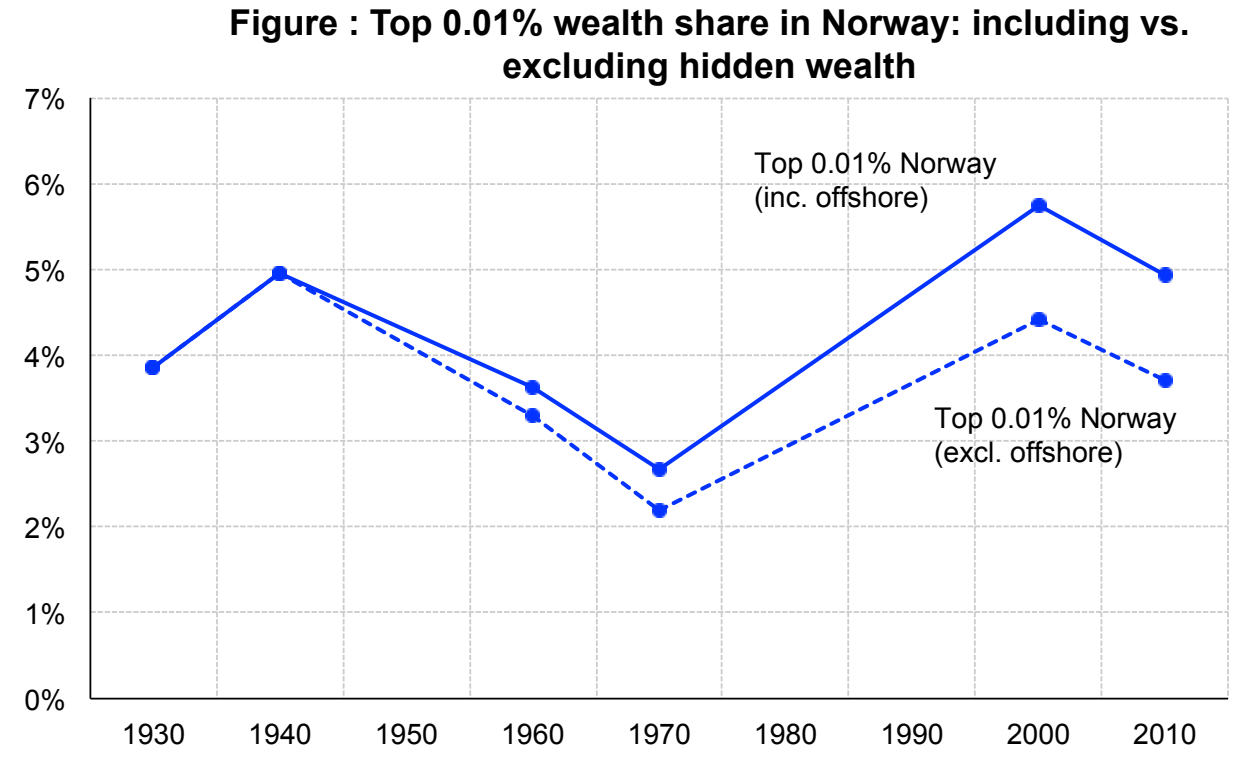


Figure 8: Top 0.1% wealth share and top 0.01% wealth share in Norway including offshore wealth

Notes: This graph compares the top 0.1% wealth share (top panel) and top 0.01% wealth share as estimated from uncorrected tax data vs. corrected by including offshore wealth. Source: Authors' computations, Appendix Table B2 and B4.

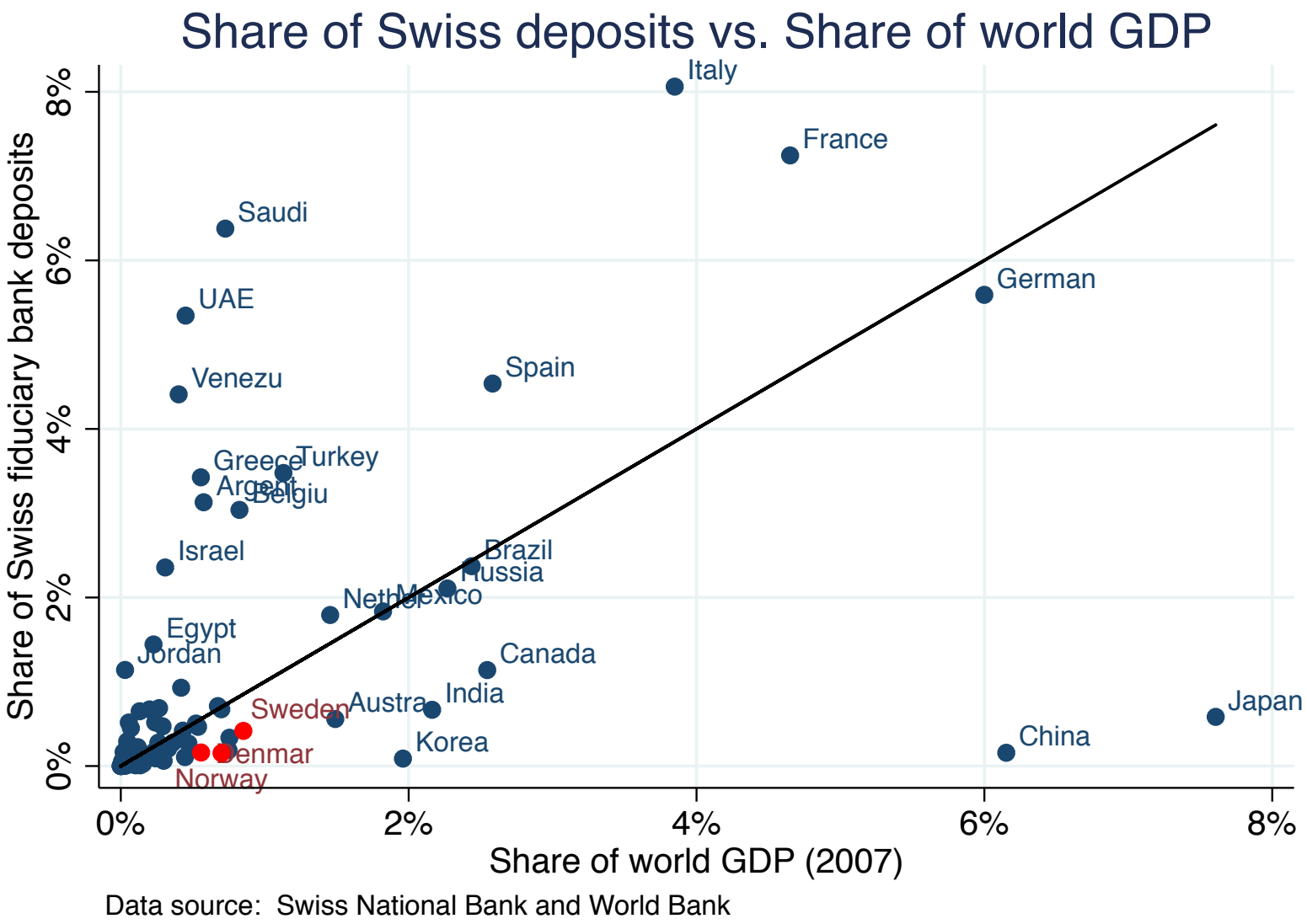


Figure 9: Who Owns Swiss Bank Accounts?

Notes: This figure plots the country distribution of the wealth managed by all Swiss banks against the share of each country in world GDP. Countries above the 45 degree line own more wealth in Switzerland than their share of world GDP would imply, and vice versa for countries below the 45 degree line. Source: Swiss National bank and World Bank.

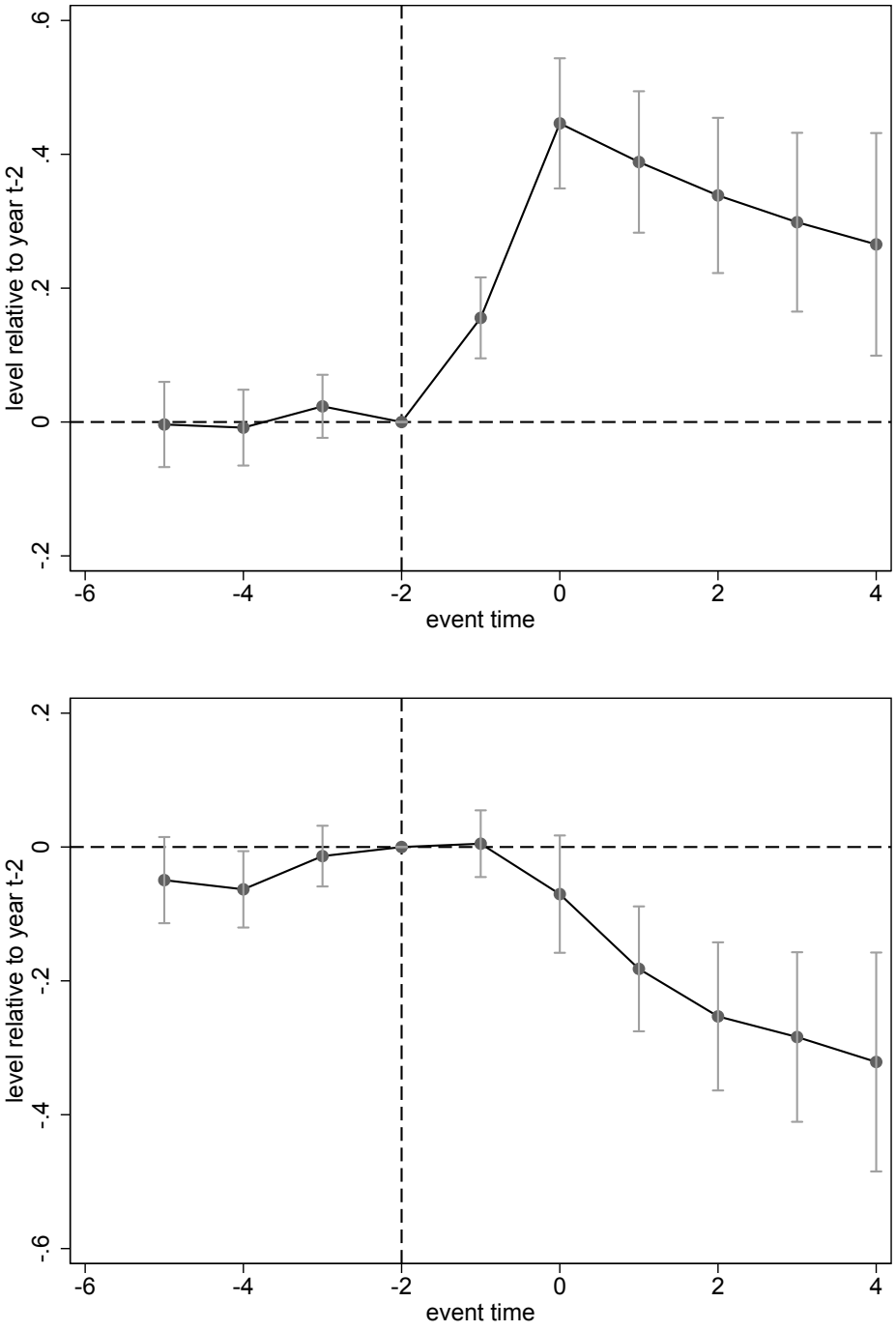


Figure 10: The impact of disclosing previously hidden wealth on reported taxable wealth

Notes: The top panel shows the effect of disclosing previously hidden wealth on wealth reported on the tax return. The outcome variable is wealth reported on the tax return before any correction, and it is regressed on individual fixed effects, year fixed effects, even-time dummies (indicating the year relative to disclosure of hidden wealth), and non-parametric controls for wealth (deciles of disclosers' 2007 wealth and deciles of disclosers' 2007 equity wealth), income (deciles of disclosers' 2007 taxable income), and age (6 age groups) interacted with time. The omitted time category is $t-2$ (two years before disclosure), since evaders disclosing offshore wealth in the beginning of year t can incorporate the disclosed wealth into the tax return for year $t-1$. The bottom panel shows the same specification but using corrected taxable wealth (i.e., after including the wealth previously hidden in the years it was hidden) as the outcome. That is, corrected taxable in $t-3$ for instance is based on the amended tax return for $t-3$ that was filed by disclosing taxpayers after disclosing hidden wealth in t, hence includes wealth that was hidden in $t-3$. Source: Authors' computations.

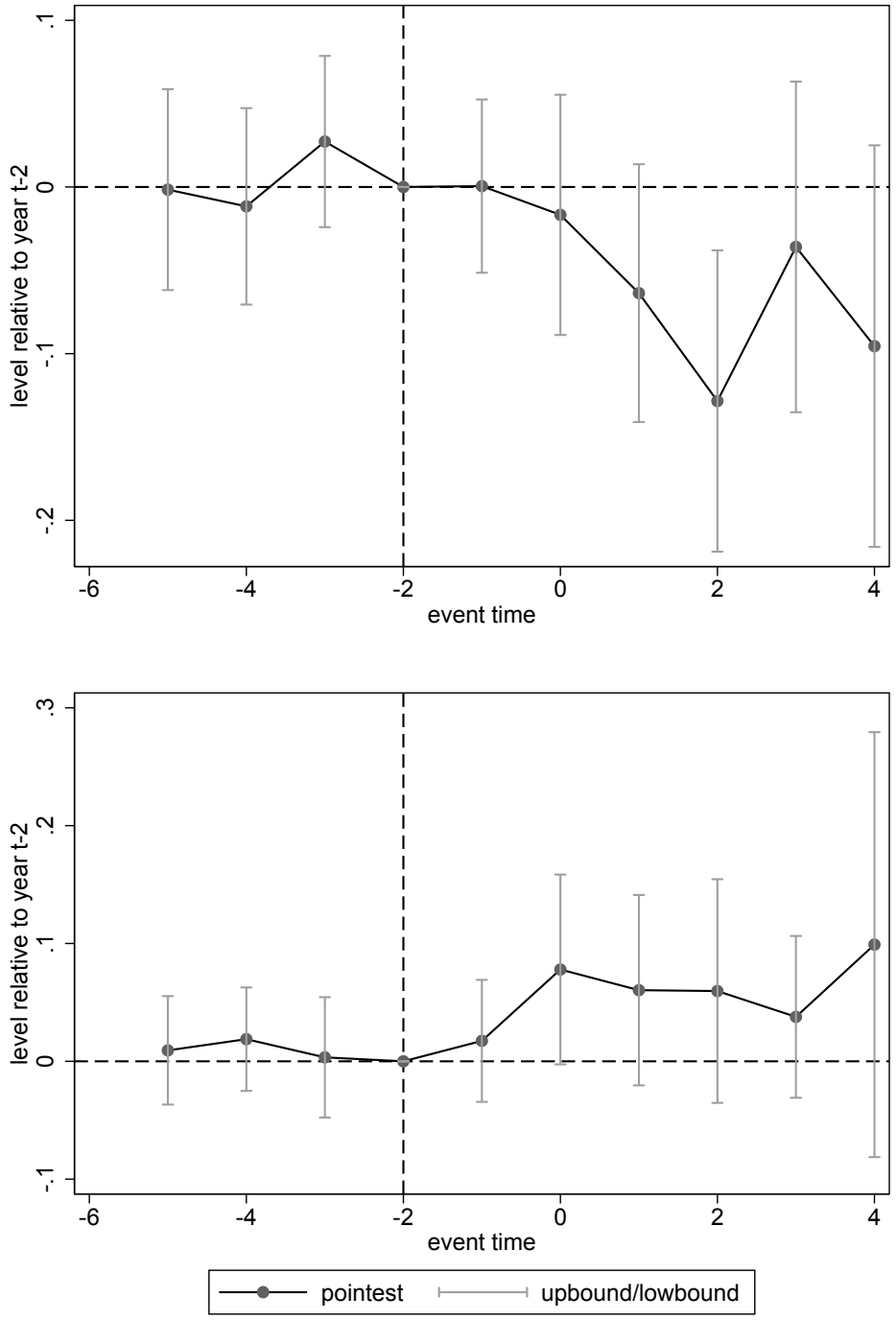


Figure 11: The impact of disclosing previously hidden wealth on taxes paid and portfolio composition

Notes: The top panel shows the effect of disclosing previously hidden wealth on total taxes paid. The outcome variable is total taxes paid, including all income taxes at all levels of government plus wealth taxes, and including taxes paid on wealth tax was hidden before disclosure (i.e., taxes paid in $t-3$ for instance are based on the amended tax return for $t-3$ that was filed by disclosing taxpayers after disclosing hidden wealth in t). The outcome is regressed on individual fixed effects, year fixed effects, even-time dummies (indicating the year relative to disclosure of hidden wealth), and non-parametric controls for 2007 wealth (deciles of disclosers' wealth and deciles of disclosers' equity wealth), income (deciles of disclosers' income), and age (6 age groups) interacted with time. The omitted time category is $t-2$ (two years before disclosure). The bottom panel shows the same specification but using the share of unlisted equities in taxable wealth as the outcome variable. Source: Authors' computations.