Income Taxes and the Residential Mobility of the Rich: Evidence from US and UK Households in Switzerland

Marko Koethenbuerger[1], Costanza Naguib[2], Christian Stettler[1], and Michael Stimmelmayr[3]

[1]ETH Zurich
[2]University of Bern
[3]University of Bath

December 15, 2024

Abstract We provide quasi-experimental evidence on the income tax-induced residential mobility of foreign high-income households living in Switzerland by exploiting the differential tax treatment of UK and US households. While the two groups are similar in terms of non-tax sorting preferences, US households are effectively insulated from Swiss income taxation due to the US world-wide income tax system. Comparing the residential choices of the two groups within a one-hour commuting zone of Zurich, we robustly find a residential location elasticity with respect to the net-of-tax rate of around eight. This estimate captures the 'pure' income tax effect on residential choice, not being downward biased by non-tax location incentives (that positively correlate with income taxes) and by coordination costs between job and residential choices, for instance.

Keywords: High-income Households, Location Choice, Income Taxes, Sorting, Income Sourcing Rules.

JEL classification: H24, H71, J44, R32

1 Introduction

Governments are increasingly concerned about the effect of taxes on the mobility of high-income earners (Epple & Romer, 1991). A significant share of migration is residential migration, where households choose their residence for a given job choice. In a variety of countries, including Switzerland, Scandinavian countries and US metropolitan areas that are composed of multiple states[1], local variation in income taxes influences residential choices. In these settings, the *residential elasticity* is key for studying issues in local public finance and urban economics, such as the efficiency of local policy choices and agglomeration patterns. In the coming decades, the residential elasticity will likely gain in importance at a larger geographical scale. Globalization allows for a rising geographical disconnect of job and residential choices for high-income households. The increasing reliance on remote working arrangements after COVID19 has intensified this trend (Barrero et al., 2021, 2023; Brueckner et al., 2023). To date, little is known about this behavioral response. Existing empirical studies predominantly rely on settings in which households simultaneously choose their tax residence and workplace, as is common in long-distance migration such as international migration.[2] The elasticity estimates are structurally different to a residential elasticity.

This study provides evidence of tax-induced residential mobility of high-income households in Switzerland. To establish this evidence, we focus on household migration within the local labor market around the City of Zurich where, for a given job choice, households are able to choose their tax residence among a high number of municipalities with different income tax rates. Within this area, we compare two groups of high-income households that are differently exposed to local taxation. Given the

[1] In the US, these include some of the largest metropolitan statistical areas (MSA), with a population of roughly 75 million living in multi-state MSAs (Coomes & Hoyt, 2008; Agrawal & Hoyt, 2018).

[2] See Kleven et al. (2013); Akcigit et al. (2016); Moretti & Wilson (2017); Agrawal & Foremny (2019), for instance.

world-wide household income tax system of the US, US high-income households are effectively insulated from income taxation in Switzerland. In contrast, other countries, including the UK, follow a territorial household income tax system, and UK high-income households living in Switzerland can influence their tax treatment by moving to low-tax municipalities. Thus, local income taxes have differential effects on US and UK high-income households, while other determinants of location choice are well aligned.[3]

Our main finding aligns with the reasoning that UK households at the top of the income distribution systematically sort into low-tax municipalities, while US households with comparable incomes do not. They tend to sort into high-tax municipalities to benefit from amenities offered in these locations. Using different specifications of the multinomial choice model, we consistently obtain residential elasticities of UK households of around eight (i.e., a 1% percent increase in the net-of-tax rate (retention rate) increases the probability of moving to the municipality by 8%). This elasticity estimate captures the 'pure' tax effect on residential choice. Our empirical design controls for different confounding non-tax factors that potentially lower location elasticity, such as uncontrolled amenities, costs of coordinating residential and job location choices, and potentially unobservable wage differentials, as more precisely detailed below. The residential elasticity is substantially larger than the location elasticity of nearly 2 or less reported in existing studies on long-distance migration, which involves joint decisions of where to live and where to work.[4] Res-

[3]The analysis makes use of the fact that the two groups of high-income households moved to Switzerland. Possibly strong and idiosyncratic preferences for US and UK locations are 'sunk'. Conditional on being in Switzerland, the two groups share many cultural aspects in addition to having a common language that are relevant in the context of location choice in Switzerland (e.g. access to anglophone networks, emphasis on private (international) schools and on private transportation). We provide different empirical results that are in line with this reasoning. We observe that, for household income levels at which US households are effectively included in the Swiss income tax system, they sort similarly as their UK counterparts. In the same vein, the two groups respond similarly when adding non-tax location factor in the regressions, suggesting that they sort similarly along non-tax dimensions of location characteristics.

[4]For instance, Kleven et al. (2013); Akcigit et al. (2016); Moretti & Wilson (2017) study long-distance migration of top inventors or football players and Agrawal & Foremny (2019) analyzes long-distance migration in Spain.

idential choices are more tax sensitive and the associated welfare effects are more pronounced than existing estimates of (joint) location elasticity possibly suggest.[5]

The estimate is informative for the effect of income taxes on residential choices in local public finance, as local labor markets typically include a high number of municipalities among which households choose residence. The elasticity estimate will be of increasing relevance beyond conventionally-defined local labor markets due to the rising geographical disconnect of job and residential choices. Residential location elasticities and job location elasticities will diverge in a work-from-home economy and, importantly, will become increasingly dependent on the tax regime (Brueckner et al., 2023). For instance, high-income households have more opportunities to work project-based and have jobs in different locations. The residential choice estimate for UK households in our analysis will be informative for their residential choice.[6] Further, sourcing rules among US states imply that income is taxed where it is earned (Agrawal & Brueckner, 2022; Agrawal & Tester, 2023). If this is a high-tax state (relative to other potential residential states), the residential choice of households will be similar in nature to the choice of US households in our analysis. Consistent with our estimates, households have an incentive to choose residence in high-tax locations which allows them to benefit from the better supply of amenities without being subject to local income taxes. Our estimate is useful in quantifying urban models that analyzes the effects of these tax systems on household sorting, for instance. To the best of our knowledge, we provide the first empirical analysis of residential location in the presence of this type of tax system.[7]

A common challenge in estimating the causal effects of tax policies on household

[5]The residential elasticity estimate is equally informative in settings with heterogeneous moves, some of which involve a residential relocation while others necessitate a joint decision of residential and job choice. Given the significant difference in elasticity estimates for the two types of location choices, the underestimation of the aggregate elasticity and welfare effects might well be substantial in such environment.

[6]As discussed in Section 7, our estimate of the residential elasticity is also of relevance for migration over longer distances in a work-from-home economy.

[7]Different to our analysis, Agrawal & Tester (2023) analyze the effect of sourcing rules in the US on job location (contract location of professional golf players). Relatedly, Agrawal & Hoyt (2018) look at the effect of sourcing rules on commuting behavior.

mobility is finding sufficiently large variations in tax rates that are plausibly orthogonal to other factors affecting individuals' location choices, such as labor market conditions, job characteristics, local amenities, public goods, or other taxes. These jurisdiction-specific factors might vary over time, such as with the opening of an international school or improvements to the transport network, or other publicly provided amenities. These variations may align with variations in income tax[8], and their importance for location choices might vary with household income (Diamond, 2016; Fajgelbaum & Gaubert, 2020; Brülhart, Danton, et al., 2022). Accounting for these confounders can be difficult due to data availability or measurement problems.[9] Instead of adding explanatory variables, one might seek to isolate the effect of income tax in location choice models by relying on an appropriate choice of treatment and control group, between which the only difference is exposure to the local income tax system. This approach is quite intuitive, but poses an almost insurmountable obstacle in the estimation of residential elasticity, as households with similar incomes are typically exposed to the same tax schedule of the residence jurisdiction.

Our approach addresses these challenges by comparing tax payers from the US and the UK, who are exposed to different local income tax rates in Switzerland, but have similar location preferences and incomes.[10] This approach flexibly controls for observable and unobservable jurisdiction-specific factors that potentially vary with the level of income tax. To establish evidence, we use household-level

[8]In fact, local governments might 'tax' locational advantages and, thus, are able to levy higher taxes without losing competitive balance vis-à-vis other jurisdictions (Baldwin & Krugman, 2004).

[9]Controlling for the effect of amenities in location choice models requires an understanding of the relevant set of amenities and their importance for different income levels. This poses an informational challenge on top of the standard problem of measuring the value of amenities across jurisdictions, such as the quality of schools or transportation systems, for instance. The insight that it is difficult to observe all relevant amenities and to add them as control variables in regression analysis squares well with the practice of quantifying urban models. Using more precisely observable data such as wages, land prices and commuting times, the model structures are used to back out the level of amenities required for the observed data to be an equilibrium of the model (Redding, 2023).

[10]The group of US and UK households are independent in the sense that we can expect treatment (exposure to the Swiss income tax) not to affect the control group by swapping the nationality. Still, US residents might be expected to renounce their US citizenship for tax reasons and to pay lower Swiss income taxes only, thereby leaving the control group. It appears that US citizens living abroad give up on their US citizenship after some years of living outside the U.S. due to increased compliance costs rather than tax liabilities (Organ, 2021).

data and conduct a municipal-level analysis and an household-location choice analysis (multinomial choice model) for different top income levels. We include standard measures of amenities and homophily in our regressions. These do not substitute for the treatment-control approach. For instance, in the estimations of the multinomial choice models, US households still tend to locate in jurisdictions with a high income tax rate after controlling for these factors, suggesting that households value location-specific factors in high-tax jurisdictions beyond those usually measured. Flexibly controlling for these non-tax location incentives in our treatment-control group analysis eliminates a downward bias in the estimate. In fact, including location-specific factors such as homophily, natural amenities, public spending and distance from the central district as controls in our estimation affects the US and UK households similarly but does not change the location elasticity with respect to the net-of-tax rate. However, our elasticity estimate is affected by the treatment-control approach adopted in this study. The residential location elasticity with respect to the net-of-tax rate increases by a factor of roughly two, which suggests that the US control group is useful in absorbing unobservable non-tax determinants of location behavior, which would not have been controlled for otherwise.

Residential location choice depends on a comparison of average tax rates across different potential locations. Although consistent with economic theory, the use of this tax measure might give rise to estimation bias. This bias is less of a concern in our empirical analysis for several reasons. First, average taxes might correlate with other location-specific factors that are potentially unobserved, leading to an omitted variable bias. As discussed, the comparison between US and UK households accounts for these (potentially unobservable) factors in a flexible way. Moreover, gross wages generally change when moving to a new location.[11] Since it is likely that households move to locations with higher gross wages, not controlling for coun-

[11] Most notably, locations differ in productivity, and thus wages, or in industry composition, which in the presence of industry switching costs imply a gross wage change in response to a move (Borusyak et al., 2022).

terfactual gross wages (which are naturally unobserved) leads to an overestimation of the average tax rate in a progressive tax system (Kleven et al., 2013; Agrawal & Foremny, 2019). The high number of municipalities in Switzerland and the huge tax variations across municipalities allow us to focus the analysis on a local labor market (within a one-hour-commuting zone around Zurich). Households are able to move within the local labor market of Zurich, while job choice and the gross wage remain unchanged. Relatedly, the decoupling of job choice and residential choice attenuates concerns that wage contracts adjust in response to (potential) moves and that gross wages at least partially absorb lower income tax payments.[12] Thus, within the local labor market around Zurich, changes in net wages across locations are related to changes in observable income tax payments, which allows us to compute counterfactual net wages across potential residential locations.

Our analysis uses variations in municipal-level income taxes to estimate the effect of taxes on mobility. Swiss municipalities enjoy a large degree of tax autonomy.[13] Income taxes vary considerably across municipalities, even neighboring municipalities. Thus, moving across municipalities generally leads to a significant change in income tax payments (Roller & Schmidheiny, 2016; Basten et al., 2017). Through tax-induced sorting, households at the top of the income distribution utilize stark income tax differences to reduce their tax burden by an average of 23%.[14] Such differences in income tax burdens are central to public debates and are salient at the household level. The huge variation in municipal income taxes likely overcome adjustment frictions, providing an ideal environment for estimating the effects of income taxes on residential location choice.

Finally, we consider super-rich households, who are typically a core element of

[12] In a similar vein, migration across local labor markets tends to equalize net wages and to neutralize the effect of income tax rate differentials via adjustments in gross wages (Feldstein & Wrobel, 1998). Moving within the local labor market of Zurich leaves labor supply unchanged, rendering gross wages unaffected by internal moves.

[13] Such within-country variation is rarely observed in practice. In most federations, income tax policies are centrally chosen and, if not, local tax authority is mostly limited in scale (OECD, 2021).

[14] See Section 4 and Section C of the Appendix for a detailed analysis.

public debates on the progressive taxation of top incomes. For instance, the income level of the top 1% households in our analysis is around CHF 2,400k (roughly USD 2,600k in current values), which is higher than the income levels commonly considered in existing empirical work.[15] For these households, estimates of residential elasticities are largely lacking.[16]

The paper proceeds as follows. Section 2 provides information on the taxation of US and UK households in Switzerland. Section 3 describes the data and provides key descriptive statistics. Section 4 presents stylized facts on the relationship between household sorting behavior and average tax burdens. Section 5 analyzes the effect of income tax on location choice at the municipal level, while Section 6 provides micro-estimates of household location by estimating a multinomial location choice model at the household level. Finally, Section 7 offers a discussion of the results, followed by concluding remarks in Section 8.

Literature discussion A few papers study the effect of income taxes on residential choices in a local labor market. For instance, Schmidheiny (2006) analyzes sorting behavior around the City of Basel. Different to the study, we adopt a treatment-control group design in which the treated group is exposed to the Swiss tax system. The control group is insulated from the Swiss income tax system and flexibly absorbs location specific factors of the residential choice that correlate with income taxes.[17]

[15] The average income level of the top 1% households in our analysis is CHF 2,400k (roughly USD 2,600k in current values), which differs from previous studies. For instance, the household income levels are around EUR 164k (roughly USD 175k in current values) in Kleven et al. (2014), CHF 120k (roughly USD 130k in current values) in Schmidheiny & Slotwinski (2018) and EUR 200k (roughly USD 213k in current values) in Agrawal & Foremny (2019).

[16] Akcigit et al. (2016) look at similar top 1% income levels, but their estimates apply to a particular group of migrants - top inventors - and the inventors' long-distance migration behavior across OECD countries, which involves joint decisions of where to live and where to work. Young et al. (2016) analyze the sorting behavior of very high-income households across US states and find only small effects.

[17] Similar to Schmidheiny (2006), we include determinants of residential choice other than taxes, such as homophily, natural amenities, public spending and distance from the central district, as controls in the estimation. Including the variables in our estimation of a multinomial choice model affects the US and UK households similarly but does not change the location elasticity with respect to the net-of-tax rate.

In our paper, the elasticity estimate increases by a factor of two when adding the control group. This suggests that the US control group is useful in absorbing unobservable non-tax determinants of location behavior, which would not have been controlled for otherwise.[18]

Coomes & Hoyt (2008) make use of discontinuous changes in taxes at borders to study migration behavior in US multi-state metropolitan areas. Analyzing the relative rate of migration to the states within a multi-state metropolitan area, they find no significant response of in-movers when the income tax is residence-based. As discussed by Coomes and Hoyt, the non-significant result may be due to the limited variation in income taxes in the data or due to the selection of states into income tax regimes. Another possible explanation is that non-tax location factors are not absorbed by a control group, as proposed in this paper. In the absence of the control group, our estimate of the moving behavior is much lower, suggesting that not controlling for non-tax location factors biases the estimate towards zero.

Relatedly, Agrawal & Hoyt (2018) study the effect of income taxes on commuting behavior in US multi-state metropolitan areas. The commuting elasticity is a sufficient statistic to measure the spatial welfare effects of tax policy. However, it only qualitatively indicates that commuting times will be altered by the residential relocation of individuals. In their framework, a quantification of the residential elasticity requires a more detailed structural modelling and calibration of the model. This paper provides a direct estimation of the residential elasticity.

[18] Schmidheiny & Slotwinski (2018) and Martinez (2022) also study migration responses in Switzerland. Martinez uses a difference-in-differences model where, following a reduction of the top marginal income tax rates in the Swiss Canton of Obwalden, high-income households (mean income of CHF 849k) in Obwalden are compared to households (mean income of CHF 206k) in all other Swiss cantons. Schmidheiny and Slotwinski analyze sorting behavior of foreigners with a gross income below CHF 120k who are subject to a special tax regime (which is not municipality specific) for the first five years of residence in Switzerland. Our approach contrasts with their approaches. Most notably, we focus on residential choices in a local labor market to estimate a residential elasticity.

2 Taxation of US and UK Households in Switzerland

In this section, we provide information on the income tax system for US and UK households who reside in Switzerland. We start with a short summary of the Swiss tax system, followed by a description of the global income tax system of the US and its implications for US citizens in Switzerland.

Tax decentralization in Switzerland Switzerland has one of the highest degrees of tax decentralization to sub-central levels of governments worldwide (Brülhart et al., 2015; Schmidheiny, 2017). In 2020, 26 cantons (similar to states) and 2,202 municipalities raised around 54% of the country's total tax revenue. While indirect taxes are important at the federal level, the municipalities and cantons mostly raise tax revenue through residence-based income taxation. The combined cantonal and municipal share of personal tax revenue amounts to 81%.[19] The federal government and the 26 cantonal governments have the authority to design their own income tax schemes, including exemption levels, tax deductions, and the degree of tax progression. Municipalities are bound to these cantonal tax schemes. However, they annually decide on a income tax rate (so-called 'tax multiplier') that applies to the cantonal tax schedule. Cantonal and municipal income taxes are salient to tax payers and can easily be compared on various official and private information portals, many of which are available in English.[20]

Income taxation of US households residing in Switzerland Worldwide, the income of US citizens is generally subject to US federal income tax, regardless of their country of residence (Internal Revenue Service, 2020b).[21] In addition, US citizens living abroad are required to pay income taxes in their country of residence.

[19] All reported fiscal statistics are taken from the Swiss Federal Finance Administration (SFFA): Switzerland's financial statistics for 2020.

[20] See for example the information portal of the Swiss Federal Tax Administration: https://swisstaxcalculator.estv.admin.ch.

[21] Moreover, US citizens living abroad may also be subject to an US state tax.

However, the US provides two important instruments to protect its citizens living abroad from double taxation. The use of one or both instruments is optional for tax payers. First, US citizens living abroad may choose to exclude a limited amount of foreign-earned income (up to USD 107,600 in the 2020 tax year)[22] from their taxable income; this is known as 'foreign-earned income exclusion'. Second, they may credit the tax liability owed to their country of residence against the US federal tax liability on the non-excluded income, referred to as 'foreign tax credit'[23]. Above a certain income threshold, that we compute in Appendix B, it is optimal for households to only use the 'foreign tax credit'.[24] Those households, when the US federal tax liability is higher than the foreign one, pay the difference as income tax to the US. Otherwise, the US federal tax liability is zero.[25] Hence, US residents are effectively insulated from the Swiss income tax system when the US tax is higher than that of Switzerland[26].

For US households living in Switzerland with an income between USD 107,600 and USD 194,600, the tax-minimizing choice depends on location and thus the tax burden in Switzerland. US households with incomes of USD 194,600 (and USD 278,100 for married couples) and higher always minimize their tax burden by only opting for the foreign tax credit, but not the 'foreign-earned income exclusion' (see Appendix A for more detailed information, including an illustrating example)[27].

[22]Together, married couples can generally exclude the double amount, that is up to USD 215,200.

[23]The choice of opting for the foreign-earned income exclusion can be revoked in any tax year.

[24]The decision of whether US residents living in Switzerland should opt for foreign-earned income exclusion depends on gross income level, municipality of residence, and household type. US citizens living in Switzerland married to a non-US spouse can choose to treat the non-US citizen as a US citizens for tax purposes. If they chose to treat the spouse for income tax purposes as US resident, then they must file a joint income tax return for the year they make the choice.

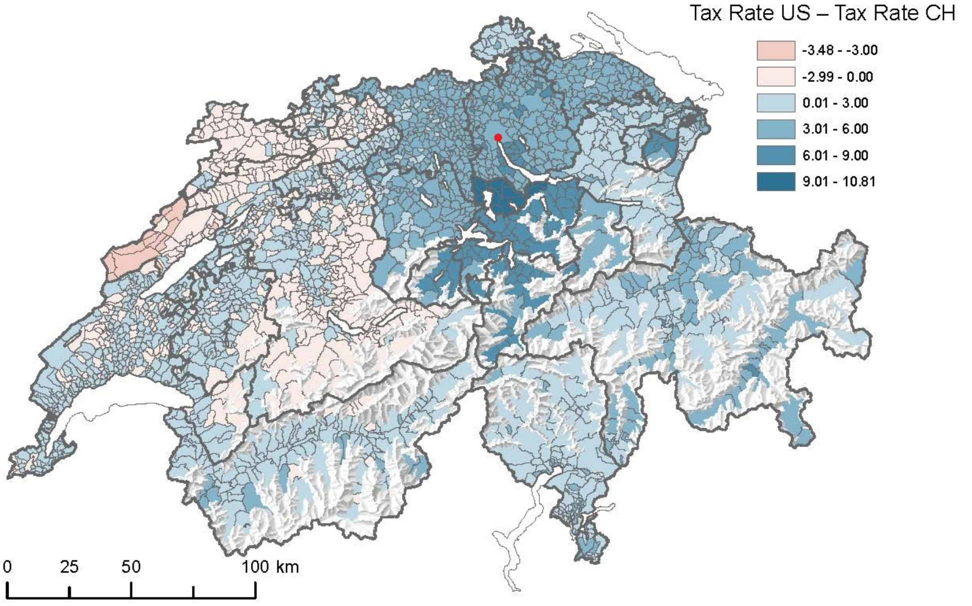
[25]The Internal Revenue Service (IRS) describes it as follows: 'The foreign tax credit is intended to relieve you of a double tax burden when your foreign source income is taxed by both the United States and the foreign country. If the foreign tax rate is higher than the U.S. rate, there will be no U.S. tax on the foreign income. If the foreign tax rate is lower than the U.S. rate, U.S. tax on the foreign income will be limited to the difference between the rates (Internal Revenue Service, 2020a)'.

[26]In principle not all US households are eligible for the foreign earned income exclusion, as there are residency requirements associated with eligibility, such as the *bona fide* residency requirement or the physical presence test. We deem that this is not a serious source of bias in our framework, as we use data from the Swiss Social Security Earnings Records, i.e. from people who pay social security contributions in Switzerland and hence are most likely to effectively reside in Switzerland.

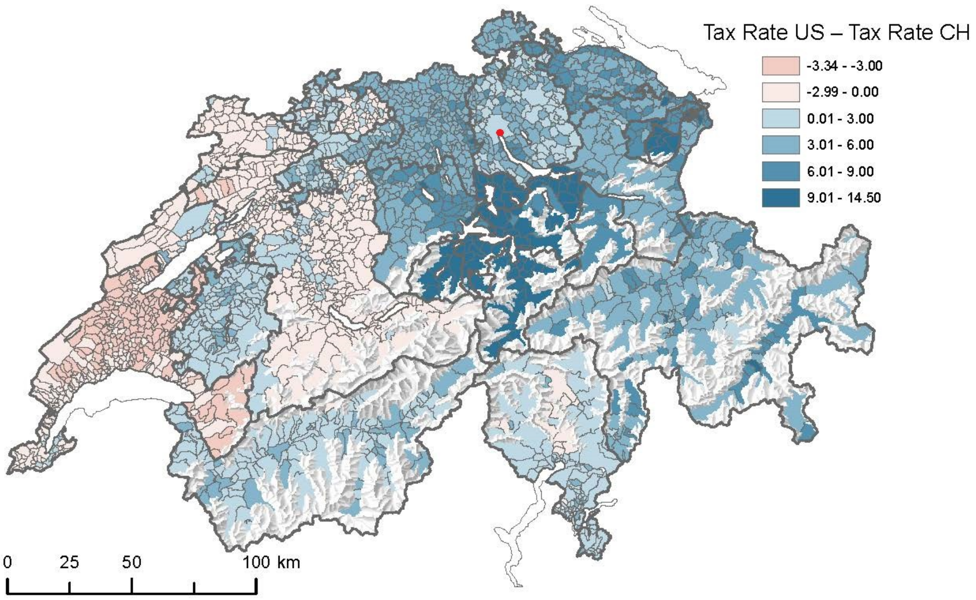
[27]Note that in the period of our analysis (2012-2020), the exchange rate between USD and CHF

Figure 1: Difference between Swiss and US income tax rates (married, two children, 2017)

Gross income: CHF 300,000 p.a.



Gross income: CHF 1 million p.a.



Notes: The figures illustrate the difference between the US federal income tax payment (before application of the tax credit) and the aggregated income tax payment in Switzerland (federal, cantonal and municipal) in 2017 for a married couple with two children and a taxable income of CHF 300,000 (upper graph) and CHF 1,000,000 (lower graph). The federal income tax rate is calculated using the NBER tax simulator. Income tax rates in Switzerland are taken from the Swiss Federal Tax Administration. USD are converted into CHF using the yearly average currency exchange rate for 2017 used by the IRS: 1.024 CHF per USD. The red dot in both panels indicates the City of Zurich.

has been notably stable, and relevantly more stable than in the previous decades.

Figure 1 shows that this is the case for the majority of the 2200 municipalities in Switzerland. This figure compares the US federal income tax with the aggregated income tax in Switzerland (federal, cantonal, and municipal) in 2017 for a married couple with two children and a yearly taxable income of CHF 300,000 (upper graph) and CHF 1 million (lower graph). At these income levels, the US federal income tax rate exceeds the Swiss income tax rate in roughly three out of four municipalities. In the German-speaking region of Switzerland (roughly the middle and north-east areas in Figure 1), high-income households from the US are largely insulated from the Swiss income tax system.[28] Important for our empirical analysis, the US tax is higher in municipalities around the City of Zurich. The location of the city is indicated by the red dot in Figure 1. In this area, US households are fully insulated from Swiss income tax; this also applies to income levels higher than those depicted in Figure 1. In contrast, in the French-speaking part of Switzerland (roughly the western and south-west areas in Figure 1), the income tax for the reference household generally exceeds the US federal income tax.[29]

Income taxation of UK households residing in Switzerland Contrary to the US, the UK has a territorial income tax system and does not tax its non-resident citizens. Consequently, UK citizens living in Switzerland are only exposed to the Swiss tax system.[30] Switzerland generally taxes the income of a UK citizen residing in Switzerland in the same way as it does the income of a Swiss citizen.[31]

[28]The exception are some municipalities in the Canton of Bern and Basel-Land, where US households are exposed to the Swiss income tax schemes to a limited extent. As further explained below, these municipalities are not part of our analysis since they are not within the one-hour-commuting zone around the City of Zurich.

[29]Since the US income tax scheme is generally more progressive at higher income levels, US federal income tax rates also exceed income tax rates in the french-speaking part of Switzerland for annual household incomes above CHF 2 million.

[30]An individual is a tax-resident under Swiss domestic tax law if it meets at least one of the following three conditions: intention to permanently establish his/her usual abode in Switzerland; staying in Switzerland with the intention to exercise gainful activities for a consecutive period (ignoring short absences) of at least 30 days; staying in Switzerland with no intention to exercise gainful activities for a consecutive period (ignoring short absences) of at least 90 days.

[31]The exception are UK citizens (and other foreign citizens) whose yearly gross income is below CHF 120,000 (CHF 240,000 for double earners). They are subject to a special tax regime (so called

Exposure of US and UK households to the Swiss wealth tax One of our key identification assumptions is that differential tax treatment is limited to income taxes. There should be no other location-specific taxes that systematically cause different incentives for location choice of US and UK households. After income tax, the second most important (albeit far less important in terms of tax revenues) direct tax on individuals is a tax on net wealth. Wealth taxes are levied on individuals' worldwide assets in all cantons and municipalities; the top marginal wealth tax rates vary between 0.11% and 1.02% in 2020.[32] The double tax treaty between Switzerland and the US does not cover wealth taxes. Moreover, the US does not unilaterally accept a deduction of the wealth tax owed to Swiss authorities from the US income tax (Internal Revenue Service, 2020a). This exposes US and UK citizens equally to Swiss wealth taxes.

3 Data Construction and Descriptive Statistics

3.1 Merged Census and Social Security Tax Data

Our analysis combines different data sources. We merge the register-based population census data of Switzerland (STATPOP) for the years between 2012-2020 via a social security number with 100% of the Social Security Earnings Records (SSER) from the Old-Age and Survivors' Insurance (OASI).[33] Both, the SSER data and the census data cover all foreign and native populations residing in Switzerland. In the SSER data, employed or self-employed individuals generate one record per job per year. Labor earnings are uncapped and also include variable salary components,

Quellenbesteuerung) until they get a permanent residence permit, for which they can apply after five years of stay in Switzerland. Within this regime, the income tax burden differs across cantons, but does not differ across municipalities within a canton (Schmidheiny & Slotwinski, 2018). The income thresholds are below the income levels (top 10 percent to top 1 percent) we consider in the empirical analysis, c.f. Table 1.

[32] See Brülhart, Gruber, et al. (2022) for a more detailed description of the wealth tax.

[33] See, for instance Martinez et al. (2021) who construct the same data set, albeit for an earlier time period.

such as stock options or bonuses[34]. The register-based census data includes several variables that are important for our analysis, such as information about the individual's age,[35] marital status, gender, municipality of residence, nationality,[36] residence permit, and date of immigration to Switzerland.

An advantage of our data-set is that individuals of the same household are linked through a household identification variable. This variable allows us to construct a data-set at the level of the taxable entity, which we refer to as a 'household'. Based on available household characteristics (wage, number of children, marital status, and income splits), we match these household-level observations with average income tax rates in Swiss jurisdictions (federal, cantonal and municipal) obtained from the tax calculator of the Swiss Federal Tax Administration and with US federal tax rates (for US taxable entities) obtained from the NBER tax simulator.

A notable share of US nationals residing in Switzerland are married to individuals of different citizenship. If the non-US spouse of a US national generates a considerable share of the household income, the location incentives of this household might be significantly influenced by the Swiss income tax system. To mitigate this problem, we restrict our sample of interest to households where all members have US citizenship or, in cases of non-aligned citizenship, i) at least 90% of the income is earned by the US household member, and ii) the non-US spouse earns no more than CHF 100,000.[37]

[34]Note that in our dataset we do not have any information about capital earnings.

[35]We focus on households in which the main earner is between 26 and 64 years old for two reasons: First, the matched SSER-census data only contains labor income, but does not include information on capital income and contributions from pension plans. Second, by focusing on this age range, we exclude young individuals between 18 and 25, who by law constitute a separate taxable unit, but whose location choice is largely driven by parents.

[36]The census data set only contains the nationality that an individual used to register at the Swiss residential registration office. We therefore miss part of the US and UK residents with a multiple citizenship.

[37]For reasons of symmetry, we apply the same rule when selecting the group of UK households.

3.2 Zurich Commuting Zone

For our analysis, we select all municipalities with a commuting time (by car) of up to one hour to the City of Zurich, the largest city in Switzerland.[38] This gives an area consisting of 676 municipalities, as shown in Figure 2.[39] We chose the commuting area around the City of Zurich for various reasons: (i) within this area, income tax rates are below US federal tax rates (c.f. Figure 1), (ii) income tax variation is large, (iii) the area is home to almost half of all anglophone high-income households residing in Switzerland, and (iv) unlike Basel and Geneva, Zurich is not a border city with limited Swiss data coverage of potential location choices, as detailed below.

Income taxes are below US federal taxes First, within this area, combined federal, cantonal, and municipal income tax applied to high-income households are almost exclusively below the US federal income tax, so tax differences across municipalities and cantons are irrelevant to US citizens.[40] As shown in Figure 1, this is not always the case for municipalities outside the one-hour-commuting zone. Thus, restricting the scope of analysis to the area around Zurich allows a much cleaner estimation of the income tax-induced sorting effects.

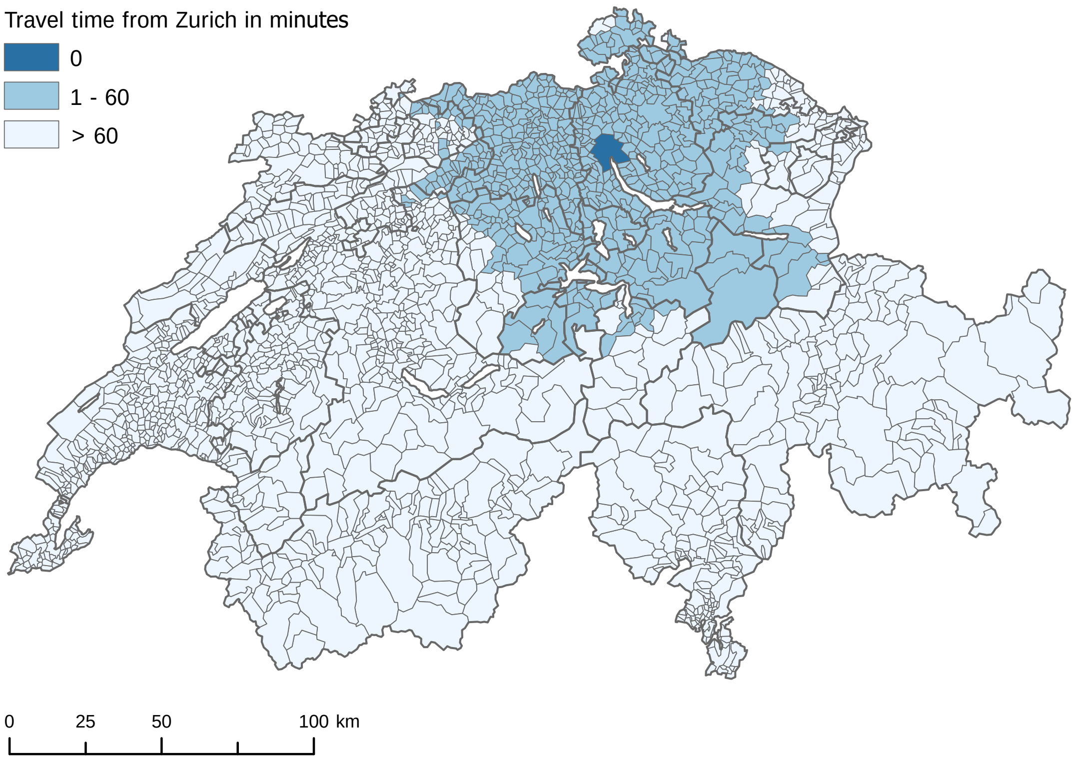
Huge variation in income taxes The one-hour-commuting area around Zurich allows us to exploit huge tax variations across 676 jurisdictions. The area includes the local tax havens of Zug (average commuting time of 31 minutes to Zurich) and Schwyz (39 minutes), which levy the lowest income taxes in Switzerland. Figure 3 provides an overview of the huge variations in income tax rates within the one-

[38] We use the georoute command from Weber & Péclat (2017). Travel time is in minutes and is computed based on the road network under average traffic conditions. We also provide sensitivity checks in which we select all municipalities with a commuting time by car of up to 45 minutes to the City of Zurich. See Section D of the Appendix.

[39] See Section B of the Appendix for detailed descriptive statistics for these municipalities. During our period of observation, the number of municipalities within a commuting time of 1 hour around Zurich dropped from 705 in 2012 to 676 in 2020 due to municipal mergers.

[40] We provide more detailed information, including income thresholds above which US households are insulated from Swiss income tax, in Section 3.3.

Figure 2: Zurich Commuting Zone

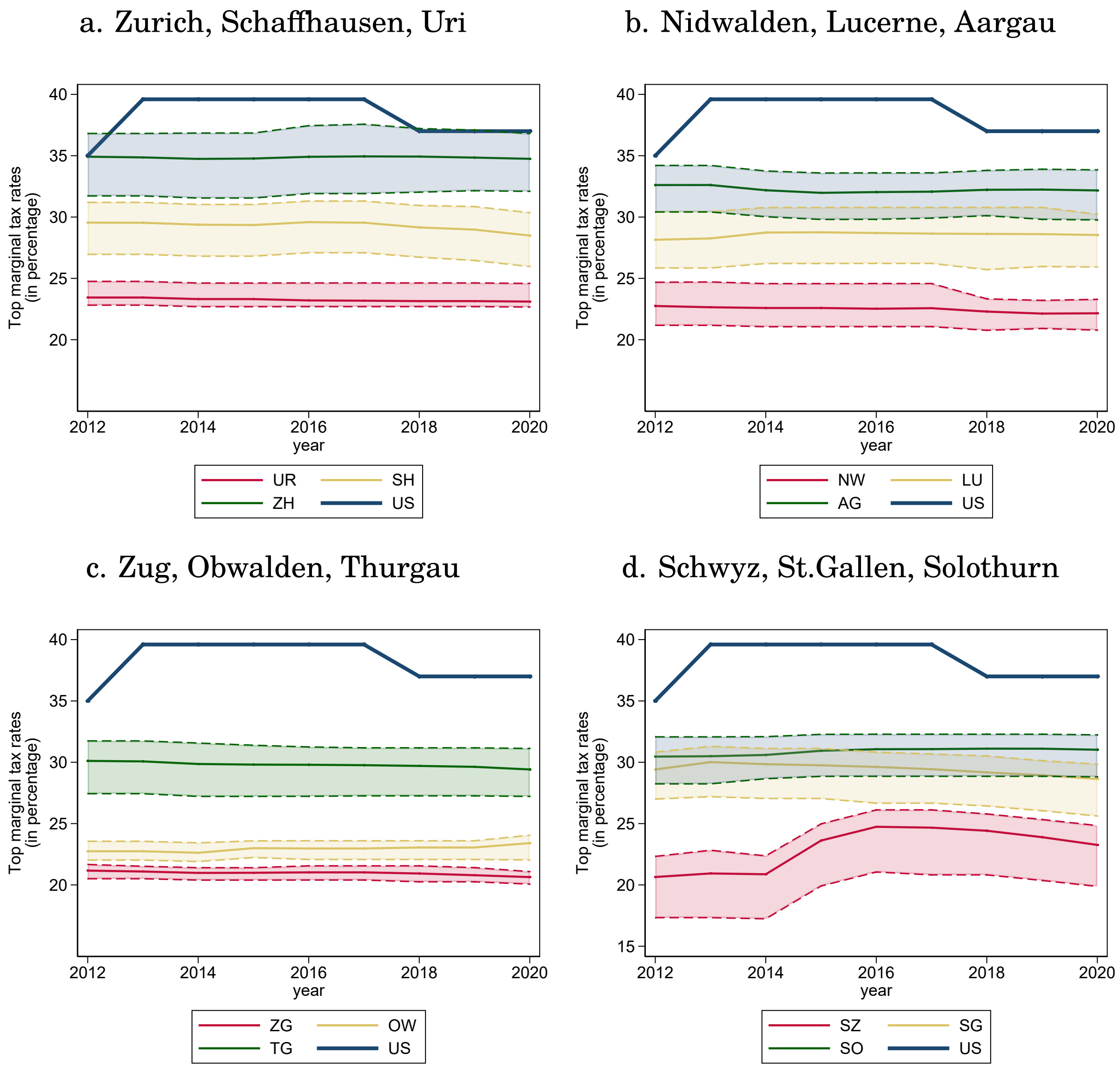


Notes: This figure shows the municipalities that are within a travel time (by car) of one hour to the Zurich main station (main station of the railway network). The commuting time is calculated using the georoute command from Weber & Péclat (2017). Travel time is in minutes and is computed based on the road network under average traffic conditions.

hour-commuting area. The figure depicts the top marginal income tax rate from 2012-2020. The top marginal income tax rates are aggregates, including federal, cantonal, and municipal taxes. The solid lines correspond to the cantonal means, and the shaded areas represent the municipal taxes between the 5^{th} and 95^{th} percentiles. We include the top marginal federal income tax rates of the US. Figure 3 shows that the cross-sectional variation in the top marginal income tax rates is large; the majority stems from the variation across cantons rather than across the different municipalities within a canton. The mean top marginal tax rates of the (relatively) high-tax cantons Zurich and Aargau exceed those of the cantons in central Switzerland, in particular Uri, Schwyz, Obwalden, Nidwalden and Zug, by around 10-15 percentage points. In comparison, cross-municipal variation within cantons is around 2-5 percentage points.

The time variation is limited over the sample period in which individual wage and population census data are available to us.[41] Tax rate changes of one percentage point or more are rare. Given this observation, and as documented in Sections 5 and 6, our analysis primarily exploits the large cross-sectional variations in tax rates.

Figure 3: Top Individual Income Marginal Tax Rates 2012-2020



Notes: Top marginal income tax rates in Switzerland for the area with a travel time (by car) of one hour to Zurich and by canton. For Switzerland, income tax rates are the aggregate of the municipal, cantonal, and federal tax rates. For the United States, state income taxes are not included. The bold lines correspond to the cantonal means. Shaded areas show municipal tax rates between the 5^{th} and 95^{th} percentiles.

[41]For this time window, variation in income tax rates across time is limited to a within variance of the top marginal tax rate of 0.24 percentage points, as compared to a between variance of 14.18 percentage points.

Largest labor market area in Switzerland The one-hour-commuting area around Zurich is the largest labor market area in Switzerland, with major industrial and financial industries. The area hosts many national and international firms and offers a variety of job opportunities for high-skilled workers. This has two advantages for our analysis. First, the area is home to almost half of anglophone high-income households in Switzerland.[42] As shown in more detail in the next subsection, the focus on the Zurich local labor market leaves us with a sufficient number of high-income households.

Second, within this area, households can choose their tax residence while leaving their workplace and gross wage unchanged. The job choice and residential choice can be decoupled.[43] This is important, since household location decisions are driven by after-tax income, which is jointly determined by gross wages and income tax rates. Gross wages might differ across potential locations. Since counterfactual gross wages are naturally unobserved, disentangling the effect of taxes from wage effects poses a key challenge when estimating the causal effects of income tax policies on household location choice.

Area is restricted to Switzerland The commuting time from the City of Zurich to the nearest jurisdiction abroad amounts to around one hour.[44] In contrast to the commuting zones of Geneva – the second most populated city in Switzerland – and Basel, the Zurich commuting zone is restricted to Swiss territory (c.f. Figure

[42] Averaged over the observation period between 2012 and 2020, the one hour commuting area around Zurich accounts for 47 percent of the households in Switzerland with an income above CHF 300,000 for which the top earner is a US or UK national and for 43 percent of the overall anglophone population living in Switzerland. For the year 2020, the area of one hour around Zurich captures 676 out of 2,202 municipalities.

[43] The reasoning is in line with a sequential household location choice where, first, the household chooses the job location and, afterwards, chooses the exact residential location among the 676 municipalities within the local commuting zone around Zurich. The latter is done for given job choice and gross wage. In Appendix F, we show that US and UK households in Switzerland exhibit similar commuting behavior. This in line with the view that the two groups share a similar willingness to choose a residential location different from the job location.

[44] The commuting time to the nearest city abroad (Waldshut-Tiengen, located in the German state of Baden-Württemberg) amounts to 58 minutes.

2). Thus, it is covered by Swiss data sources and can be captured by the empirical analysis.

3.3 Summary Statistics

Table 1 presents descriptive statistics for 2016, the midyear of our annual observation period ranging from 2012-2020.[45] In our analysis, we focus on the top 10% of US and UK households in terms of gross wages. We define the 10% threshold for each year, jointly for all working-age US and UK households living within a commuting distance of one hour around Zurich. The anglophone working-age population in this area has higher incomes compared to the overall population. Depending on the year, the 90^{th} percentile ranges between CHF 331,473 (in 2012) and CHF 375,177 (in 2020) and is roughly twice as large as the 90^{th} percentile of the overall working age population residing in the Zurich commuting zone.[46]

In the one-hour-commuting zone around Zurich, US households are effectively insulated from the Swiss income tax system if their household income exceeds around CHF 207,200 (individual) or CHF 296,200 (couples).[47] The exact threshold value slightly varies across years and household characteristics. We introduce different safeguards to ensure that US households in our sample are effectively excluded from the Swiss income tax system. First, the focus on the top 10% creates an income margin of around 10-20%, safely excluding households for which the Swiss tax liability is higher than the US tax liability. Second, the thresholds refer to US federal tax liability only, not accounting for possible US state-level taxes. Thus, we provide a lower boundary for the total US tax liability. These safety margins allow us to rule out possible bunching around the threshold (i.e., self-selection of US households into an income range slightly below the income threshold), which could bias the selection

[45] See Section B of the Appendix for summary statistics on the pooled years from 2012 to 2020.

[46] Depending on the year, the 90^{th} percentile of household labor income in the Zurich commuting zone ranges between CHF 174,655 (in 2012) and CHF 179,530 (in 2020).

[47] See Appendix A for detailed information.

Table 1: Summary Statistics: US and UK Households (2016)

	All households		US households		UK households		US–UK
	mean	sd	mean	sd	mean	sd	p-value
Panel A: top 1% (N_{total} =107 / N_{US}=35 / N_{UK}=72)							
Wage (in 1000 CHF)	2390	1281	2667	1639	2255	1052	0.119
Gender (female=1)	0.11	0.32	0.14	0.36	0.10	0.30	0.487
Age	51.08	6.97	52.51	6.08	50.39	7.30	0.139
Years in Switzerland	7.32	7.16	5.97	3.91	7.97	8.23	0.176
Number of children	0.81	1.07	0.77	1.06	0.83	1.09	0.781
Marital status (married=1)	0.65	0.48	0.69	0.47	0.64	0.48	0.637
Tax rate	26.12	5.42	29.41	4.28	24.52	5.21	0.000
Panel B: top 1%-5% (N_{total}=418/ N_{US}=111/ N_{UK}=307)							
Wage (in 1000 CHF)	757	190	752	173	759	196	0.762
Gender (female=1)	0.17	0.37	0.18	0.39	0.16	0.37	0.618
Age	48.01	6.82	47.09	7.37	48.35	6.59	0.097
Years in Switzerland	6.72	5.23	6.03	4.87	6.97	5.34	0.104
Number of children	1.04	1.08	1.04	1.02	1.05	1.11	0.937
Marital status (married=1)	0.67	0.47	0.59	0.49	0.69	0.46	0.067
Tax rate	24.34	4.71	26.48	4.33	23.56	4.60	0.000
Panel C: top 5%-10% (N_{total}=463/ N_{US}=126/ N_{UK}=337)							
Wage (in 1000 CHF)	429	45	430	45	429	46	0.922
Gender (female=1)	0.17	0.38	0.21	0.41	0.16	0.37	0.278
Age	46.32	7.21	45.79	7.98	46.51	6.90	0.342
Years in Switzerland	6.54	6.30	5.71	6.22	6.85	6.31	0.085
Number of children	1.00	1.04	0.99	1.04	1.01	1.04	0.877
Marital status (married=1)	0.69	0.46	0.69	0.46	0.69	0.46	0.966
Tax rate	20.78	4.19	22.00	4.04	20.32	4.15	0.000
Panel D: Top 10% (N_{total}=988 / N_{US}=272 / N_{UK}=716)							
Wage (in 1000 CHF)	780	729	849	929	754	636	0.067
Gender (female=1)	0.16	0.37	0.19	0.39	0.16	0.36	0.219
Age	47.55	7.17	47.19	7.79	47.69	6.92	0.328
Years in Switzerland	6.70	5.97	5.88	5.42	7.01	6.15	0.007
Number of children	1.00	1.06	0.98	1.03	1.01	1.07	0.738
Marital status (married=1)	0.68	0.47	0.65	0.48	0.68	0.47	0.314
Tax rate	22.86	4.98	24.78	5.00	22.13	4.78	0.000

Notes: This table presents descriptive statistics of the taxable entities by nationality and income category for 2016, the midyear of our observation period ranging from 2012 to 2020. Wage (in CHF), tax rate (in percentage) and number of children (living in the household) refer to the taxable entity. The other characteristics (gender, age, number of years residing in Switzerland) refer to the top earner of the taxable entity. Note that the total number of households (N_{total}) of the different panels might not be strictly proportional to wage percentile brackets, due to the restriction of the final sample to households in which at least 90% of the income is generated by US household members or by UK household members, respectively, and in which an eventual non-US spouse or an eventual non-UK spouse, respectively, earns no more than CHF 100,000, as described in Subsection 3.1.

of US households in the sample. It also addresses the concern that small (expected) income variations might change the effective exclusion from the Swiss tax system and thus the location incentives of the US households in the sample.

Our final data set includes 9,268 anglophone households belonging to the top 10% of the income distribution; 2,616 from the US and 6,652 from the UK. Table 1 provides descriptive statistics for the top 10% income group and the subgroups (top 1%, the top 1-5%, and the top 5-10%), for which we provide separate estimates later. Gross wages are relatively similar across US and UK households. The average annual gross wages of the top 10% amount to around CHF 780,000 and are large in international comparison. Even more striking, the average labor income of anglophone households belonging to the top 1% amounts to around CHF 2.4 million. The average age of the top 10% principal earners is largely similar across US (47.19) and UK (47.69) nationals. The same holds true for the average number of children living in the same household (0.98 for US households versus 1.01 for UK households) and the share of female principal earners (19% for US households versus 16% for UK households). In contrast, principal earners of high-income households from the UK tend to have immigrated to Switzerland a slightly longer time ago (around six years, on average, for the top 10% of US nationals compared to seven years for their UK counterparts). As expected, average tax rates for US and UK households are robustly different in statistical terms for all income groups. This difference is an early indication of the diverging influence of income taxes on the location behaviors of UK versus US households. In the next sections, we analyze the role of income taxes more thoroughly using different empirical approaches.

4 Stylized Facts

We begin our analysis by showing reduced-form graphical evidence on the relation between local income taxes and the location choices of US and UK households. Figure 4 compares the shares of anglophone households that live in the tax havens of the Zurich commuting zone for different income levels. We define 'tax havens' as the 5% of the 686 municipalities with the lowest top marginal income tax rates (aggregated over the federal, cantonal, and municipal levels).[48] As apparent from the figure, we observe heterogeneity in the fraction of UK households (blue bars) residing in local tax havens across income deciles. Almost four of every five UK households at the very top of the wage distribution (top 0.1%, with annual wage incomes above CHF 3.4 million) reside in local tax havens. This fraction gradually declines as we move down the income distribution.[49,50] In contrast, the fraction of US households (red bars) living in the tax havens of the Zurich commuting zone is much more evenly distributed across income groups.

How does the differential sorting behavior translate into differences in the average tax rates that UK and US households pay in Switzerland? To shed light on this issue, we provide semi-parametric estimates of the average tax rate effectively faced by UK households (blue line) and US households (red line) across the wage distribution (c.f. Figure 5). We contrast these estimates with the average tax rate that households face, on average, when they make their location choices randomly (gray line). In what follows, we refer to this tax rate as the 'mean average tax rate'.[51]

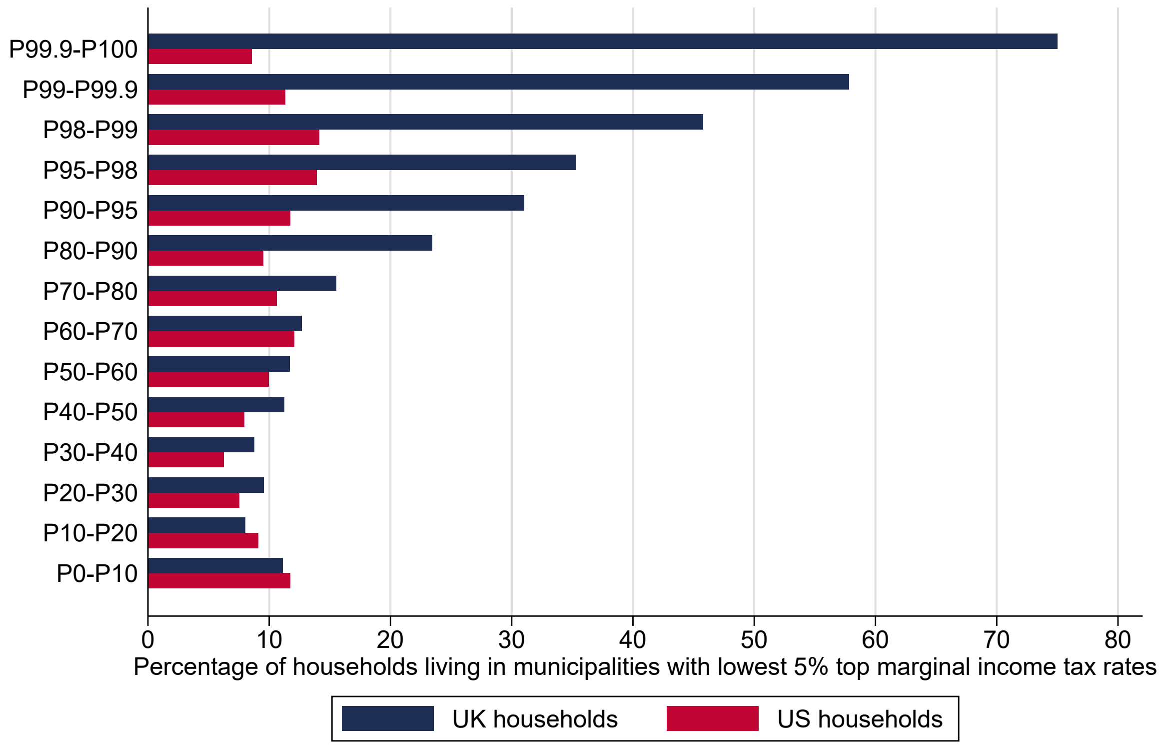
[48]Overall, this 5% account for around 8% of the overall population in the Zurich commuting zone.

[49]The share of households in the 99th to 99.9th wage percentile (with annual wages between around 1.2 and 3.4 millions) that live in the local tax heavens of the Zurich commuting zone is around 57 percent, and the one for the 98th to 99th wage percentile bracket (with annual wages between 700'000 CHF and 1.2 mio CHF) is around 43 percent. Further, the share of households in the 95th to 98th wage percentile (with annual wages between 440'000 and 700'000 CHF) that live in the local tax heavens of the Zurich commuting zone is around 35 percent.

[50]The sorting behavior is not unique to UK households. In Appendix H, we show that French, German and Swiss households, who are equally exposed to Swiss income tax as UK households, show sorting behavior similar to that of UK households. Relative to French, German and UK households, Swiss households sort the least according to income taxes, reflecting the relative importance of local social ties in location behavior for Swiss households. This renders Swiss households an inappropriate substitute for our treatment group of UK households.

[51]The figure provides estimates for all households in the area under scrutiny with an income between CHF 62,500 and CHF 4,000,000. The gray line represents the average tax rate that a specific household type would face if the population weighted mean tax rate would apply to them. In other words, it is the tax rate that a household would face on average by choosing a random municipality within the one hour commuting area around Zurich. As in Roller & Schmidheiny (2016), we calculate the mean average income tax rate by wage as $\tau_w = \sum_{t \in T} \sum_{n \in N} s_{t,n} \sum_{\theta \in \Theta} s_\theta \times \tau_{w,\theta,t,n}$. In this equation, $\tau_{w,\theta,t,n}$ is the average tax rate for a taxable entity belonging to group θ, with (joint) income w,

Figure 4: Percentage of US and UK Households in Local Tax Havens (by wage percentile, 2012-2020)



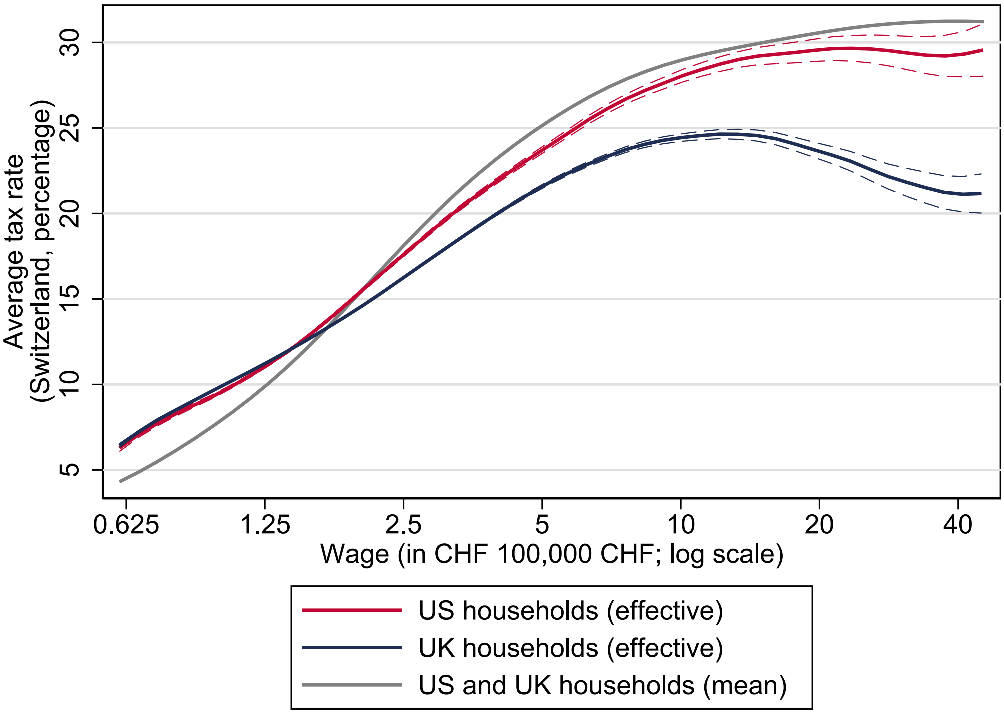
Notes: This figure compares the shares (in percentage) of US (in red) and UK (in blue) households living in low tax municipalities across wage percentiles. All municipalities with a travel time (by car) of one hour to the Zurich main railway station are considered. Low tax municipalities are defined as the 5% of municipalities with the lowest top marginal income tax rate. Percentage shares from 2012-2020 are averaged.

As shown in Figure 5, the effective average tax rate exceeds the mean average tax rate for low- and medium-income households from both the US and the UK. These households tend to be located in high-tax jurisdictions. The effective average tax rate for high income levels is lower than the mean average tax rate. This tends to be true for high-income households from both the US and the UK. However, high-income US households that are fully (or partially for moderately high incomes) isolated from the Swiss income tax schedule sort to a lesser extent into low tax jurisdictions than their counterparts from the UK.[52] This (income tax-induced) sorting effect on average

residing in municipality n at time t; s_θ is the share of a taxable entity of type θ in all taxable entity types, considering all taxable entities from the US and UK, hence $\sum_{\theta\in\Theta} s_\theta = 1$; and $s_{n,t}$ is the share of taxable entities residing in municipality n at time t in all taxable entities in Switzerland, hence $\sum_{t\in T}\sum_{n\in N} s_{t,n} = 1$.

[52]As explained above, the computed income thresholds for the Swiss tax payment to be lower than the US tax payment are not sharp such that the convergence of sorting incentives of US and UK

Figure 5: Effective and mean average tax rates (2012 to 2020)



Notes: The figure shows the average tax rate of tax payers (taxable entities) from the US and the UK in the considered area (i.e., 60 minutes travel time from Zurich's main station) by income. The red and blue lines are local polynomial regressions (of order 3) with an Epanechnikov kernel and bandwidth selection by Silverman's rule of thumb. The mean average tax rate (in gray) is calculated as follows: $\tau_w = \sum_{t \in T} \sum_{n \in N} s_{t,n} \sum_{\theta \in \Theta} s_\theta \times \tau_{w,\theta,t,n}$. In this equation, $\tau_{w,\theta,t,n}$ is the average tax rate for a taxable entity belonging to group θ, with (joint) income w, residing in municipality n at time t; s_θ is the share of a taxable entity of type θ in all taxable entity types, considering all taxable entities from the US and UK, i.e. $\sum_{\theta \in \Theta} s_\theta = 1$. $s_{t,n}$ is the share of taxable entities residing in municipality n at time t in all taxable entities in Switzerland. Hence, $\sum_{t \in T} \sum_{n \in N} s_{t,n} = 1$. The red and blue lines depict the average tax rate effectively faced by US and UK households, respectively.

income tax rates strongly increases with income level. The effective average tax rate that applies to UK households in Switzerland becomes regressive for incomes of around CHF 1.5 million and above, despite the strictly progressive tax schedule in all Swiss tax jurisdictions (Roller & Schmidheiny, 2016).[53] Using regression analysis to quantify the differential sorting effect between UK and US households on effectively paid average tax rates, UK households in the top 1% (i.e., annual household wages

households shown in Figure 5 continuously steps in as income becomes smaller.

[53]The finding that, as a consequence of the strong income segregation, average tax rates effectively faced by high-income households are much lower and decline for higher income levels is reminiscent of the findings in Roller & Schmidheiny (2016), who do not restrict the analysis to US and UK households and include all Swiss cantons and municipalities.

above CHF 1.2 million) are found to have a 23% lower average tax rate than their US counterparts. Consistent with Figure 5, the tax savings gradually decrease as we move down the income distribution and disappear for lower incomes (the full analysis can be found in Appendix C).

This observation reflects the fact that, for lower income levels, the tax payment in Switzerland is higher than the payment owed to the US federal government. US households are effectively included in the Swiss tax income system and have a sorting behavior that aligns with that of their UK counterparts. This finding is consistent with Figure 4, which shows similar sorting behavior across tax havens in the Zurich commuting zone for lower incomes (below the 70th income decile). These two observations suggest that US and UK households in Switzerland have similar moving behaviors and sort similarly when exposed to the same tax environment. These similarities extend to higher income levels. In Sections 5 and 6, we show that these two high-income groups respond similarly when non-tax location factors are included in the regressions, suggesting that they sort similarly along the non-tax dimensions of location characteristics. Overall, the evidence is in line with the reasoning presented above: conditional on being in Switzerland, the two groups share many cultural aspects in addition to having a common language, which makes them useful as treatment and control groups to isolate the effect of income taxes on location behavior.

5 Municipality-level Analysis

Figure 4 provides suggestive evidence of a strong relation between income tax rates and the location choices of UK high-income households. To emphasize the empirical pattern, we now seek to quantify the role of income tax rates in each municipality's ability to attract high-income households by means of a regression analysis.

5.1 Specification and Estimation

We begin by estimating the effect of the net-of-tax rate on the number of high income households from each country $o \in \{US, UK\}$ residing in a municipality n in year t. The distribution of US and UK households across municipalities is skewed to the left (c.f., Figure A2 in Appendix B). While some of the 12,362 municipal observations are rather high, the majority of observations contain zeros or single single-digit integers. Part of the reason for this is that the majority of UK households at the top of the income distribution live in a small number of local tax havens, as shown in Figure 4. Using the standard ordinary least square (OLS) estimator in the presence of a heavily skewed and non-negative distribution with a large share of zeros is criticized on two grounds. First, taking logs of the outcome variable drops all the observations that contain zeros. Second, the estimator does not account for the presence of heteroskedasticity. We address these issues using the Poisson pseudo-maximum likelihood (PPML) estimator proposed by Silva & Tenreyro (2006) in the trade literature, where zero values in the data are particularly common. Our main empirical specification can be described as follows:[54]

$$N_{o,n,t} = exp\,(\,\alpha\, UK_{n,t} + \beta\, UK_{n,t} \times ln(1 - \tau_{n,t}) + \mu_{n,t}) + \epsilon_{o,n,t}. \tag{1}$$

$N_{o,n,t}$ denotes the number of high income households originating from each country $o \in \{US, UK\}$ in municipality n and year t. $UK_{n,t}$ is a dummy variable equal to one for UK households (the treatment group) and zero for US households (the control group). α is the corresponding coefficient. $\tau_{n,t}$ denotes the (hypothetical) municipality- and year-specific average income tax rate averaged across all US and UK households in the Zurich commuting zone and belonging to the respective top income groups.[55] Thus, $ln(1-\tau_{n,t})$ is the log retention rate in municipality n and year

[54]We estimate the model by using the command of Correia et al. (2020), which allows fast Poisson estimation with multiple high-dimensional fixed effects.

[55]Formally, we calculate average tax by municipality n in year t as $\tau_{n,t} = \frac{1}{M}\sum_{i=1}^{M} \tau_{n,t,i}$. In this equation, $\tau_{n,t,i}$ is the (hypothetical) average tax rate that household $i = 1, ..., M$ would face by resid-

t. The use of this tax measure allows for a theory-guided estimation of the elasticity of location decisions with respect to the net-of-tax rate, given that location choices are influenced by differences in average tax rates across locations.[56] β is the coefficient of the interaction term between the UK dummy and the log retention rate. We can directly interpret this parameter as the income-tax-induced elasticity of the number of top income households located in a specific municipality with respect to the net-of-tax rate. $\mu_{n,t}$ are municipality-by-year fixed effects. Thus, we use within-municipality time variations in income taxes and locational attractiveness, which differs from income tax but correlates therewith, of each municipality-year pair to estimate the effect of interest.

While we consider equation (1) our main specification, we also provide estimates from specifications in which we replace the municipality-by-year fixed effects with just year and/or municipality fixed effects, or in which we renounce fixed effects altogether. While parsimonious in controlling for unobserved factors, the municipality-by-year fixed effect in equation (1) absorbs the effect of $ln(1-\tau_{n,t})$, thereby not allowing a separate estimation of the effect of (potentially non-observable) location-specific factors correlating with income taxes. We estimate this effect in specifications without municipality-by-year fixed effects. Finally, $\epsilon_{o,n,t}$ is the disturbance term.

In addition to the baseline model described above, we present additional specifications in Subsection 5.4 to analyze whether our estimates are driven by modeling

ing in municipality n at time t. M are all US and UK households that belong to a respective top income group (top 10% in our baseline analysis, top 1%, top 1-5%, or top 5-10% in our analysis on heterogeneous effects) of the one-hour-commuting area around Zurich.

[56]In general, average tax rates might be endogenous and potentially bias the estimates. A common concern is that average tax rates depend on gross wages. This likely leads to an overestimation of counterfactual average tax rates, since households tend to select themselves in locations with higher gross wages. Our empirical design addresses this concern. Households are able to choose the residential location within the local labor market around Zurich, while leaving the workplace and gross wage constant. A related concern is that amenities influence average tax rates as well the location choice. The set of relevant amenities is potentially imperfectly observed, which gives rise to an estimation bias whose sign depends on the correlation between amenities and average tax rates. In our setting, the comparison between US and UK households allows to account for this concern and to absorb the effect of the (potentially unobserved) set of amenities on location choice.

choices, moving costs or systematic differences in preferences for amenities between US and UK households. We also examine whether our results are sensitive to our definition of the commuting zone of one hour around Zurich by replicating all our estimates for a commuting zone of 45 minutes around Zurich in Section D of the Appendix. Throughout the analysis, standard errors are clustered at the municipality level[57].

5.2 Baseline Results

We begin our analysis by estimating the effect of interest for the top-10%-income households from the US and UK, gradually including fixed effects (c.f. Table 2). We can directly interpret the obtained parameter values as elasticities. The coefficient on $ln(1-\tau_{n,t})$ provides elasticity estimates of the top-10% -income households from the US (the control group) with respect to the retention rate. While US households do not react to income tax itself, they likely base their location choices on unobservable factors that are linked to income-tax rates, such as geographical amenities, local public services, and housing costs. The coefficient of $ln(1-\tau_{n,t})$ thus captures the indirect or non-income-tax-induced effect of the log retention rate on household location choice.

The obtained values are almost zero and insignificant in specifications without municipality dummies (columns 1 and 2). In contrast, we obtain negative estimates of around –5 to –6 when controlling for unobserved municipality-specific characteristics, which allows for a much more precise estimation of the effect of non-tax factors on location behavior. It is important to note that we cannot estimate the coefficient on $ln(1-\tau_{n,t})$ in the specification with the year-specific municipality dummies presented in column 5 due to perfect collinearity.

[57] In unreported robustness checks, we cluster standard errors by both municipality and year. The estimated standard errors only marginally change and, as a consequence, the significance levels of the estimated coefficients remain essentially the same in all cases. The results are available upon request.

Table 2: Municipal-Level Elasticities: Baseline Effects (2012-2020)

	(1)	(2)	(3)	(4)	(5)
$ln(1-\tau_{n,t})$	0.025	-0.002	-5.866*	-5.346*	
	(11.941)	(12.058)	(2.592)	(2.335)	
$UK_{n,t} \times ln(1-\tau_{n,t})$	18.221*	18.423*	10.969***	10.987***	11.014***
	(7.403)	(7.438)	(1.914)	(1.927)	(1.936)
Municipality FE			✓	✓	
Year FE		✓		✓	
Municipality×year FE					✓
Observations	12,362	12,362	12,362	12,362	12,362
Estimator	PPML	PPML	PPML	PPML	PPML

Notes: This table reports the baseline estimates from equation (1) using the PPML estimator proposed by Silva & Tenreyro (2006). The dependent variable is the number of high income households (top 10%) with principal earner nationality $o \in \{US, UK\}$ in municipality n and year t that belong to the top wage decile of US and UK households residing in the one-hour-commuting zone around Zurich. The independent variables are a dummy variable, $UK_{n,t}$, equal to one for UK households and zero for US households, and the log retention rate in municipality n and year t, $ln(1-\tau_{n,t})$ (except specification 5 where the retention rate is omitted due to perfect collinearity with the municipality-by-year fixed effects); and the interaction term, $UK_{n,t} \times ln(1-\tau_{n,t})$. Standard errors clustered at the municipality level are provided in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

In general, the sign of the coefficient on $ln(1-\tau_{n,t})$ is unclear a priori, as it depends on the relationship between income taxes and amenities. If high-tax jurisdictions are more attractive for non-tax reasons, for example, because they can provide superior public amenities or when high tax rates capitalize into (lower) housing costs, then the coefficient of the log retention rate should be negative. Conversely, if non-income-tax factors of attractiveness are negatively correlated with tax rates, for example, because high tax rates correspond to spending requirements due to social problems in the municipality, then the coefficient of the log retention rate will likely be positive. Our findings indicate the former. High-income US households tend to locate in high-tax municipalities to benefit from the better amenities without being directly exposed to the high income tax.

The third row of Table 2 provides the point estimates of the effect of $UK_{n,t} \times ln(1-\tau_{n,t})$, our coefficient of interest. We interpret the coefficient value as the purely income-tax-induced elasticity of the number of top-income households with respect to the retention rate. The obtained values are positive and highly significant. In our preferred specification with interacted municipality-year fixed effects, we obtain a coefficient value of around 11. A 1% increase in the retention rate leads to a rise of UK top income households by around 11%. This value is barely affected by the year controls. In contrast, coefficient estimates are larger when we do not control for municipality-specific characteristics (columns 1 and 2). However, in these cases, the obtained coefficients become less significant as the standard errors increase. This reflects the higher standard error on the estimated coefficient on $ln(1 - \tau_{n,t})$ for our control group, which affects the precision of the estimation of the interaction effect $UK_{n,t} \times ln(1 - \tau_{n,t})$.

We relegate a more in-depth discussion of the estimates and the relation to existing empirical studies to Section 7.

5.3 Heterogeneous Results

In this section, we analyze whether high-income earners' responses to income taxes vary with income levels. In doing so, we run our preferred specification with municipality-by-year fixed effects, as specified by equation (1), for three different top income groups: households in top 1%, the top 1-5%, and the top 5-10%.

Columns 1-3 of Table 3 present the point estimates and standard errors on $UK_{n,t} \times ln(1 - \tau_{n,t})$. The estimates obtained for all three top income groups are highly significant. We obtain point estimates of around 13.5 for the top 5–10% group (with an average annual income of CHF 417,000) and 10.6 for the top 1–5% group (average of CHF 722,000). We find a larger income-tax-induced elasticity of around 16 for the top 1% group, with an average annual wage income of around CHF 2.4 million. This finding is in line with the prediction of a higher propensity of households with higher incomes to move to low-tax jurisdictions. In contrast, we do not find much difference in point estimates between the top 1-5% and top 5-10% income groups (i.e., among

Table 3: Overall Elasticities by Income Bracket (2012-2020)

	top 1%	top 1–5%	top 5–10%
	(1)	(2)	(3)
$UK_{n,t} \times ln(1-\tau_{n,t})$	16.060***	10.627***	13.542***
	(1.774)	(2.072)	(3.273)
Municipality×year FE	✓	✓	✓
Observations	12,362	12,362	12,362
Estimator	PPML	PPML	PPML

Notes: This table reports estimates from equation (1) for the different top wage percentile ranges using the PPML estimator proposed by Silva & Tenreyro (2006). The dependent variable is the number of high-income households with principal earner nationality $o \in \{US, UK\}$ in municipality n and year t that belong to the specified top wage percentile ranges. 'Top 1%' refers to households that belong to the 99th to 100th wage percentile while 'top 1–5%' refers to households that belong to the 95th to 98.99th wage percentile, etc. The independent variables are a dummy variable, $UK_{n,t}$, equal to one for UK households and zero for US households, and the interaction term between the UK dummy and the retention rate in municipality n and year t, $UK_{n,t} \times ln(1-\tau_{n,t})$. Standard errors clustered at the municipality level are provided in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

high-income households with annual incomes generally below CHF one million). In terms of statistical significance, the three point estimates are not overly different, as the confidence intervals of the estimates largely overlap. Indeed, a bootstrapped test for the difference of coefficients with B=200 bootstrap replications confirms that the estimated coefficient reported in column (3) is not statistically different from both the one reported in column (1) and the one reported in column (2). The p-value of this test is 0.383 in the first case and 0.290 in the second case. The null hypothesis in both cases is that there is no difference in the estimated coefficients.

Income tax burdens differ across locations due to income tax variations across cantons and municipalities. The setting provides a testing ground for the question of whether the responsiveness of high-income households to tax differences within a canton differs from the responsiveness to the overall tax variations (i.e., including cantonal tax variations). To test for this, we add canton-by-origin-by-year dummies to equation (1). Compared to the regressions in Table 3, we absorb the cantonal-

Table 4: Within-Canton Elasticities by Income Bracket (2012-2020)

	top 10%	top 1%	top 1–5%	top 5–10%
	(1)	(2)	(3)	(4)
$UK_{n,t} \times ln(1-\tau_{n,t})$	13.835**	12.555***	13.628*	19.140**
	(4.561)	(3.303)	(5.441)	(6.599)
Municipality×year FE	✓	✓	✓	✓
Canton×origin×year FE	✓	✓	✓	✓
Observations	12,362	12,362	12,362	12,362
Estimator	PPML	PPML	PPML	PPML

Notes: This table reports within-canton estimates for different top wage percentile ranges using the PPML estimator proposed by Silva & Tenreyro (2006). The specification is equivalent to equation (1) but includes canton-by-origin-by-year fixed effects. The dependent variable is the number of high income households with principal earner nationality $o \in \{US, UK\}$ in municipality n and year t that belong to the specified top wage percentile ranges. 'Top 10%' refers to households that belong to the 90th to 100th wage percentile while 'top 1%' refers to households that belong to the 99th to 100th wage percentile, etc. The independent variables are a dummy variable, $UK_{n,t}$, equal to one for UK households and zero for US households, and the interaction term between the UK dummy and the retention rate in municipality n and year t, $UK_{n,t} \times ln(1-\tau_{n,t})$. Standard errors clustered at the municipality level are provided in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

by-time variation in income taxes that is specific to the origin of households, such as the tax effect for UK households, which is due to cantonal tax variation. Table 4 provides the obtained coefficient estimates for the different top income intervals. The estimated elasticity ranges from around 13-19. For the top 1-5% and the top 5-10% income groups, the estimated elasticity exceeds what we obtain using the overall tax variation. This finding is in line with the general reasoning that household mobility among proximal municipalities is stronger than mobility across cantons. For the top 1% group of households, we observe a higher point estimate, indicating that the very high-income group responds more strongly to cantonal tax variation. As shown in Figure 3, cross-municipal variations within cantons are limited to around 2 to 5 percentage points, and larger tax savings can be realized by moving across cantons. This appears to be relevant for the very high-income group. However, we should take the interpretation with a grain of salt, as the confidence intervals between 'overall' and 'within canton' estimates overlap.[58]

5.4 Sensitivity Checks

We run a set of sensitivity checks to test whether our main estimates are driven by modelling choices, moving costs, or systematic differences in preferences for specific amenities between US and UK households.

Naive OLS model We begin by estimating a log-log specification using OLS. This model drops all observations for which there are zero households from a specific country of origin (US or UK) in a given municipality and year, thus dropping 84% of all observations. The obtained elasticity on the interaction term between the log retention rate and the UK dummy from a specification without fixed effects is presented in column 1 of Table 5, while that from a model with interacted municipality-by-year fixed effects is shown in column 4. The obtained estimate for the top-10% income households is comparable to that obtained through the PPML estimator when including municipality-by-year fixed effects, and it is lower without the fixed effects (c.f., columns 1 and 5 in Table 2). In general, the naive log-log model introduces a bias that likely increases as the share of dropped observations increases. Given the high number of dropped values, the OLS estimate tends to be less preferable. In other specifications, we experimented with adding a small constant, c, to the outcome variable before taking logs. When using $ln(N_{o,n,t} + c)$ or the inverse hyperbolic sine transformations, the estimates are largely affected by the choice of the constant.[59] This leads us, as suggested by Silva & Tenreyro (2006), to choose the PPML as our preferred model.

[58]Indeed, a bootstrapped test for the difference in coefficients with 200 bootstrap replications confirms that there is no statistically significant difference between estimate (1) in Table 3 and estimate (2) in Table 4 (p-value is 0.631), between estimate (2) in Table 3 and estimate (3) in Table 4 (p-value is 0.591), and between estimate (3) in Table 3 and estimate (4) in Table 4 (p-value is 0.675). In all three cases, the null hypothesis of the test is that there is no difference between the estimated coefficient in Table 3 and the corresponding coefficient in Table 4.

[59]The findings are available upon request.

Table 5: MUNICIPAL-LEVEL ELASTICITIES: SENSITIVITY CHECKS (2012-2020)

	(1)	(2)	(3)	(4)	(5)	(6)
$ln(1-\tau_{n,t})$	-0.248	6.320	5.313			
	(1.848)	(8.014)	(7.263)			
$UK_{n,t} \times ln(1-\tau_{n,t})$	8.053***	16.387***	15.531**	9.673***	10.314***	8.932**
	(1.650)	(3.266)	(4.753)	(1.080)	(0.951)	(2.962)
$ln(Time\ to\ Zurich_n)$			-2.210***			
			(0.486)			
$UK_{n,t} \times ln(Time\ to\ Zurich_n)$			0.506			0.418
			(0.322)			(0.261)
$Access\ to\ lake_n$			1.365***			
			(0.365)			
$UK_{n,t} \times Access\ to\ lake_n$			-0.446			-0.133
			(0.259)			(0.170)
$ln(Public\ spending_{n,t})$			5.132***			
			(1.371)			
$UK_{n,t} \times ln(Public\ spending_{n,t})$			-0.928			-0.839
			(1.072)			(0.561)
$ln(Wealth\ tax_{n,t})$			-0.207			
			(0.641)			
$UK_{n,t} \times ln(Wealth\ tax_{n,t})$			0.864			0.366
			(0.489)			(0.377)
$Homophily_{n,t}$			0.694			
			(1.026)			
$UK \times Homophily_{n,t}$			1.502*			0.369
			(0.650)			(0.627)
Municipality controls			✓			✓
Only mover		✓			✓	
Municipality×year FE				✓	✓	✓
Estimator	OLS	PPML	PPML	OLS	PPML	PPML
Observations	1,708	12,362	12,362	1,708	12,362	12,362

Notes: This table reports a series of robustness checks based on the same data set as in Tables 2 and 3. The dependent variable is the number of high-income households with principal earner nationality $o \in \{US, UK\}$ in municipality n and year t that belong to the top wage decile of US and UK households residing in the one-hour-commuting zone around Zurich. The independent variables of all specifications include a dummy variable, $UK_{n,t}$, equal to one for UK households and zero for US households. In specifications 1 and 4, observations with a value of zero are dropped. Section B of the Appendix provides a detailed description of the municipality covariates used in specifications 3 and 6. Standard errors clustered at the municipality level are provided in parentheses: $^{*}\ p < 0.05$, $^{**}\ p < 0.01$, $^{***}\ p < 0.001$.

Restriction to moving households Next, we restrict our sample to observations in which households relocated to a new municipality.[60] Focusing on 'movers' is interesting because moving to a new location is costly. Limiting the analysis to moving households eliminates the potential bias of including households that stayed in a sub-optimal location per se because of the high monetary and psychological costs of moving (Schmidheiny, 2006; Agrawal & Foremny, 2019).[61] Our analysis of movers is based on 1,708 moves of households within the top 10% of the income distribution. The obtained elasticity is presented in columns 2 (without fixed effects) and 5 (with municipality-by-year fixed effects) of Table 5. Our mover-based estimates barely differ from our baseline estimates (presented in columns 1 and 5 of Table 2). This suggests that our baseline estimates are not downward-biased by moving costs. The finding is in line with the view that moving costs are relatively small within the one-hour commuting zone around Zurich and do not represent a significant barrier to moving to a more optimal location in this area.

Preference differences Next, we aim to test the validity of our key assumption. Our natural experiment allows for a simple identification of the pure income-tax-induced effect on the location choice of anglophone high-income households. However, our identification strategy would not be valid if high-income households from the US and UK had systematically different preferences for location-specific amenities or wealth tax rates. We address this concern by collecting data on top marginal wealth tax rates, per-capita public spending, and natural amenities, such as munic-

[60]The re-locations include both, re-locations from places in- (32 percent of all moves) and outside of the Zurich commuting zone (68 percent of all moves). As detailed in the data section, some households consist of jointly taxed couples. In most cases, these couples either move together or from different places to form a household in a new municipality. In rare cases where one person of a married couple joins the other, only moves of the principal earners enter our data-set. Moreover, the vast majority of households only move once over the course of our sample. Some households move multiple times, in which case they enter our data multiple times.

[61]Another motivation for a restriction to moving households is that tax rates are likely a function of all households location decisions, which may raises endogeneity concerns (Brülhart et al., 2015) in an analysis on the overall population. In our case, however, equilibrium tax rates are less likely driven by the US and UK population as they only account for a relatively small share (2.6%) of the overall population in the Zurich commuting zone.

ipalities' access to a lake and travel time to Zurich, which proxies for the ability to consume amenities (cultural and business activities, for instance) offered by Zurich (see Section 2 of the Appendix for descriptive statistics and detailed information on these variables). We also control for homophily effects, that is, the preference to locate next to anglophone households of the same income group.[62] We directly add these variables and their interactions with the UK dummy to our model. Significant interaction effects indicate systematic differences in the preferences for specific amenities across US and UK households. However, as shown in columns 3 and 6 of Table 5, we do not find evidence of systematic differences between our control and treatment group.[63,64]

The findings in Table 5 suggest that controlling for observable amenities does not allow for uncovering the pure tax effect on location choice. The estimate on $UK_{n,t} \times ln(1 - \tau_{n,t})$ is only slightly affected by the inclusion of these variables (c.f., column 1 and 5 in Table 2 and columns 3 and 6 in Table 5). Thus, controlling for observable location-specific factors of attractiveness does not appear to be a substitute for the empirical approach of comparing two groups of households with fundamentally different exposures to local taxes, but otherwise similar location behaviors. The treatment-control group approach allows for the inclusion of municipality-by-year fixed effects to comprehensively absorb confounding factors, which might otherwise bias the estimate of the tax effect on location choice.[65] This renders the finding in

[62]To this end, we define the homophily variable as the number of anglophone households in a specific municipality and year that belong to the top 10% income group over the total number of the Swiss and foreign population in the same municipality and year.

[63]As expected, we obtain negative coefficients on the commuting time to Zurich (with an elasticity with respect to the control variable, rather than the interaction term, of -2.3), but positive effects on public spending (5.2) and access to a lake (1.4). However, as these variables affect UK and US households simultaneously, we cannot estimate these level effects in a specification with municipality-by-year dummies.

[64]Public spending and wealth taxes are policy variables and may respond to the sorting behavior. In Appendix G, we show that there is no statistically significant impact of the number of US/UK households on neither wealth tax rates nor public spending.

[65]In the absence of the control group (i.e., in a setting in which all households are treated and subject to local taxes), municipality-by-year fixed effects are perfectly collinear with the main independent variable (the average tax rate at the municipal level). They are incapable of selectively absorbing unobservable location-specific factors that differ from income taxes, but correlate with income taxes.

column 6 of Table 5 our preferred estimate, indicating an elasticity value of slightly below nine.

Alternative Definition of the Zurich Commuting Zone Finally, we replicate our estimates of this section by considering all municipalities located within 45 minutes rather than one hour of the City of Zurich. These results are shown in Section D of the Appendix. The estimates are consistent with those presented in this section. For instance, for our preferred specification with municipality-by-year fixed effects, the point estimate is 11.4 in the smaller commuting zone compared to 11.0 in the one-hour commuting zone (c.f., column 5 of Table A4 and column 5 of Table 2).

6 Household-level Analysis

We now draw on household-level location choice models to analyze how taxes affect the location choices of anglophone households in the one-hour-commuting zone around Zurich. The empirical approach adopted in the previous section might be criticized on the grounds that unobserved heterogeneity between US and UK households biases the municipality-level estimates. In that analysis, we test for different preferences of US and UK households for key geographical variables, publicly provided amenities, and wealth taxes, and the estimates do not indicate systematic differences. However, other time-varying jurisdiction-specific factors might correlate with possible differences in characteristics across US and UK households. A household-level analysis allows us to address this concern by controlling for household-specific characteristics. Moreover, it allows us to exploit available information on gross wages, marital status, income splits, and number of children to calculate household-specific average tax rates for each location.

For the sake of computational feasibility, and this section only, we conduct our estimation at the cantonal level, using the five cantons within which all municipalities within a commuting distance of one hour to the City of Zurich are located as our choice variables. These are the cantons of Aargau, Lucerne, Schwyz, Zug, and Zurich (see Table A2 in Appendix B for detailed summary statistics on these cantons).

6.1 Utility Specification

Following Kleven et al. (2013) and Akcigit et al. (2016), the total utility from choosing canton c at time t for household i is given by

$$U_{c,t,i} = u\left(w_{t,i}(1 - \tau_{c,t,i})\right) + \mu_{c,t,i}. \tag{2}$$

The gross wage of household i is constant within the Zurich commuting zone and denoted by $w_{t,i}$. Canton- and household-specific average income tax rates are denoted by $\tau_{c,t,i}$.[66] In addition to disposable income after taxes $w_{t,i}(1 - \tau_{c,t,i})$ household i also gains utility from locating in canton c in year t, represented by the idiosyncratic preference component $\mu_{c,t,i}$. Household i chooses to locate in canton c only if

$$u\left(w_{t,i}(1 - \tau_{c,t,i})\right) + \mu_{c,t,i} \in max_{c' \in \mathcal{C}}\left\{u\left(w_{t,i}(1 - \tau_{c',t,i})\right) + \mu_{c',t,i}\right\},$$

where $\mathcal{C}$ denotes the set of cantons from which to choose. Based on the utility function defined by equation (2), we obtain the econometric specification for the discrete choice model at the level of household i in year t by assuming the log utility of consumption:[67]

$$U_{c,t,i} = \alpha\, log(1 - \tau_{c,t,i}) + \gamma_c \boldsymbol{x}_{t,i} + \lambda \boldsymbol{x}_{c,t} + \eta_c + v_{c,t,i}. \tag{3}$$

[66]Since choice alternatives are at the cantonal level, average tax rates (aggregated over the municipal, cantonal and federal level) are household-specific averages over all municipalities of a given canton. Using population-weighted averages hardly changes the results.

[67]Assuming log utility of consumption allows us to write utility as additively separable in gross wages and retention rates. Note, given our focus on the one-hour commuting zone around the City of Zurich, households can select among the different residential alternatives for a given job choice. Thus, gross wages are not alternative specific and the variable cancels out of equation (3).

In our most comprehensive specification, household i's idiosyncratic preference component of canton c and year t, $\mu_{c,t,i}$, is assumed to depend on the following:

(i) Household-level characteristics, $\boldsymbol{x}_{t,i}$, including age, gender and years of residency in Switzerland of household i's principal earner, as well as marital status and the number of children (see Table 1 for the summary statistics of these variables). The effect of all these household characteristics, denoted by γ_c, may vary by canton.

(ii) Canton-specific variables, $\boldsymbol{x}_{c,t}$, which include a variety of characteristics, such as commute time to Zurich, public spending per capita, wealth tax rates, and the share of municipalities that have access to a lake. See Table A2 in Appendix B for summary statistics for these variables. Moreover, we account for homophily, meaning the preference of households to locate in the same canton as their peers in the same income range. Specifically, we define this variable as the number of anglophone households in a specific canton and year that belong to a respective top income group – top 1%, top 1–5%, top 5–10%, top 10% – over the total number of Swiss and foreign households in that canton and year.

(iii) The choice-specific constant, η_c, which captures all unobserved canton-specific effects that are constant across anglophone top-income households.

We estimate the model in equation (3) separately for households of US and UK origin, yielding the estimates of interest α_{UK} and α_{US}.

6.2 Estimation and Computation of Elasticities

We denote $P_{c,t,i} \equiv Prob(U_{c,t,i} \geq U_{c',t,i} \; \forall \; c')$ as the probability of household i to locate in canton c at time t. If the error term $v_{c,t,i}$ is independent and identically distributed with a type I extreme value distribution, we can estimate this model using a multinomial logit regression.

Based on the coefficients of the retention rate, we can compute the elasticity of the location choice (Kleven et al., 2013; Akcigit et al., 2016). In the multinomial choice model, the elasticity of the probability of household i locating in canton c with respect to a net-of-tax rate $(1-\tau_{c,t,i})$, denoted by $\epsilon_{c,t,i}$, is

$$\epsilon_{c,t,i} \equiv \frac{dlog(P_{c,t,i})}{dlog(1-\tau_{c,t,i})} = \Delta\alpha(1-P_{c,t,i}), \tag{4}$$

where $\Delta\alpha = \alpha_{UK} - \alpha_{US}$ is the coefficient of the retention rate of UK nationals relative to the control group of US nationals. The difference $\Delta\alpha$ captures the pure income tax effect on location choice. We can then define the tax-induced elasticity in canton c, ϵ_c. Letting I_c be the set of treated households residing in canton c provides the following:

$$\epsilon_c \equiv \frac{dlog(\sum_{i\in I_c} P_{c,t,i})}{dlog(1-\tau_{c,t})} = \Delta\alpha\frac{\sum_{i\in I_c} P_{c,t,i}(1-P_{c,t,i})}{\sum_{i\in I_c} P_{c,t,i}}. \tag{5}$$

Average elasticities ϵ are then defined as the weighted average elasticities across all cantons:

$$\epsilon \equiv \sum_{c\in\mathcal{C}}\left(\frac{\sum_{i\in I_c} P_{c,t,i}}{\sum_{c\in\mathcal{C}}\sum_{i\in I_c} P_{c,t,i}}\right)\epsilon_c. \tag{6}$$

6.3 Baseline Results

Table 6 provides our baseline results. Column 5 presents the results from our preferred specification, which contains the controls described in Subsection 6.1. Columns 1-4 provide the results from the less saturated models. In addition to the regression estimates, the table provides the values and standard errors of the location elasticity for the top 10% income group, as derived in equation (6). The computed elasticities are fairly constant across the five specifications, ranging between 8.2-8.3. The constancy of elasticity when adding household and location characteristics is in line with the empirical evidence provided in Section 4, suggesting that US and UK households share many cultural aspects in addition to a common language, which makes them useful as treatment and control groups to isolate the effect of income taxes on location behavior. Otherwise, we would expect elasticity values to vary when controlling for household and location characteristics that may correlate with, for instance, utility gained from natural and publicly-provided amenities, for instance.

Table 6: Location Choice Models: Baseline Results

	(1)	(2)	(3)	(4)	(5)
Coefficient estimate UK (treatment group)					
Log retention rate	11.548***	7.985***	11.854***	11.573***	7.629***
	(1.185)	(1.482)	(1.250)	(1.304)	(1.757)
Coefficient estimate US (control group)					
Log retention rate	-2.509	-6.081**	-2.237	-2.815	-6.873**
	(1.580)	(2.342)	(1.621)	(1.721)	(2.097)
Average elasticity (UK – US)	8.219***	8.166***	8.269***	8.289***	8.339***
	(0.703)	(0.693)	(0.707)	(0.715)	(0.715)
Homophily control		✓			✓
Canton controls			✓		✓
Household controls				✓	✓
N	61,310	61,310	61,310	61,310	61,310

Notes: Multinomial logit regressions. Regressions based on the matched Census-SSER-Tax data described in Section 3. The observed time period is 2012–2020. Top 10% of the combined income distribution of UK and US households in Switzerland are considered. The location alternatives include five cantons: Aargau, Lucerne, Schwyz, Zug, and Zurich. The first row reports the coefficient of the log retention rate interacted with an indicator variable for being a UK household, while the second row reports the standard error. Analogously, the subsequent two row pairs report the coefficients and standard errors for US households. The average elasticities are computed following equation (6) with US households as a control group (Kleven et al., 2013; Akcigit et al., 2016). Standard errors are clustered at the household level: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

The computed elasticities are somewhat below those obtained from the baseline municipal-level specification of around 11 in column 5 of Table 2. The most apparent reason for this difference is that the household-level analysis focuses on the variation in retention rates across cantons. Thus, our household location choice model excludes within-cantonal variations in retention rates. When putting both models on the same footing by comparing estimates that follow from across-cantonal variations, the difference tends to vanish. In the municipal-level regression, within-cantonal (across municipalities) variation yields a higher elasticity value than the total variation in retention rates – a point estimate of 13.8 compared to 11 (c.f. column 5 in Table 2 and column 1 in Table 4). The across-cantonal variation thus gives an estimate below 11. A different reason is that we do not control for amenities, homophiliy, and wealth taxes in the baseline estimation of the municipal-level specification. When doing so (as we do in the household-level location choice model in Table 6), the elasticity value is slightly below nine (c.f. column 6 in Table 5).

As shown in Table 6, and consistent with our municipal-level analysis, we obtain negative coefficient values for US households. The coefficient values of the multinomial logit model are hard to interpret directly.[68] Nevertheless, as described in Section 5, the negative coefficient values for US households indicate a positive correlation between tax levels and non-observed determinants of jurisdictional attractiveness. A possible explanation is that higher tax levels lead to lower housing prices[69], and better natural or publicly provided amenities. US households tend to locate in high-tax jurisdictions to benefit from these advantages without being subject to high income taxes.

We obtain lower coefficient estimates when controlling for homophily, that is, households' preference to locate in the same canton as their peers of the same income range (c.f., columns 2 and 5 in Table 6). Controlling for homophily shifts the obtained coefficients of UK and US households equally downward. It clearly unravels the preference of US households to locate in high-tax jurisdictions for amenity rea-

[68]The coefficient values represent changes in the log-odds ratios of locating in canton c associated with a one-unit increase in the independent variable (i.e., log retention rate).

[69]Income taxes capitalize into housing rents in Switzerland, with a capitalization rate of roughly one half (Basten et al., 2017).

sons, as shown by the magnitude of the estimate and its statistical significance. This downward shift in coefficient estimates suggests that the preference to locate near their peers makes the two groups more willing to locate in low-tax jurisdictions. This preference applies equally to both household groups, leaving the computed elasticities – calculated as the responsiveness of UK households relative to the US control group – largely unaffected. This finding is in line with the view that the treatment-control group analysis absorbs effects of location determinants other than the pure effect of income taxes.

Finally, adjusting the location elasticity by netting out location-specific factors that correlate with income taxes has a non-negligible impact on elasticity value. For instance, using the estimates of the specification with a full set of controls (c.f. column 5 in Table 6), the non-adjusted location elasticity would be roughly half the size of the adjusted elasticity.[70] In the absence of the adjustment, location elasticity combines two opposing incentives: locating in a low-tax jurisdiction to save on taxes and locating in a high-tax jurisdiction due to location-specific factors that positively correlate with tax level. The former incentive is the pure tax effect on location choice and should influence location elasticity only.

6.4 Heterogeneous Results

Next, we analyze whether the households' reactions to local income tax differentials vary across income levels. Therefore, we interact retention rates with an indicator variable for whether a UK household is in the top 1%, the top 1–5% or the top 5–10% of the income distribution. Table 7 provides the computed values and corresponding standard errors of the elasticities for the three top-income intervals (see Table A8 in Appendix E for the obtained coefficient estimates on the log retention rates). Column 5 of Table 7 shows the results from our preferred specification, which includes all the

[70] From equations (5) and (6), the adjustment can be undone by multiplying the average elasticity in column 5 in Table 6 by a factor of $\Delta\alpha/\alpha_{UK} \approx 0.52$.

controls detailed in Subsection 6.1. The computed values from more parsimonious models are reported in columns 1-4. For all three top-income groups, the computed elasticities are highly significant and stable across the different specifications. The computed values tend to be larger for higher top-income groups, suggesting that the reaction of households to local tax differentials increases with income level.

Table 7: LOCATION CHOICE MODELS: HETEROGENEOUS RESULTS

	(1)	(2)	(3)	(4)	(5)
Average elasticities (UK – US)					
Top 1%	8.942***	8.942***	9.012***	9.017***	9.068***
	(1.292)	(1.292)	(1.305)	(1.337)	(1.325)
Top 1–5%	7.294***	7.293***	7.351***	7.315***	7.358***
	(0.865)	(0.866)	(0.872)	(0.874)	(0.881)
Top 5–10%	8.628***	8.627***	8.696***	8.867***	8.926***
	(1.117)	(1.117)	(1.124)	(1.144)	(1.149)
Homophily control		✓			✓
Canton controls			✓		✓
Household controls				✓	✓
N	61,310	61,310	61,310	61,310	61,310

Notes: Multinomial logit regressions. Regressions based on the matched Census-SSER-Tax data described in Section 3. The observed time period is 2012–2020. The location alternatives include five cantons: Aargau, Lucerne, Schwyz, Zug, and Zurich. The average top 1 elasticity is the elasticity of the top 1% households with respect to the net-of-tax rate (one minus the average tax rate), while the top 1–5 elasticity is the corresponding elasticity for the top 1–5% households, etc. The average elasticities are computed following equation (6), with US households as the control group (Kleven et al., 2013; Akcigit et al., 2016). Coefficient estimates of the log retention rates (interacted with top income group dummies) are provided in Table A8 in Appendix E. Standard errors are clustered at the household level: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

The elasticity of top 1% income households to the retention rate ranges between 8.9 and 9.2, with a value of 9.1 in our preferred specification with household, canton and homophily controls. For the top 1–5% and the top 5–10% brackets, the computed elasticities range between 7.3 and 7.4, and 8.8 and 8.9, respectively. As in our baseline analysis, these elasticities are below those obtained from our municipal-level analysis. Again, the most apparent reason for this difference is that the household-level location choice models only exploit the variation in retention rates across cantons, while the PPML models presented in Section 5 exploit both, cantonal and municipal variations.

Finally, when comparing point estimates over income distribution, the elasticity values are smallest for the top 1-5% income group. The slight U-shape pattern is robust across specifications, a finding reminiscent of our municipal-level analysis in Section 5.[71] However, the pattern should be interpreted with care. Confidence intervals of the estimates overlap, implying that the difference in the point estimates is not statistically significant at conventional levels. Indeed, when performing a bootstrapped test for the difference in coefficients with $B = 200$ bootstrap replications on the model specification (5) in Table 7, we find that, for example, the average elasticity for the top 5-10% is not statistically significantly different from the average elasticity for the top 1-5% (the corresponding p-value is 0.603 for the null hypothesis of these two average elasticities being equal).

7 Discussion

The environment we use in this study lends itself to the credible estimation of a residential elasticity for different reasons. First, households sort within a labor market area among a sizable set of municipalities while leaving their workplaces unchanged. Location choice may follow a sequential decision-making process in which a household chooses a tax residence after its workplace has been chosen and the decision to move to Switzerland has been set. Costs of coordinating job and residential choices do not downward bias our estimate of residential elasticity. Our estimate of the residential elasticity is structurally different to estimates obtained in the context of migration with a simultaneous choice of tax residence and workplace, where the

[71] Using the overall variation in tax rates in Table 3, the estimated elasticities follow a U-shape pattern. The pattern vanishes when relying on within-canton variation only (c.f. columns 2 to 4 in Table 4). Thus, the across-canton variation in point estimates appears to follow a U-shape pattern.

coordination of residential and job choices likely reduces migration responses. The 'coordination discount' appears to be substantial. Existing studies find smaller elasticity estimates of nearly two or less.[72] To isolate the effect of coordination costs, we might compare this estimate with the UK residential elasticity that is not adjusted for non-tax location factors (absorbed by the US household estimate), but uses conventional non-tax location factors as controls. The non-adjusted elasticity is around four.[73]

Second, amenities influence location choices and potentially downward bias estimates of the residential elasticity. The supply of amenities may correlate positively with tax level, and the preference for amenities in isolation implies a location choice in high-tax jurisdictions, which is in line with our findings for US households (Tables 2 and 6).[74] Comparing UK and US households in the local labor market of Zurich allows to flexibly control for the effect of amenities. In line with the general reasoning, it yields a tax effect on residential choice that is higher than the estimate including the effect of amenities that correlate with income taxes. The difference in estimates is non-negligible. Accounting for the effect of amenities, the elasticity estimate increases by a factor of roughly two.

Third, the analysis allows to address various additional endogeneity concerns. Households can move within the local labor market while keeping the workplace and thus gross wages unchanged. The local labor market approach addresses concerns that counterfactual gross wages across different potential residential locations might be imprecisely inferred, which gives rise to an estimation bias in a progressive income tax system. Also, shocks to a jurisdiction that affect migration incentives and that correlate with income taxes might not be observed. The potential omitted

[72]See Kleven et al. (2013), Kleven et al. (2014), Akcigit et al. (2016), Moretti & Wilson (2017) and Agrawal & Foremny (2019), for instance.

[73]As shown in Section 6, using the estimates of the specification with a full set of controls (c.f. column 5 in Table 6), the non-adjusted UK residential elasticity would be roughly half the size of the adjusted elasticity, i.e. it would take a value of around four.

[74]This particularly applies when preferences for amenities are stronger for high-income households, as suggested by the stronger sorting of high-skilled households in high-amenity locations (Diamond, 2016; Fajgelbaum & Gaubert, 2020).

variable bias will be less of a concern in this setting since the comparison with the control group absorbs the unobserved effect. The same holds true for correlated shocks across different jurisdictions within the local labor market around the City of Zurich (Borusyak et al., 2022).

Our findings deliver several insights for the urban and public economics literature. The estimated residential elasticity values relate to the location-choice environment in a local labor market, allowing the estimates to reflect the role of income taxes for a given job choice. The propensity to select a location based on residential amenities rather than job characteristics has increased over time (Diamond, 2016; Diamond & Gaubert, 2022); this trend has intensified in the aftermath of the COVID19 pandemic and will likely persist in the future (Barrero et al., 2021). The decoupling of residential locations and jobs is pronounced for households with higher incomes, which is the basis of our estimation. Our estimate is useful in informing the debate and modeling of location choices in such environments. For instance, the equilibrium condition in the work-from-home economy in Brueckner et al. (2023) aligns with the residential arbitrage condition that applies to US households in our analysis. The equilibrium condition in a work-from-home economy is $A_s + H(N_s) = A_d + H(N_d)$, where A_i and $H(N_i)$ denote the supply of amenities and the net-housing utility (as a function of the local population) in region $i = d, s$.[75] In the work-from-home economy, wages are the same across residential locations and drop out of the equilibrium condition, whereas in our setting net wages of US households are identical across locations in the local labor market around Zurich since US households are subject to the US worldwide income tax system. As such, the sorting behavior of US households follows those incentives that govern residential sorting in a work-from-home economy.

A natural question is whether the elasticity estimate has relevance in a work-from-home setting with larger geographical distances. When extending the residen-

[75]See equation (3) in Brueckner et al. (2023).

tial choice model in Section 6.1 to include migration costs, modelled as an iceberg mobility cost in terms of utility, and scaling up geographical distances between jurisdictions, the residential elasticity will not change.[76] Intuitively, a new migrant in the enlarged geographical area still evaluates relative utility across locations. The re-scaled geographical distance does not change this relative evaluation. Thus, the estimated elasticity value continues to inform the model and is helpful in quantifying how income taxes influence residential choices in a work-from-home economy over larger distances.[77] A similar conclusion applies to residential choices in a conventional, non work-from-home economy, provided distances do not become too large to alter job choices.[78]

In addition, the estimates form the empirical basis for quantifying the effect of different income tax rules (sourcing rules) which exist in various countries and where income is taxed where it is earned. These rules will likely gain in importance in a work-from-home economy. The analysis of the effects of these tax systems is still in its infancy.[79] Part of the reason is that such an analysis requires estimates for

[76]Formally, the utility of household i from residing in jurisdiction c in year t gets discounted by the iceberg factor $\omega\kappa_{c,t,i}$, where $\kappa_{c,t,i} \in (1, \infty)$ and $\omega > 0$ is a scaling parameter, reflecting geographical distance. The log utility function in equation (3) now reads

$$U_{c,t,i} = \alpha \, log(1 - \tau_{c,t,i}) - log(\omega\kappa_{c,t,i}) + \gamma_c \boldsymbol{x}_{t,i} + \lambda \boldsymbol{x}_{c,t} + \eta_c + v_{c,t,i}. \tag{7}$$

Given the distributional assumption related to $v_{c,t,i}$, the probability of a new household i to locate in jurisdiction c in year t is

$$P_{c,t,i} = exp(\Phi_{c,t,i}) / \sum_{c' \in \mathcal{C}} exp(\Phi_{c',t,i}) \tag{8}$$

with $\Phi_{c,t,i} = \alpha \, log(1 - \tau_{c,t,i}) - log(\omega\kappa_{c,t,i}) + \gamma_c \boldsymbol{x}_{t,i} + \lambda \boldsymbol{x}_{c,t} + \eta_c$. Differentiating equation (8) with respect to ω yields $\partial P_{c,t,i}/\partial\omega = 0$. The probability $P_{c,t,i}$ does not change with the scaling parameter, ω. As a consequence, the residential elasticity in equation (6) is also insensitive to changes in the scaling parameter.

[77]For household i who already resides in one of the locations, the probability to locate in jurisdiction c, in which the household already resides, follows from setting $\omega \equiv 0$. For all other jurisdictions the location probability continues to be given by equation (8). In this case, the residential elasticity in equation (6) will change with ω. The effect on the residential elasticity will be small provided the probability to stay in the jurisdiction is sufficiently small, which tends to be the case when the number of jurisdictions becomes large, for instance. The estimated elasticity approximately describes residential choice behavior.

[78]In such environments, commuting costs matter. They can be similarly modelled as migration costs or as a factor that discounts wages (Monte et al., 2018). In either case, for a given job choice, scaling up commuting distances leaves the residential elasticity in equation (6) unchanged.

[79]See Agrawal & Tester (2023) and Agrawal & Hoyt (2018), for instance.

residential choices that are tax effects on location behavior for a given job choice, or a residential elasticity that is not related to income taxes but influenced by non-tax factors only (Agrawal & Brueckner, 2022). The residential elasticity estimate for UK households and the estimate for US households correspond to these estimates. Thus, this study opens up possibilities to quantify urban models and to analyze the consequences of these tax systems in an increasingly relevant work-from-home economy.

8 Conclusion

In this paper, we provide novel evidence of the effect of income taxes on the residential location choices of high-income households. To isolate the tax effect on sorting from non-tax effects (such as preferences for amenities), we make use of the institutional feature that US households are subject to a world-wide income tax system, which effectively insulates US households with a sufficiently high taxable income from Swiss income tax. Their moving decisions are thus unaffected by Swiss income taxes. In contrast, UK households are subject to Swiss income tax. Comparing the sorting behavior of US households with those of UK households with similar location preferences (w.r.t. locations in Switzerland) in the one-hour commuting zone around Zurich provides the opportunity to isolate the effect of income taxes on sorting behavior for a given job choice and to provide an estimate of the residential elasticity.

Using different empirical models, we consistently estimate a residential elasticity of around eight. US households sort into high-tax municipalities to benefit from the improved supply of amenities without being subject to local income tax. The elasticity estimate provides the empirical basis for the evaluation of the efficiency of local tax policy as well as of tax issues that gain prominence in an increasingly relevant work-from-home economy.

References

Agrawal, D. R., & Brueckner, J. (2022). Taxes and telework: The impacts of state income taxes in a work-from-home economy. CESifo Working Paper No. 9975.

Agrawal, D. R., & Foremny, D. (2019). Relocation of the rich: Migration in response to top tax rate changes from Spanish reforms. *Review of Economics and Statistics*, *101*(2), 214–232.

Agrawal, D. R., & Hoyt, W. H. (2018). Commuting and taxes: Theory, empirics and welfare implications. *Economic Journal*, *128*(616), 2969-3007.

Agrawal, D. R., & Tester, K. (2023). The effect of taxes on where superstars work. *American Economic Journal: Economic Policy*, *Forthcoming*.

Akcigit, U., Baslandze, S., & Stantcheva, S. (2016). Taxation and the international mobility of inventors. *American Economic Review*, *106*(10), 2930–81.

Baldwin, R., & Krugman, P. (2004). Agglomeration, integration and tax harmonization. *European Economic Review*, *48*(1), 1-23.

Barrero, J. M., Bloom, N., & Davis, S. J. (2021). Why working from home will stick. NBER Working Paper 28731.

Barrero, J. M., Bloom, N., & Davis, S. J. (2023). The evolution of work from home. *Journal of Economic Perspectives*, *37*(4), 23-50.

Basten, C., von Ehrlich, M., & Lassmann, A. (2017). Income taxes, sorting and the costs of housing: Evidence from municipal boundaries in Switzerland. *The Economic Journal*, *127*(601), 653–687.

Borusyak, K., Dix-Carneiro, R., & Kovak, B. K. (2022). Understanding migration responses to local shocks. mimeo, University College London.

Brueckner, J. K., Kahn, M. E., & Lin, G. C. (2023). A new spatial hedonic equilibrium in the emerging work-from-home economy? *American Economic Journal: Applied Economics*, *15*(2), 285-319.

Brülhart, M., Bucovetsky, S., & Schmidheiny, K. (2015). Taxes in cities. In *Handbook of Regional and Urban Economics* (Vol. 5, pp. 1123–1196). Elsevier.

Brülhart, M., Danton, J., Parchet, R., & Schläpfer, J. (2022). Who bears the burden of local taxes? CEPR Discussion Paper 15826, London.

Brülhart, M., Gruber, J., Krapf, M., & Schmidheiny, K. (2022). Behavioral responses to wealth taxes: Evidence from Switzerland. *American Economic Journal: Economic Policy*, *14*(4), 111-150.

Coomes, P. A., & Hoyt, W. H. (2008). Income taxes and the destination of movers to multistate MSA's. *Journal of Urban Economics*, *63*, 920-937.

Correia, S., Guimarães, P., & Zylkin, T. (2020). Fast poisson estimation with high-dimensional fixed effects. *The Stata Journal*, *20*(1), 95–115.

Diamond, R. (2016). The determinants and welfare implications of US workers' diverging location choices by skill: 1980-2000. *American Economic Review*, *106*(3), 479-524.

Diamond, R., & Gaubert, C. (2022). Spatial sorting and inequality. *Annual Review of Economics*, *Forthcoming*.

Epple, D., & Romer, T. (1991). Mobility and redistribution. *Journal of Political economy*, *99*(4), 828–858.

Fajgelbaum, P. D., & Gaubert, C. (2020). Optimal spatial policies, geography and sorting. *Quarterly Journal of Economics*, *135*(2), 959-1036.

Feldstein, M., & Wrobel, M. V. (1998). Can state taxes redistribute income? *Journal of Public Economics*, *68*(3), 369–396.

Internal Revenue Service. (2020a). *Publication 514: Foreign tax credit for individuals.*

Internal Revenue Service. (2020b). *Publication 54: Tax guide for u.s. citizens and resident aliens abroad.*

Kleven, H. J., Landais, C., & Saez, E. (2013). Taxation and international migration of superstars: Evidence from the European football market. *American Economic Review*, *103*(5), 1892–1924.

Kleven, H. J., Landais, C., Saez, E., & Schultz, E. (2014). Migration and wage effects of taxing top earners: Evidence from the foreigners' tax scheme in Denmark. *The Quarterly Journal of Economics*, *129*(1), 333–378.

Martinez, I. (2022). Mobility responses to the establishment of a residential tax haven: Evidence from Switzerland. *Journal of Urban Economics*, *129*, 1034–41.

Martinez, I., Saez, E., & Siegenthaler, M. (2021). Intertemporal labor supply substitution? Evidence from the Swiss income tax holidays. *American Economic Review*, *111*(2), 506–46.

Monte, F., Redding, S. J., & Rossi-Hansberg, E. (2018). Commuting, migration, and local employment elasticities. *American Economic Review*, *108*(12), 3855-3890.

Moretti, E., & Wilson, D. J. (2017). The effect of state taxes on the geographical location of top earners: Evidence from star scientists. *American Economic Review*, *107*(7), 1858–1903.

OECD. (2021). *Fiscal federalism 2022: Making decentralisation work*. OECD Publishing, Paris.

Organ, P. R. (2021). Citizenship and taxes: Evaluating the effects of the US tax system on individuals' citizenship decisions. *mimeo,* University of Michigan.

Redding, S. J. (2023). Quantitative urban models: From theory to data. *Journal of Economic Perspectives*, *37*(2), 75-98.

Roller, M., & Schmidheiny, K. (2016). Effective tax rates and effective progressivity in a fiscally decentralized country. *mimeo*.

Schmidheiny, K. (2006). Income segregation and local progressive taxation: Empirical evidence from Switzerland. *Journal of Public Economics*, *90*(3), 429–458.

Schmidheiny, K. (2017). Emerging lessons from half a century of fiscal federalism in Switzerland. *Swiss Journal of Economics and Statistics*, *153*(2), 73–101.

Schmidheiny, K., & Slotwinski, M. (2018). Tax-induced mobility: Evidence from a foreigners' tax scheme in Switzerland. *Journal of Public Economics*, *167*, 293–324.

Silva, J. S., & Tenreyro, S. (2006). The log of gravity. *The Review of Economics and Statistics*, *88*(4), 641–658.

Weber, S., & Péclat, M. (2017). Georoute: Stata module to calculate travel distance and travel time between two addresses or two geographical points. *mimeo,* Boston College.

Young, C., Varner, C., Lurie, I. Z., & Prisinzano, R. (2016). Millionaire migration and taxation of the elite: Evidence from administrative data. *American Sociological Review*, *81*(3), 421–446.

Online Appendix

Contents

A Additional Information on the Taxation of US Households in Switzerland

In what follows, we calculate the income threshold at which US households are fully isolated from the income tax system in the one-hour commuting area around Zurich for a single filer (without children) and a married couple filing jointly (2 children).

As described in the main paper, the US provides two important instruments to protect its citizens living abroad from double taxation. The use of one or both instruments is optional for tax payers. First, US citizens living abroad may choose to exclude a limited amount of foreign-earned income (up to USD 107,600 in the 2020 tax year)[80] from their taxable income; this is known as 'foreign-earned income exclusion'. Second, they may credit the tax liability owed to their country of residence against the US federal tax liability on the non-excluded income, referred to as 'foreign tax credit'.

Since the use of both instruments is optional for US households, we first calculate the optimal strategy dependent on household characteristics, and location choice.

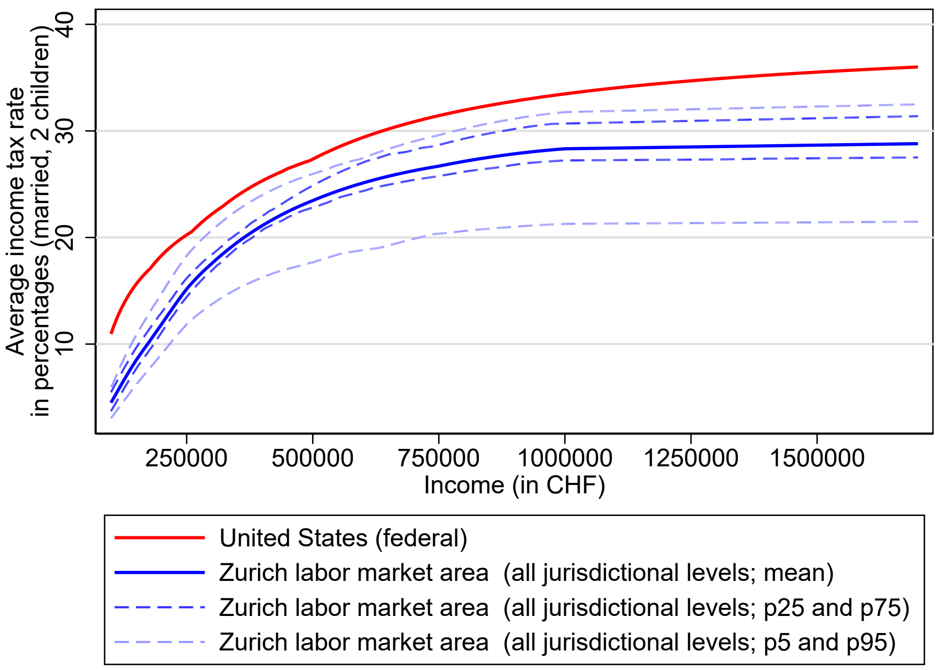
If US households opt for the 'foreign-earned income exclusion', then they can only credit foreign income taxes against their US tax liability on their non-excluded foreign earnings. On this non-excluded part, they can only deduct the average tax on their *total* income - rather than the average tax on the non-excluded part. As a consequence, households should only opt for the foreign earned income exclusion, if the average US tax rate on the excluded amount (that is 107,600 or twice the amount for a married couples) is higher than the average tax on their total income in Switzerland. The tax liability of these households depends on both, Swiss and US average tax rates.

By contrast, high-income US households for which the average tax on their total income in Switzerland exceeds the average US federal tax rate on the excludable

[80] Together, married couples can generally exclude the double amount, that is up to USD 215,200.

amount should not opt for the 'foreign-earned income exclusion'. For this households, we can express the relevant average tax rate for high-income US household i, residing in municipality n, simply as $max[\tau_{i,n}, \tau_{i,US}]$. These households are hence effectively isolated from the Swiss income tax system if the US tax rate is higher than the one in Switzerland (Example 1 illustrates this case based on actual values for 2020). Figure 3 and Figure A3 show that this is indeed the case for the 676 municipalities that are located within the one-hour commuting area around the city of Zurich.

Figure A1: Comparison of Average Income Tax Rates - Zurich Local Labor Market



Notes: This figure compares the federal income tax rate in the US with the aggregated income tax rate in Switzerland (federal, cantonal, and municipal) in 2017 for a married couple with 2 children. The federal income tax rate is calculated using the tax simulator (version 32) from NBER. Income tax rates in Switzerland (aggregated over federal cantonal and municipal level) are taken from the Swiss Federal Tax Administration. USD are translated into CHF using the yearly average currency exchange rate for 2017 used by the Internal Revenue Service: 1.024 CHF per USD.

We calculate the income threshold above which the lowest average tax rate on the total income within the municipalities of the one-hour commuting area around Zurich exceeds the US average tax rate on the excludable amount. This income threshold amounts to CHF 207,100 (USD 194,600)[81] for a single households without children and CHF 296'000 (USD 278,100) for a married couple with 2 children. Within the one-hour commuting area around Zurich, the tax liability of these households is hence independent of tax rates in Switzerland.

Example 1: Taxation of a US citizen residing in Switzerland

- Ms. Smith (US citizen, unmarried, no children) lives in Horgen (Switzerland) and has a gross income of CHF 400,000 (USD 425,985).
- The average income tax rate on a gross income of CHF 400,000 is 25.6% in Horgen (Switzerland) and 28.6% in the US. The **foreign earned income exclusion** amounts to CHF 101,036 (USD 107,600). The US average tax rate on this amount is 18.5%.
- Ms. Smith pays **CHF 102,400** (400,000 × 25.6%) to Swiss tax authorities.
- When only opting for the **foreign tax credit** she additionally pays **CHF 12,000** [400,000 × (28.6% - 25.6%)] to the IRS.
- When additionally opting for the **foreign earned income exclusion** the amount owed to the IRS after excluding the first CHF 101,036 of her income (but before applying the tax credit) reduces to **CHF 95,708** (400,000 × 28.6% - 101,036 × 18.5%).
- However, now she can only deduct 74.7% (298,964/400,000) of her Swiss tax liability. Therefore, she pays **CHF 19,216** (95,708 - 400,000 × 25.6% × 74.7%) to the IRS. This amount is larger than the **CHF 12,000** that she would owe to the IRS when not opting for the foreign earned income exclusion.

[81] Yearly average exchange rates used by the IRS: 0.939 (in 2020) CHF per USD; values are available here: https://www.irs.gov/individuals/international-taxpayers/yearly-average-currency-exchange-rates.

B Additional Information on Data

B.1 Information on Jurisdiction-specific Variables

Commuting Time to Zurich We calculate the commuting time on the road network between the center coordinates of each municipality to the center coordinate of the City of Zurich under average traffic conditions using the georoute command from Weber & Péclat (2017). Center coordinates are defined by the Swiss Federal Office of Topography and generally placed on the main road or main square of a municipality. The center coordinates of the City of Zurich are next to the main station of the public railway network. We also experiment with centroid coordinates. However, center coordinates appear superior since centroid coordinates are often located away from main roads. For the City of Zurich, we compute the average commuting time needed for a random sample of 1000 households from their exact location of residence to the Zurich main station.

Access to major lake Access to a major lake is a dummy variable equal to one if a municipality borders a lake with a surface of at least $5km^2$. The following lakes within the Zurich commuting zone fulfill this definition: Lake Constance, Lake Lucerne, Lake Zurich, Untersee, Lake Zug, Lake Walen, Sihlsee, Lake Hallwil, Lake Greifen, Lake Sarnen, Aegerisee and Lake Baldegg.

Income tax rates The individual census data contain a household identification variable, which allows us to construct a data-set at the level of the taxable entity. For unmarried individuals, this taxable entity corresponds to the individual. For married households, the taxable entity consists of the two partners. Based on the available characteristics (wage, number of children, marital status, income splits), we then match these observations at the level of the taxable entity with average income tax rates in Swiss jurisdictions (federal, cantonal, and municipal) obtained from the

Swiss Federal Tax Administration: https://swisstaxcalculator.estv.admin.ch.

Wealth tax rates We obtain the average wealth tax rate of Swiss jurisdictions (federal, cantonal and municipal) from the Swiss Federal Tax Administration and use the average wealth tax rate that applies to a net wealth of 100 million as a very close approximation of the top marginal wealth tax rate.

Municipal, Cantonal and Federal Expenditure Data We obtain annual federal and cantonal expenditure data for years 2012-2020 from the Swiss Federal Finance Department. By contrast, the accounting data of the universe of Swiss municipalities are not compiled at the federal level. This forces us to collect municipal expenditure data from cantonal offices. We succeeded in obtaining annual municipal expenditure data between 2012 and 2020 for the universe of 676 municipalities that are located in a commuting time of one hour around Zurich.

Table A1: Summary statistics on the municipalities of the Zurich commuting zone (2012-2020, pooled)

	mean	sd	min	median	max	N
Time to Zurich (in min.)	40.2	11.9	8.3	41.1	59.9	6181
Access to major lake	0.147	0.354	0	0	1	6181
Public expenditure p.c. (in CHF)	22024	2903	16634	21578	40945	6181
Top marginal income tax rate	31.32	3.83	16.91	31.66	39.07	6181
Top marginal wealth tax rate	0.44	0.15	0.10	0.46	0.77	6181
Population	5232	16785	86	2619	421878	6181
# of Top 1% US households	0.050	0.537	0	0	15	6181
# of Top 1–5% US households	0.169	1.732	0	0	60	6181
# of Top 5–10% US households	0.204	2.44	0	0	91	6181
# of Top 10% US households	0.423	4.616	0	0	158	6181
# of Top 1% UK households	0.111	0.858	0	0	21	6181
# of Top 1–5% UK households	0.456	2.497	0	0	51	6181
# of Top 5–10% UK households	0.509	2.872	0	0	67	6181
# of Top 10% UK households	1.076	5.989	0	0	123	6181

Notes: This table presents descriptive statistics of the municipalities in the Zurich commuting zone for the time period between 2012 and 2020. Population refers to the total Swiss and foreign population. Top marginal income tax rates and top marginal wealth tax rates are in percent and aggregated over the municipal, cantonal, and federal level. Total public expenditure per capita denotes the sum of municipal, cantonal, and federal spending divided by the population of the respective jurisdiction.

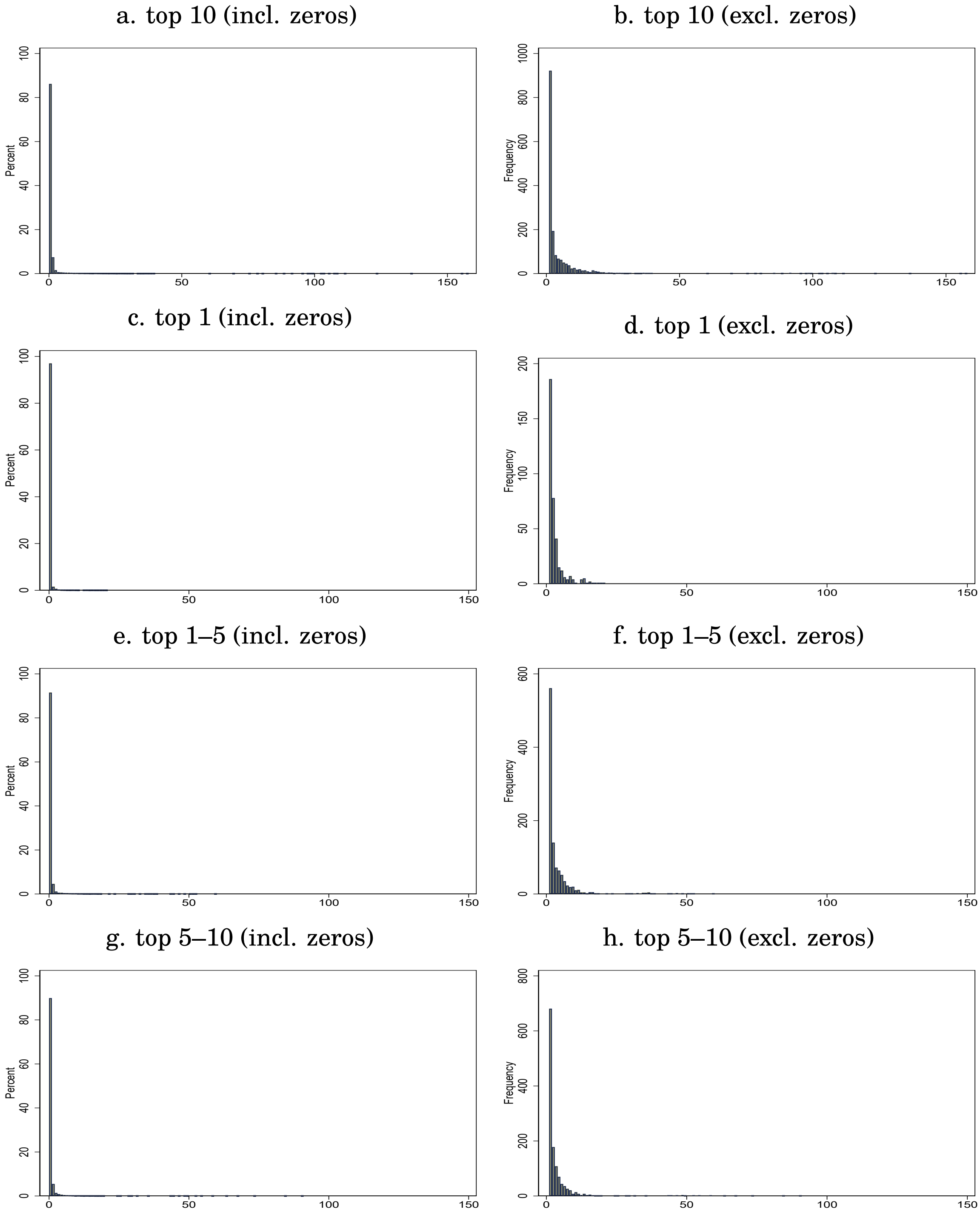
Table A2: Summary statistics on the choice cantons of the household-level analysis (2012-2020, pooled)

	AG	LU	SZ	ZG	ZH
Time to Zurich (in min.)	38.9	49.9	39.3	31.1	27.8
Access to major lake	0.028	0.298	0.400	0.636	0.137
Public expenditure p.c. (in CHF)	20093	22019	21498	25733	25278
Top marginal income tax rate	32.00	28.78	22.74	20.85	35.23
Top marginal wealth tax rate	0.46	0.28	0.19	0.29	0.61
Population (in thousands)	653	398	154	122	1438
# of Top 1% US households	0.4	1.4	1.2	5.1	38.0
# of Top 1–5% US households	3.2	4.6	6.4	16.7	122.7
# of Top 5–10% US households	3.7	7.3	4.3	18.3	145.1
# of Top 10% US households	7.3	13.3	12.0	40.1	305.8
# of Top 1% UK households	0.8	2.1	20.1	46.0	40.8
# of Top 1–5% UK households	11.7	9.8	38.8	121.3	223.6
# of Top 5–10% UK households	29.8	11.8	31.0	115.6	280.9
# of Top 10% UK households	42.2	23.7	89.9	282.9	545.2

Notes: This table presents descriptive statistics of the choice cantons of the household-level analysis: AG, LU, SZ, ZG and ZH are the abbreviations for the cantons of Aargau, Lucerne, Schwyz, Zug, and Zurich. All values are averaged over the time period between 2012 and 2020. Population refers to the total Swiss and foreign population. Top marginal income tax rates and top marginal wealth tax rates are in percent and aggregated over all jurisdictional levels. Total public expenditure per capita denotes the sum of municipal, cantonal, and federal spending divided by the population of the respective canton.

B.2 Municipal-level Analysis: Outcome Variable

Figure A2: Number of households by municipality, year and origin



Notes: This figure shows discrete histograms of our outcome variables used in Section 5 for the specified top income groups. The outcome variable, $N_{o,n,t}$, denotes the number of households with citizenship $o \in \{US, UK\}$ that reside in municipality n and in year t. Therefore, the histogram is based on the count of US and UK households in a given municipality and year. Observations on the number of UK and US households are then pooled together for constructing the histograms. The panels on the left show the whole distribution for the respective top income groups with the heights of bars scaled in percent. The panels on the right show the distribution without zero values (in frequencies, so that heights are equal to the number of observations in each discrete bin). Total number of observations (including zeros): 12,362. Pooled data for the years 2012-2020.

C Tax-induced Sorting Effect on Effective ATR

In this section of the Appendix, we aim to quantify the effect of income tax-induced sorting on the effective average income tax rates – as displayed in Figure 5 – by means of regression analysis.

Specification and Estimation In our most comprehensive specification, we quantify the effect of tax-induced sorting on effective average income tax rates based on the following equation:

$$ln(\tau_{i,n,t}) = \alpha UK_{i,t} + \boldsymbol{X}_{i,t}\lambda + \mu_t + \epsilon_{i,n,t}. \tag{9}$$

$ln(\tau_{i,n,t})$ is the log average tax rate applicable to household i in municipality n and year t. $UK_{i,t}$ is a dummy variable equal to one for UK nationals and equal to zero for US nationals. μ_t is a year-fixed effect. $\boldsymbol{X}_{i,t}$ contains the set of control variables summarized in Table 1. This includes the principal-earner specific age, gender and years of residency in Switzerland, and the number of children in the household. Finally, $\epsilon_{i,n,t}$ is the disturbance term. In addition, we run specifications without any control variables and year-fixed effect, or with only one of the two. In all estimations reported, standard errors are clustered at the municipality level.

Results The estimates presented in Table A3 are consistent across the different specifications and confirm the findings obtained by the semi-parametric approach. We find large differences in effectively paid average tax rates depending on household nationality. For the top wage percentile with annual household wages above CHF 1.2 millions, we find that being a UK rather than US household results in a reduction of effectively paid average tax rate by around 23%. This finding is statistically highly significant and economically large. For a household with an average tax rate of 29% and an annual wage income of CHF 2.4 million - the mean values for this income group - the obtained coefficient values are equivalent to an annual income tax reduction of around CHF 160,000 (2,4 million $\times$ 0.29 $\times$ 0.23).

Table A3: Effect on Effective Average Tax Rate (2012-2020)

	(1)	(2)	(3)	(4)
Panel A: Top 1%				
Tax-induced sorting effect on ATR (UK=1)	-0.229***	-0.228***	-0.228***	-0.227***
	(0.032)	(0.029)	(0.031)	(0.028)
Panel B: Top 1–5%				
Tax-induced sorting effect on ATR (UK=1)	-0.113***	-0.109***	-0.113***	-0.109***
	(0.024)	(0.016)	(0.024)	(0.016)
Panel C: Top 5–10%				
Tax-induced sorting effect on ATR (UK=1)	-0.099***	-0.092***	-0.098***	-0.091***
	(0.021)	(0.017)	(0.020)	(0.017)
Control variables		✓		✓
Year FE			✓	✓

Notes: The table presents estimates based on equation $ln(\tau_{i,n,t}) = \alpha UK_{i,t} + \boldsymbol{X}_{i,t}\lambda + \mu_t + \epsilon_{i,n,t}$. In all regressions, the dependent variable is the logarithm of the average income tax rate (aggregated over the municipal, cantonal, and federal level). Standard errors clustered at the municipality level are provided in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Total number of observations in Panel A: 999, Panel B: 3862, and Panel C: 4407.

As hypothesized, the estimated effects decrease when going down in the wage distribution. For households that belong to the top 1–5% (top 5–10%) of the income distribution, we find a tax induced sorting effect on the effectively paid average tax rate of around 11% (9–10%).

D Alternative Definition of Zurich Commuting Zone

Table A4: Municipal-Level Elasticities: Baseline Effects (2012-2020): 45 minutes Commuting Zone around Zurich

	(1)	(2)	(3)	(4)	(5)
$ln(1-\tau_{n,t})$	0.335	0.306	-5.880*	-5.193	
	(12.681)	(12.848)	(2.810)	(2.681)	
$UK_{n,t} \times ln(1-\tau_{n,t})$	17.548*	17.781*	11.380***	11.399***	11.432***
	(8.632)	(8.690)	(1.920)	(1.931)	(1.935)
Municipality FE			✓	✓	
Year FE		✓		✓	
Municipality×year FE					✓
Observations	7,478	7,478	7,478	7,478	7,478
Estimator	PPML	PPML	PPML	PPML	PPML

Notes: This table reports the baseline estimates from equation (1) using the PPML estimator proposed by Silva & Tenreyro (2006). All municipalities are in the 45 minutes (by car) commuting zone around the City of Zurich. The dependent variable is the number of high income households (top 10%) with principal earner nationality $o \in \{US, UK\}$ in municipality n and year t that belong to the top wage decile of US and UK households residing in the one-hour-commuting zone around Zurich. The independent variables are a dummy variable, $UK_{n,t}$, equal to one for UK households and zero for US households, and the log retention rate in municipality n and year t, $ln(1-\tau_{n,t})$ (except specification 5 where the retention rate is omitted due to perfect collinearity with the municipality-by-year fixed effects); and the interaction term, $UK_{n,t} \times ln(1-\tau_{n,t})$. Standard errors clustered at the municipality level are provided in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A5: Overall Elasticities by Income Bracket (2012-2020): 45min Commuting Zone around Zurich

	top 1%	top 1–5%	top 5–10%
	(1)	(2)	(3)
$UK_{n,t} \times ln(1-\tau_{n,t})$	16.217***	11.086***	14.825***
	(1.846)	(1.906)	(3.180)
Municipality×year FE	✓	✓	✓
Observations	7,478	7,478	7,478
Estimator	PPML	PPML	PPML

Notes: This table reports estimates from equation (1) for different top wage percentile ranges using the PPML estimator proposed by Silva & Tenreyro (2006). All municipalities are in the 45 minutes (by car) commuting zone around the City of Zurich. The dependent variable is the number of high-income households with principal earner nationality $o \in \{US, UK\}$ in municipality n and year t that belong to the specified top wage percentile ranges. 'Top 1%' refers to households that belong to the 99th to 100th wage percentile while 'top 1–5%' refers to households that belong to the 95th to 98.99th wage percentile, etc. The independent variables are a dummy variable, $UK_{n,t}$, equal to one for UK households and zero for US households, and the interaction term between the UK dummy and the retention rate in municipality n and year t, $UK_{n,t} \times ln(1-\tau_{n,t})$. Standard errors clustered at the municipality level are provided in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A6: Within-Canton Elasticities by Income Bracket (2012-2020): 45 minutes Commuting Zone around Zurich

	top 10%	top 1%	top 1–5%	top 5–10%
	(1)	(2)	(3)	(4)
$UK_{n,t} \times ln(1-\tau_{n,t})$	13.923**	11.819***	13.828*	19.656**
	(4.591)	(3.303)	(5.438)	(6.575)
Municipality×year FE	✓	✓	✓	✓
Canton×origin×year FE	✓	✓	✓	✓
Observations	7,478	7,478	7,478	7,478
Estimator	PPML	PPML	PPML	PPML

Notes: This table reports within-canton estimates for different top wage percentile ranges using the PPML estimator proposed by Silva & Tenreyro (2006). All municipalities are in the 45 minutes (by car) commuting zone around the City of Zurich. The specification is equivalent to equation (1) but includes canton-by-origin-by-year fixed effects. The dependent variable is the number of high income households with principal earner nationality $o \in \{US, UK\}$ in municipality n and year t that belong to the specified top wage percentile ranges. 'Top 10%' refers to households that belong to the 90th to 100th wage percentile while 'top 1%' refers to households that belong to the 99th to 100th wage percentile, etc. The independent variables are a dummy variable, $UK_{n,t}$, equal to one for UK households and zero for US households, and the interaction term between the UK dummy and the retention rate in municipality n and year t, $UK_{n,t} \times ln(1-\tau_{n,t})$. Standard errors clustered at the municipality level are provided in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A7: MUNICIPAL-LEVEL ELASTICITIES: SENSITIVITY CHECKS (2012-2020): 45 MINUTES COMMUTING ZONE AROUND ZURICH

	(1)	(2)	(3)	(4)	(5)	(6)
$ln(1-\tau_{n,t})$	-1.225	6.696	2.516			
	(2.124)	(7.845)	(7.996)			
$UK_{n,t} \times ln(1-\tau_{n,t})$	9.774***	15.071***	16.031**	10.140***	10.672***	10.254**
	(1.665)	(3.492)	(5.406)	(1.085)	(1.012)	(3.713)
$ln(Time\ to\ Zurich_n)$			-2.059***			
			(0.565)			
$UK_{n,t} \times ln(Time\ to\ Zurich_n)$			0.551			0.345
			(0.360)			(0.334)
$Access\ to\ lake_n$			1.375***			
			(0.393)			
$UK_{n,t} \times Access\ to\ lake_n$			-0.437			-0.105
			(0.277)			(0.178)
$ln(Public\ spending_{n,t})$			5.497***			
			(1.515)			
$UK_{n,t} \times ln(Public\ spending_{n,t})$			-0.702			-0.980
			(1.104)			(0.679)
$ln(Wealth\ tax_{n,t})$			-0.742			
			(0.752)			
$UK_{n,t} \times ln(Wealth\ tax_{n,t})$			0.977			0.530
			(0.564)			(0.438)
$Homophily_{n,t}$			0.596			
			(0.993)			
$UK \times Homophily_{n,t}$			1.278			0.343
			(0.656)			(0.652)
Municipality controls			✓			✓
Only mover		✓			✓	
Municipality×year FE				✓	✓	✓
Estimator	OLS	PPML	PPML	OLS	PPML	PPML
Observations	1,440	7,478	7,478	1,440	7,478	7,478

Notes: This table reports a series of robustness checks considering all municipalities within a commuting area of 45 minutes around Zurich. The dependent variable is the number of high-income households with principal earner nationality $o \in \{US, UK\}$ in municipality n and year t that belong to the top wage decile of US and UK households residing in the one-hour-commuting zone around Zurich. The independent variables of all specifications include a dummy variable, $UK_{n,t}$, equal to one for UK households and zero for US households. In specifications 1 and 4, observations with a value of zero are dropped. Section B of the Appendix provides a detailed description of the municipality covariates used in specifications 3 and 6. Standard errors clustered at the municipality level are provided in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

E Location Choice Models: Additional Results

Table A8: Location Choice Models: Heterogeneous Results

	(1)	(2)	(3)	(4)	(5)
Coefficient estimate UK (treatment group)					
Log retention rate × top 1%	10.337*	10.335*	12.290**	10.277*	13.212**
	(4.239)	(4.239)	(4.655)	(4.252)	(4.619)
Log retention rate × top 1–5%	8.471***	8.468***	8.926***	9.041***	10.045***
	(2.910)	(2.909)	(3.037)	(3.224)	(3.466)
Log retention rate × top 5–10%	9.188**	9.191**	9.564**	11.253***	12.577***
	(2.112)	(2.297)	(2.275)	(2.337)	(2.782)
Coefficient estimate US (control group)					
Log retention rate × top 1%	-5.027	-5.030	-3.102	-5.425	-2.516
	(4.609)	(4.608)	(4.930)	(4.658)	(4.933)
Log retention rate × top 1–5%	-4.061	-4.063	-3.630	-3.698	-2.714
	(2.449)	(2.448)	(2.529)	(2.536)	(2.659)
Log retention rate × top 5–10%	-5.637	-5.634	-5.287	-4.188	-2.903
	(3.297)	(3.298)	(3.409)	(3.680)	(3.909)
Average elasticities (UK – US)					
Top 1%	8.942***	8.942***	9.012***	9.017***	9.068***
	(1.292)	(1.292)	(1.305)	(1.337)	(1.325)
Top 1–5%	7.294***	7.293***	7.351***	7.315***	7.358***
	(0.865)	(0.866)	(0.872)	(0.874)	(0.881)
Top 5–10%	8.628***	8.627***	8.696***	8.867***	8.926***
	(1.117)	(1.117)	(1.124)	(1.144)	(1.149)
Homophily control		✓			✓
Canton controls			✓		✓
Household controls				✓	✓
N	61,310	61,310	61,310	61,310	61,310

Notes: This table shows the coefficient estimates on the retention rate for the multinomial logit regressions presented in Table 7. Standard errors are clustered at the household level: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

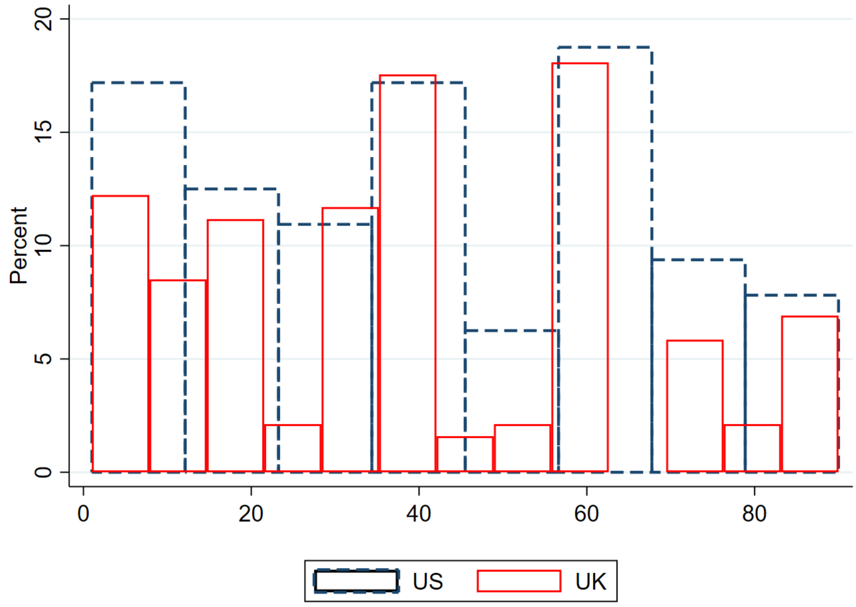
F Commuting behavior by nationality group

In this section, we aim at showing that there are not relevant differences in commuting behavior between US and UK individuals. This analysis cannot be performed on the same dataset that we use for all the estimates reported in the paper, as there only the place of residence and not the place of work is recorded. Therefore, we resort to the Swiss Household Panel (SHP), a survey that reports, among many other variables, also the number of minutes of daily commuting for each respondent, for all the years of the survey. In the following, we use SHP data for the years 1999-2020, in order to maximize our sample size, as well as the number of US and UK citizens in it. We keep individuals between 20 and 64 years old, i.e. those in working age. Our final sample consists of 164'665 individual-year observations, among those 588 are UK nationals and 156 are US nationals.

First, as far as the extensive margin of commuting is concerned, i.e. the choice between whether or not commuting, the test for the difference in means (both with equal and with unequal variances) concludes that we cannot reject the null hypothesis of an equal share of commuters among US and UK individuals (the corresponding p-value is 0.3792 with equal variances and 0.3848 with unequal variances in the two groups). Further, as far as the intensive margin of commuting is concerned, a test for the difference in means also provides evidence that we cannot reject the null hypothesis of same average daily commuting time between UK and US nationals. The corresponding p-value is 0.2691 in the test with equal variances and 0.1608 in the test with unequal variances across groups.

Finally, the similarity in the commuting behavior across the two groups can also be assessed by visual inspection of Figure A3.

Figure A3: Histogram of daily commuting time in minutes



Pooled SHP data 1999-2020, UK and US nationals only. Top 2% of commuting time has been trimmed to increase the readability of the graph. Only individuals with positive commuting time.

G Municipal-Level Elasticities: Placebo Checks

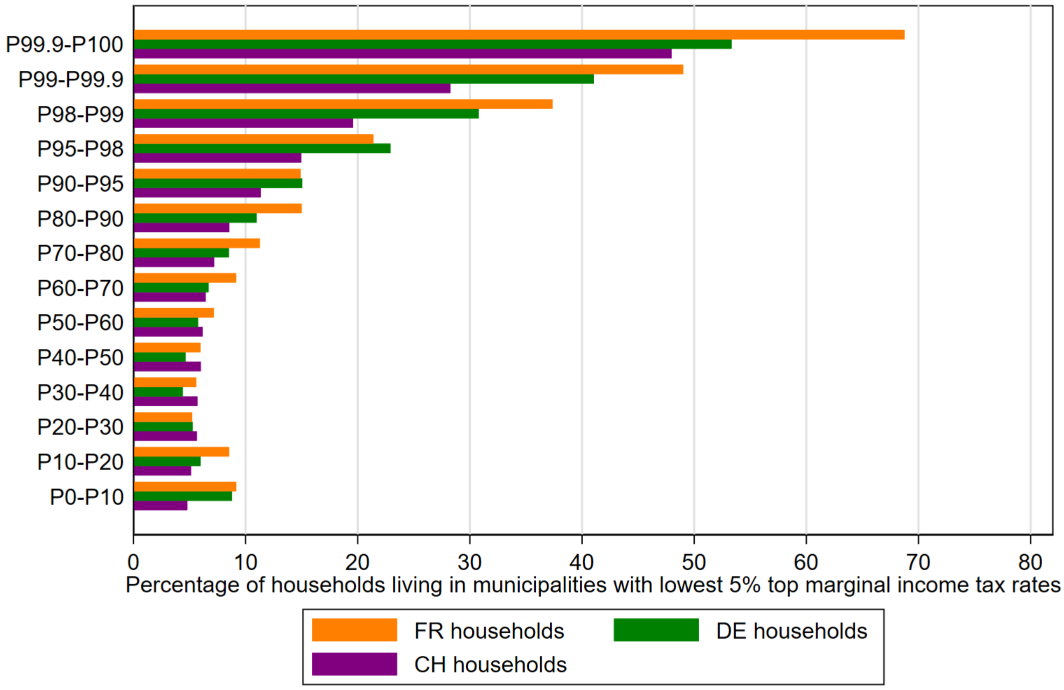
Table A9: Municipal-Level Elasticities: Placebo Checks (2012-2020)

	(1)	(2)
	Public spending	Wealth tax
N. of US/UK high-income households	0.000	-0.000
	(0.000)	(0.000)
Municipality controls	✓	✓
Estimator	OLS	OLS
Observations	12,362	12,362

Notes: This table reports a series of placebo checks based on the same data set as in Tables 2 and 3. The dependent variables are: (1) Public spending, and (2) Wealth tax. Section B of the Appendix provides a detailed description of each of these two outcome variables. The control variable used here is the number of high-income households with principal earner of nationality $o \in \{US, UK\}$ in municipality n and year t that belong to the top wage decile of US and UK households residing in the one-hour-commuting zone around Zurich. Standard errors clustered at the municipality level are provided in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

H Additional Evidence on Stylized Facts

Figure A4: Percentage of FR, DE and CH Households in Local Tax Havens (by wage percentile, 2012-2020)



Notes: This figure compares the shares (in percentage) of French (in orange), German (in green) and Swiss (in purple) households living in low-tax municipalities across wage percentiles. All municipalities with a travel time (by car) of one hour to the Zurich main railway station are considered. Low tax municipalities are defined as the 5% of municipalities with the lowest top marginal income tax rate. Percentage shares from 2012-2020 are averaged. $N_{FR}^{2020} = 8'519$, $N_{DE}^{2020} = 119'715$, $N_{CH}^{2020} = 1'099'772$. The percentiles have been computed on the basis of the wage distribution of UK and US households.