

APPENDIX

for “Taxing Top Wealth: Migration Responses and their Aggregate Economic Implications”

by Katrine JAKOBSEN, Henrik KLEVEN, Jonas KOLSRUD,
Camille LANDAIS and Mathilde MUNOZ

Contents

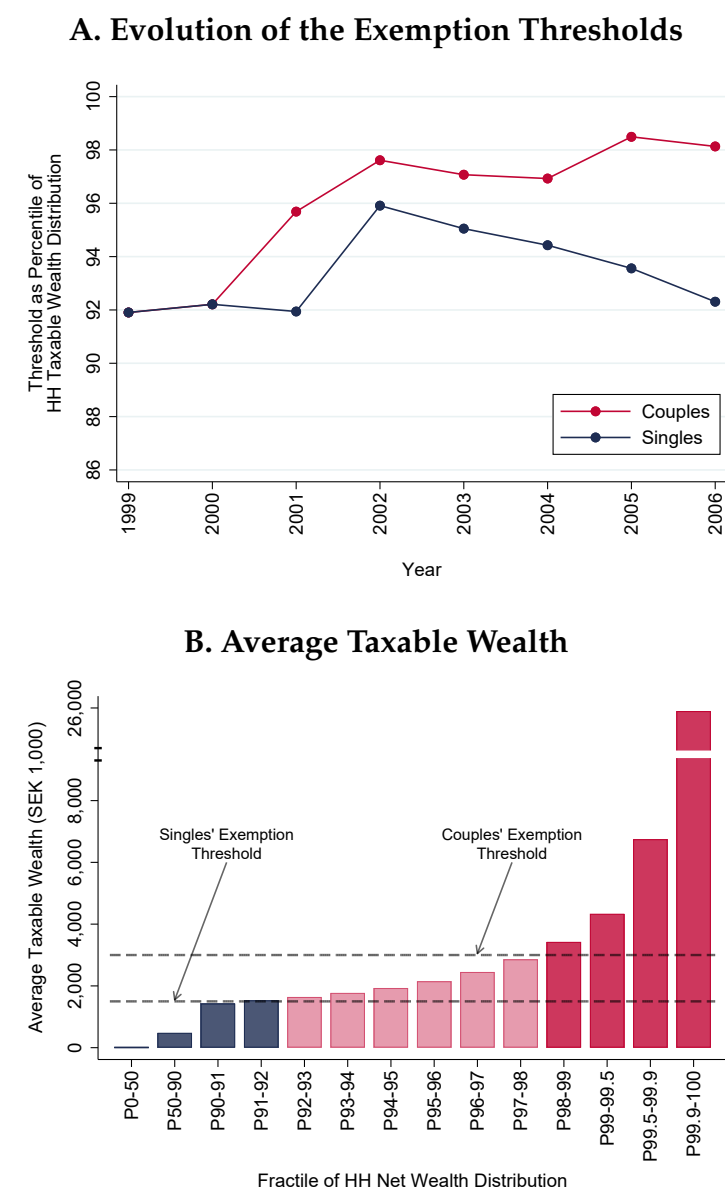
I	Additional Institutional Details	3
I.1	Additional Institutional Details: Sweden	3
I.2	Wealth Taxation in Denmark	6
II	Data	8
II.1	Prediction Model for Wealth in Sweden	8
II.2	Computation of Effective Tax Rates for Entrepreneurs	12
II.3	Wealth and Migration in Denmark	13
II.4	Identifying Ownership Links in Serrano Dataset	15
III	Additional Descriptive Evidence on Migration Patterns in Scandinavia	17
III.1	Duration analysis	17
III.2	Destination and origin countries	19
III.3	Selection: Who Migrates In and Out?	24
IV	Impact of Migration on Individual and Firm Outcomes: Additional Event Study Results	29
V	Migration Elasticity Estimates	44
V.1	Additional Evidence: Sweden	44
V.2	Out-of-Sample Validation: Migration Elasticity Estimates in Denmark . . .	50
VI	Quantification of Aggregate Implications	52
VI.1	A Simple Framework Accounting for the Dynamics of the Population of Wealthy Individuals	52
VI.2	Microfoundations of Migration Responses	53

VI.3 Impact of Wealth Tax Changes on the Steady State Population Stock	54
VI.4 Aggregate External Effects on Employment and Economic Activity	61
VI.4.1 Formula & Heterogeneous Treatment Effects of Migration	61
VI.4.2 Results & Robustness	69

I. Additional Institutional Details

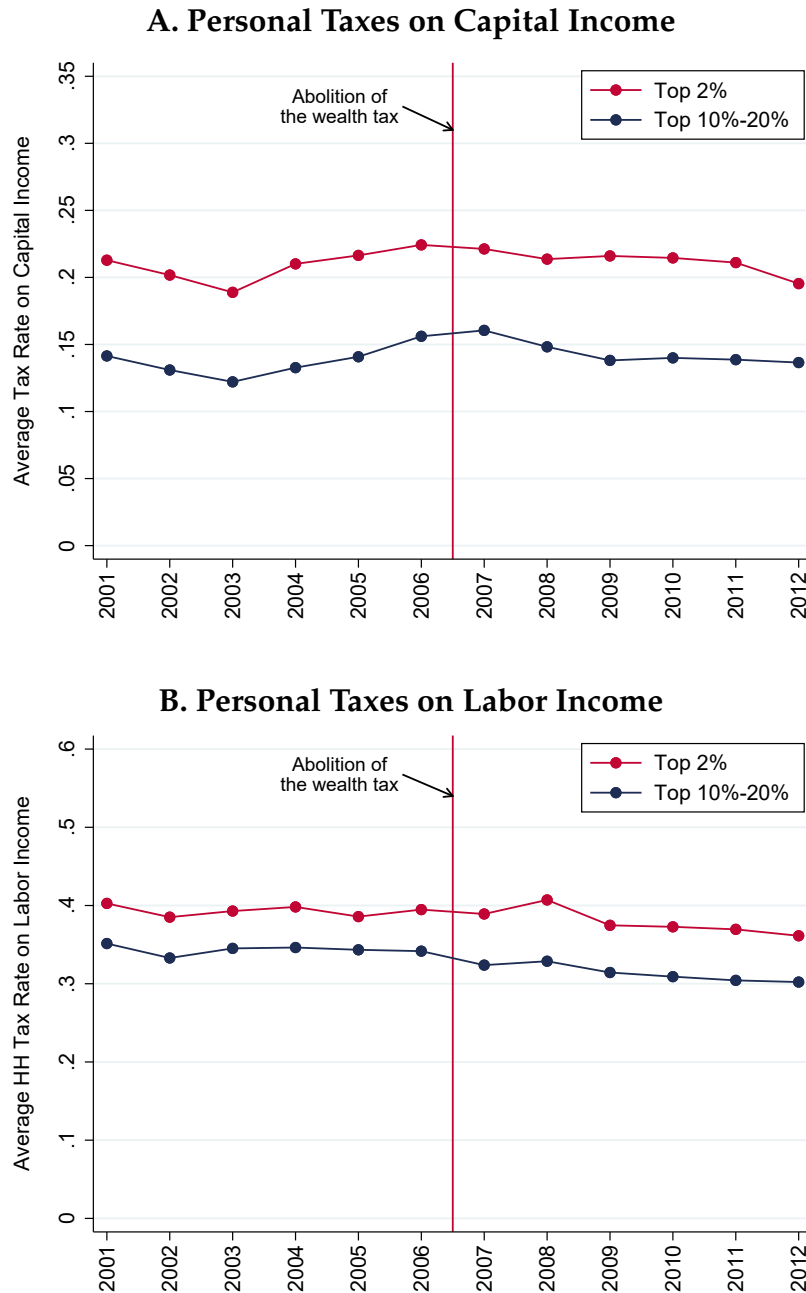
I.1. Additional Institutional Details: Sweden

Figure I.1: Exemption Threshold for the Wealth Tax in Sweden



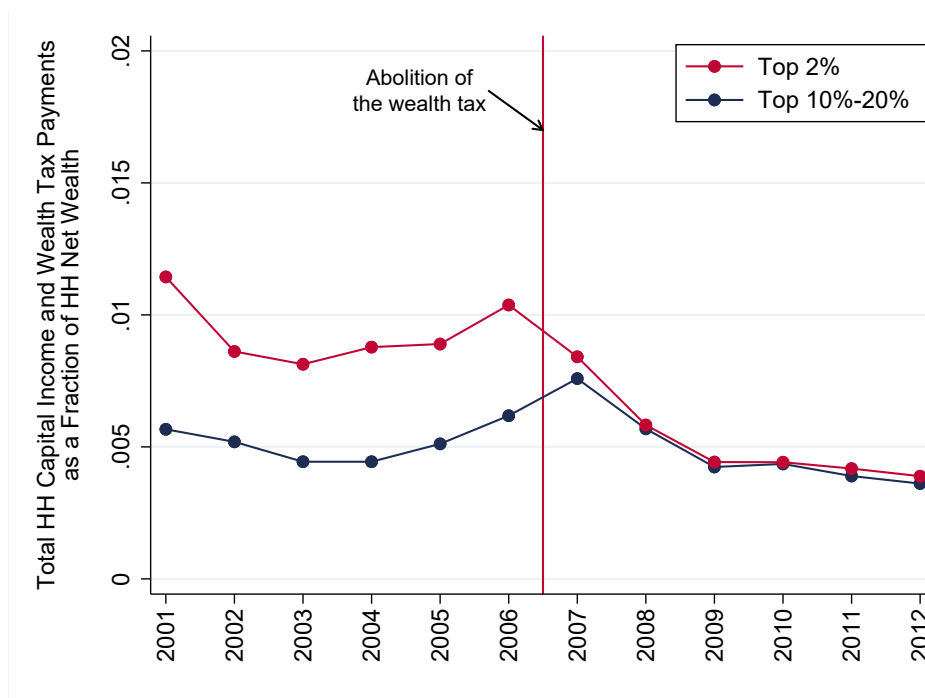
Notes: Panel A shows the evolution of the exemption threshold in the Swedish wealth tax as a function of the household taxable wealth distribution. The red line refers to the wealth tax exemption threshold set for couples, while the dark blue line refers to the wealth tax exemption threshold set for single individuals. Panel B shows the average household taxable wealth in 2006. We rank households based on their total household net wealth in 2006, and then plot the average taxable wealth by fractile of household net wealth. The dashed black lines refer to the wealth tax exemption thresholds for singles and couples that same year.

Figure I.2: Effective Tax Rates on Capital and Labor Income



Notes: This figure plots the average effective tax rate on capital income (Panel A) and on labor income (Panel B) for households in the top 2% of net wealth (wealthy households subject to the wealth tax, red line) and for households in the top 20-10% of net wealth (wealthy households not subject to the wealth tax, blue line). The tax rate on capital income (Panel A) is computed as the ratio of household capital income tax payments over household capital income each year. The top 1% of capital income tax payments among individuals in the top 20% of household net wealth was winsorized. The tax rate on labor income (Panel B) is computed as the ratio of household labor income tax payments over household taxable income. The vertical red line denotes the year where the wealth tax was repealed in Sweden.

Figure I.3: Effective Tax Rate on Capital Income and Wealth



Notes: This figure plots the average effective tax rate on the sum of capital income and wealth for households in the top 2% of net wealth (wealthy households subject to the wealth tax, red line) and for households in the top 20-10% of net wealth (wealthy households not subject to the wealth tax, blue line). We divide the sum of household tax payments on capital income and wealth by total household net wealth. The top 1% of capital income tax payments among individuals in the top 20% of household net wealth was winsorized. The vertical red line denotes the year where the wealth tax was repealed in Sweden.

I.2. Wealth Taxation in Denmark

An annual progressive wealth tax was implemented in Denmark from 1903 to 1997 and applied to Danish tax residents with net taxable wealth above an exemption threshold. This exemption threshold was always above the 98th percentile of the wealth distribution during the period we study.

Tax Base Taxable wealth was defined as the total net wealth of households, excluding pension wealth. Taxable wealth components included cash, deposits, bonds, equities, housing, large durables and business assets, net of any debts. Only taxable wealth above a given exemption threshold, whose evolution is shown in Panel B of Figure I.1, was included in the wealth tax base.

The definition of the geographic scope of the wealth tax base was the same in Denmark and Sweden. Danish residents were taxed on their worldwide assets, including financial and non-financial assets held abroad, while non-residents were only taxed on their assets held in Denmark.

From 1980 onwards, the value of certain assets was reduced when calculating taxable wealth. Specifically, the value of physical business assets was reduced by 20%, while the value of large forests was reduced by 30%.¹ The relief for physical business assets was increased to 25% in 1982 and 30% in 1983. From 1986 onwards, the relief for both business assets and forests was increased to 40%.

Finally, there was also a tax credit for the wealth tax through a tax ceiling mechanism capping the amount of wealth tax owed by taxpayers as a fraction of their taxable income. This mechanism varied for couples and single individuals and is described in [Jakobsen et al. \(2020\)](#).

Tax Rates Like in Sweden, the Danish wealth tax had a simple two-bracket structure. Wealth in Denmark was taxed at a flat rate of 2.2% above the exemption threshold until two reforms in the late 1980s and 1990s. After 1989, the tax rate was reduced to 1%, while in 1997 the wealth tax was repealed.

Reporting and Enforcement Most of the assets and liabilities were reported by third-parties to the Danish government. The value of bank deposits was reported by banks and the value of listed stocks and bonds was reported by financial institutions. The government used land and real estate registries to record non-financial assets. All other wealth

¹Since forests were also considered a physical business assets, their taxable value was reduced by 50%.

components had to be self-reported by taxpayers.

II. Data

II.1. Prediction Model for Wealth in Sweden

After the repeal of the wealth tax, reporting requirements changed, and we do no longer observe after 2007 the same comprehensive components of household wealth. Certain elements such as liquid bank accounts or listed stocks are missing, but we still observe many wealth components, such as real estate (through real estate registers) and closely-held business assets. To construct a consistent measure of wealth before and after 2007, we build a prediction model of household total net wealth that we train for the period pre-2007, and then use to predict net wealth after 2007. In this section, we provide all the details regarding our prediction model.

Overview of the Approach Our approach leverages two important features of the data. First, we continue to observe many components of household net wealth post-2007. Second, for the elements that we do not observe any more, we have precise information on the past value of assets, and on all income flows, which are both related in an accounting sense to the current value of the assets through iterating the law of motion of household wealth.

To understand the approach, we start by splitting assets between wealth W^o from asset classes o that are observable throughout the period, and wealth W^u from asset classes u that are only observed up to 2007.

$$\underbrace{W_t}_{\text{Net Wealth}} = \underbrace{W_t^o}_{\text{Assets observed pre and post 2007}} + \underbrace{W_t^u}_{\text{Assets not observed after 2007}} \quad (1)$$

We can then use the law of motion of wealth to break down W_t^u between capitalized past wealth W_{t-1}^u and net active savings/dissavings in assets of class u :

$$W_t = W_t^o + \underbrace{(1 + r_t^u) \cdot W_{t-1}^u}_{\text{Capitalized Past Wealth}} + \underbrace{p_t^u \cdot \Delta q_t^u}_{\text{Active Savings in Assets } u} \quad (2)$$

We then use the accounting identity of the household budget constraint, which imposes that the sum of net active savings in all classes of assets must be equal to the sum of all earnings and inheritances flows, net of all taxes and transfers received, minus household consumption.

$$p_t^u \cdot \Delta q_t^u + p_t^o \cdot \Delta q_t^o = \underbrace{E_t}_{\text{Earnings}} + \underbrace{I_t}_{\text{Inheritance}} + \underbrace{T_t}_{\text{Taxes \& Transfers}} - \underbrace{C_t}_{\text{Consumption}} \quad (3)$$

Combining (2) and (3), we get:

$$W_t = W_t^o + (1 + r_t^u) \cdot W_{t-1}^u + E_t + I_t + T_t - C_t - p_t^o \cdot \Delta q_t^o \quad (4)$$

Iterating the above identity, we get after X iterations:

$$W_t = W_t^o + W_{t-X}^u \prod_{j=t-X}^t (1 + r_j^u) + \sum_{k=t-X}^t (E_k + I_k + T_k - C_k - p_k^o \cdot \Delta q_k^o) \cdot \prod_{j=k+1}^t (1 + r_j^u) \quad (5)$$

Note that equation 5 is an accounting identity. Which means that if all the elements of the right-hand side of the identity are observed, the exact value of wealth can be computed. In practice, we observe, thanks to the rich administrative information available in Sweden, many of the elements from the right-hand side:

- Observable wealth W_t^o :
We observe continuously the following elements of wealth: real estate assets via the real estate register; closely-held business assets directly owned, through the K10 registry data, and indirectly owned through the Serrano database.
- Past wealth W_{t-X}^u :
We can observe all asset classes until 2007. After this date, the following financial assets are no longer observed: mutual funds, stocks in listed firms, bonds, bank holdings and balance in other liquid accounts. We use the fact that we can observe these assets up until 2007, and that we also observe average rate of returns r_j^u on these assets.
- Cumulated flow of past earnings $\sum_{k=t-X}^t E_k$:
We measure past disposable earnings recorded in the LISA dataset from 1990 to 2017. Importantly LISA accounts for all potential sources of labor income.
- Cumulated flow of past inheritances $\sum_{k=t-X}^t I_k$:
We observe all inheritance flows for the period 2001 to 2005 via the inheritance registry BELINDA.
- Cumulated flow of past taxes and transfers $\sum_{k=t-X}^t T_k$:

We observe all tax payments in the IoT tax registry, and all transfer payments received from LISA.

- Cumulated active savings in asset classes o $\sum_{k=t-X}^t p_k^o \cdot \Delta q_k^o$:
For these assets, the registers record both prices and quantities, and all related transactions, which means that we can compute active savings/dissavings in these asset classes.

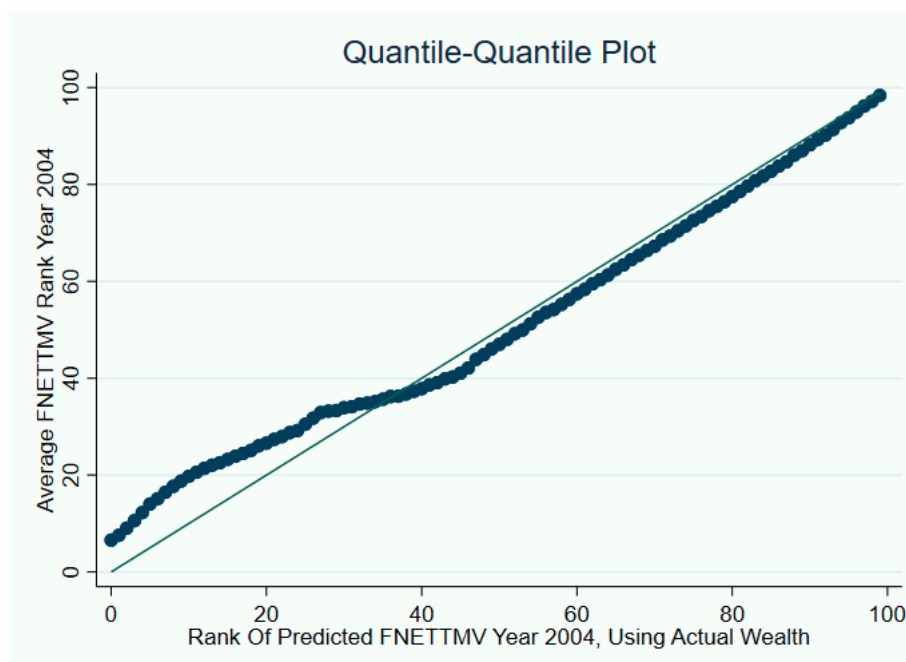
Implementation While we observe many of the elements of equation (5), we cannot simply implement this accounting identity to measure wealth post 2007 because: (i) we do not observe inheritances flows for the period pre-2001, and post-2005 and (ii) most importantly, we do not observe the cumulated flow of consumption $\sum_{k=t-X}^t C_k$. However, we have access to rich additional information that can serve as useful predictors of these unobserved inheritance and consumption flows.

- For inheritance flows: we observe parental wealth as well as the age of parents and the number and age of siblings, which allows us to predict the likelihood of receiving an inheritance using the observed inheritance flows over the period 2001-2005.
- For consumption flows: we observe some durable consumptions, such as cars (through the FORDON car register), and a rich set of demographics that correlate with consumption patterns (age, household structure, place of residence, etc).
- Furthermore, from equation (3), we know that consumption can be retrieved as residual from all flows of earnings, taxes/transfers and inheritances and active savings and dissavings. In practice we measure throughout the period all active savings and dissavings in asset classes o and we also measure active dissavings in asset classes u (e.g. dividends perceived, realized capital gains/losses, etc), but not active savings in asset classes u .

As a consequence, our methodology relies on building an ensemble classification model that exploits identity (5). For our model, we use all elements of the identity that we observe. And for the elements that we do not strictly observe in identity (5) (i.e. some inheritance flows and consumption flows), we complement the set of covariates of the model with all the various proxies and potential predictors described above: (i) parental wealth, age of parents and number and age of siblings, (ii) demographic characteristics of the household, (iii) observed durable consumption in cars, (iv) cumulated flows of active dissavings in asset classes u (e.g. dividends perceived, realized capital gains/losses, etc).

More specifically, we train random-forests on 10% of a random sample of the Swedish population and we then classify the population in groups of predicted net wealth. We choose to fix $X = 7$ for the model. This means that we can then use our model to predict wealth from 2000 to 2014.

Figure II.1: Predicting Wealth: Accuracy of our Model



Notes: This figure plots the average rank of taxable wealth in 2004 based on the rank of predicted taxable wealth the same year using our prediction model.

Validation and Model Accuracy Importantly, our prediction model performs exceptionally well, as shown in Appendix Figure II.1, and much better than capitalization methods often used to proxy wealth in the absence of proper administrative registers on wealth (Saez and Zucman, 2016).

Note that because the prediction model relies on past data, it is a good model to predict wealth for people about to move (out-migration effects) but does not work for individuals who just moved in (in-migration rates) for whom we do not have info on past wealth, parental wealth, past income, etc.

II.2. Computation of Effective Tax Rates for Entrepreneurs

We compute average effective tax rates on wealth for entrepreneurs. To do so, we need to obtain a valuation of private business assets. We value private business assets by using the balance-sheet measure of profits before taxes (EBIT) from our firm registry data. We apply a liquidity discount factor of 10%, meaning that we capitalize $0.9 \times \text{EBIT}$ for the wealthy who control unlisted firms. For smaller businesses with less than 50M in profits, we account for the human capital contribution of profits estimated in [Smith et al. \(2019\)](#): we only capitalize $0.9 \times 0.25 \times \text{EBIT}$. We then capitalize firms' profits. We use [Smith et al. \(2022\)](#) estimates for entrepreneurs' returns (10.5%) to infer the value of private business assets controlled by the rich in Sweden. The implicit assumption is that private business assets of the wealthy had the same returns in the U.S. and Sweden during the period we study. We thus obtain an estimate of private business assets wealth for each wealthy entrepreneur, each year. We then compute the implied effective tax rate on wealth, by dividing effective wealth tax payments by the total of net wealth and valued private business assets.

II.3. Wealth and Migration in Denmark

In Denmark, wealth and income registries are based on tax return data from the Danish Tax Agency (SKAT). Wealth levels are stated by the end of the year, and most of the components measuring wealth are reported by third-parties. For instance, cash value of real estate, deposits, bonds, and shares are third-reported. Until 1997, we also observe self-reported components that are subject to auditing by the tax authorities, such as the self-reported value of stocks (listed and non-listed), self-reported value of durables (cars, boats and caravans) and the self-reported value of share certificates for housing cooperatives, premium bonds and cash-holdings. Liabilities include third-party reported values of debt in financial institutions, mortgage credit debt, credit and debit card debt, student debt, and all other liabilities such as unpaid taxes which are not deposited. After 1997, the asset statement is based solely on third-party reporting. Unlike in Sweden, we do not observe closely-held businesses or unlisted shares held by taxpayers neither before or after the wealth tax repeal. More information of the Danish wealth tax data can be found in [Jakobsen et al. \(2020\)](#).

These data are linked with a longitudinal dataset containing rich information on all earnings, transfers and demographics from 1989 to 2006, which include information on age, family situation or occupation.

Like in Sweden, we have access to detailed citizenship and migration information such as daily dates of entry and exit in the country. Individuals working in Denmark must obtain a personal identification number (CPR) to pay taxes, rent an apartment, or register with health insurance. The application for a CPR contains detailed questions about citizenship, country of origin, and date of entry in Denmark. Individuals must also register their move to the CPR office to stop paying taxes in Denmark.

A summary of the different variables and datasets in Denmark and Sweden is presented in Table II.1. In Sweden, we have better information on unlisted companies, and will therefore focus on this country for our analysis of the effects of out-migration on firm-level outcomes.

Table II.1: **Sweden and Denmark**

	Sweden	Denmark
<i>Panel A. Wealth Taxation</i>		
Period of implementation	1910-2007	1903-1997
Exemption Threshold (percentile of wealth distribution)	96th-98th	98th
Tax ceiling	Y	Y
Business Assets	Exempted	Taxed
Number of brackets	2	2
Maximum Top MTR in the period	1.5%	2.2%
Revenue	0.16% of GDP in 2006	0.06% of GDP in 1996
<i>Panel B. Data Availability</i>		
Migration (in or out) date	1990-2019	Y
Duration of stay abroad (days)	1986-2019	Y
Closely-Held Businesses	2000-2017	N
Income and transfers	1990-2017	Y
Inheritance	2001-2005	Y

Notes: This table summarizes the data and institutional environments for the wealth tax in Sweden and Denmark.

II.4. Identifying Ownership Links in Serrano Dataset

Serrano is a firm-year level dataset, whose population comprises all the firms registered in Sweden between 1998 and 2021. It includes general company information retrieved from Statistics Sweden, data on financial statements, bankruptcy and mergers compiled from the Swedish Companies Registration Office (Bolagsverket), and group data from the Bisnodes group registry. For our analysis, we focus on the group data.

The group data are at the parent-subsidary-year level. For each company (“subsidiary firm”) registered in Sweden and whose ownership shares are held at least partially by other firms (“parent firms”), the group data list its parent companies and the share of the subsidiary they own in every year. We can also see links between associated companies and parent companies, where the ownership share is between 20 and 50%. The reason for the cutoff is that as a parent firm owning a minority stake at 20% or more gives certain rights vis-à-vis the majority owner.² We thus see all ownership links between domestic firms in the SERRANO dataset, except for firms that own less than 20% of another firm. We merge this information on parent companies and subsidiaries links with our firm ownership registries. To do so, we proceed in two steps.

The first step consists in reconstructing the entire paths of firm links using the group data in Serrano. For example, assume the following group structure: firm A owns 50% of firm B, firm B owns 100% of firm C and firm C owns 40% of firm D. In the Serrano data, this structure is decomposed into the links A-B, B-C and C-D. By merging these links sequentially, we are able to reconstruct the path A-B-C-D. When we implement this sequential merge, we also merge the ownership shares of parent firms in their subsidiaries. Once we reconstruct the entire path of links for each group of firms, we can multiply these ownership shares to compute the integrated ownership share. In the example, the integrated ownership share of firm A in firm D is 20%. In the dataset of firm links that we build, we do not only keep the final link paths (e.g., A-B-C-D) but also each intermediate link path (e.g., A-B-C, B-C-D, A-B, B-C, C-D) and the corresponding integrated ownership shares.³ For each link path we reconstruct, we keep the first and the last firm and we define these as parent-subsidary pairs. Two firms forming a parent-subsidary pair

²For instance, the minority owner can (i) ask for a special “minority accountant” at the firm’s expense to make sure the accounting is made properly, (ii) call for an extra general meeting to discuss firm issues (likely so that other minority owners can be alerted of potential wrongdoings by the majority owner), (iii) make the firm sue the CEO or board members if they have caused damage to the firm, and (iv) call on the court to make a compulsory liquidation of the firm.

³Since ownership cycles (e.g., firm A owns a share of firm B, which in turn owns a share of firm A) are very few in the data, we deal with them by stopping the sequential merge (e.g., we would keep links A-B and B-A but not A-B-A or B-A-B).

may be linked via multiple paths. Therefore, we sum the integrated ownership shares computed for each parent-subsidary pair across all the paths that link them and use this as our final measure of integrated ownership of the parent firm in the subsidiary firm.

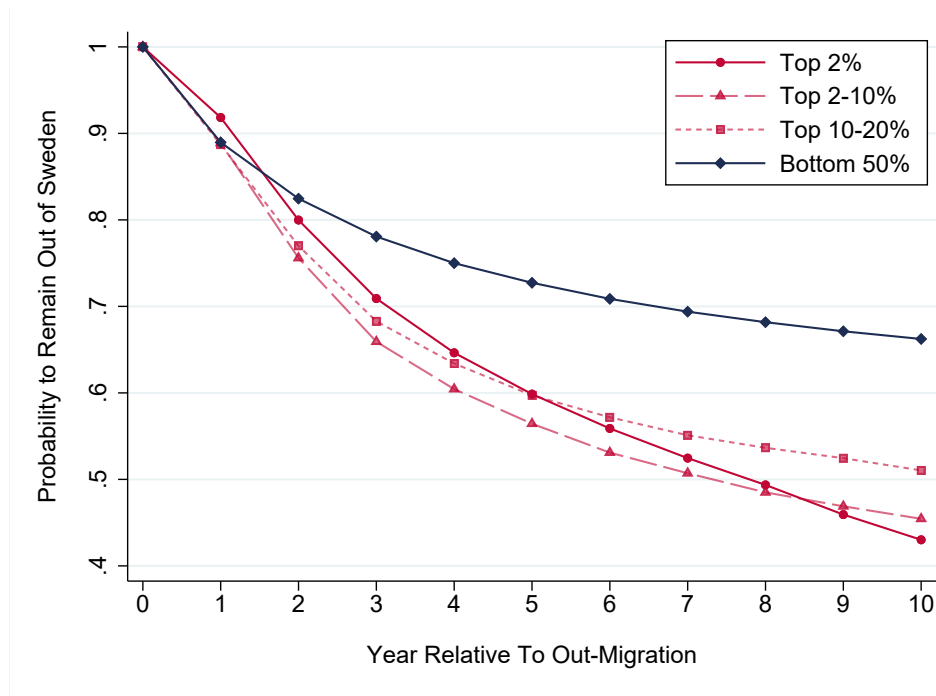
In a second step, we match the parent companies to the registry of closely-held businesses. When a parent company appears in the registry of closely-held businesses, we observe the individuals who own it. We define these individuals as the indirect owners of the subsidiary firm held by their closely-held business. The ownership share of these individuals in the subsidiary is defined by the integrated ownership share defined previously. Continuing with the example and assuming that firm A is a closely-held business, the owners of firm A are also indirect owners of firm B (with a 50% share), firm C (with a 50% share) and firm D (with a 20% share). By implementing this approach, although we can only observe the direct owners of closely-held businesses, we can observe the indirect owners of both closely-held and not closely-held businesses, provided that their parent companies are closely-held businesses. We use the integrated ownership shares to allocate employment and other economic outcomes of Swedish firms to their final owners in Sweden.

III. Additional Descriptive Evidence on Migration Patterns in Scandinavia

III.1. Duration analysis

Figure III.1 documents the duration pattern of migration events at the top of the wealth distribution. We observe the number of days spent in Sweden by each individual each year in the administrative migration register. Thus we can check that migration events reflect real mobility responses, rather than bogus changes in tax residency. In Figure III.1, we show the probability to remain outside Sweden following an out-migration event for individuals with different level of wealth. For all groups, mobility is a real response and is quite persistent over time. One year after out-migrating, the probability to remain outside Sweden is around 90% for both wealthy and non-wealthy individuals. However, migration also appears to be transitory for a large fraction of migrants: for instance, roughly 40% of the wealthy are back after 5 years. Migration events are also less persistent among the very wealthy than for rest of the population.

Figure III.1: Probability to Remain Out of Sweden After an Out-Migration Event



Notes: This figure shows the probability to remain out of Sweden after an out-migration event. The out-migration events considered for this plot cover the years 2000-2007. The wealth groups are based on the net wealth of individuals' household in the year before their out-migration date. Each year that follows an out-migration event, we measure the probability that the individual still lives outside Sweden, or has migrated back to Sweden.

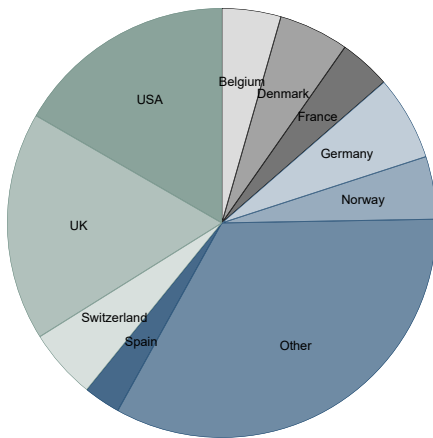
III.2. Destination and origin countries

We also describe the geographic patterns of migration at the top of the wealth distribution. Figure III.2 decomposes destination and origin countries for Swedish citizens with different level of wealth before (respectively after) their migration event. The top destination for wealthy Swedes is the UK. Capital taxes are lower in the UK compared to Sweden and there has been no tax on wealth. Furthermore, Swedish individuals can benefit from the preferential “non-dom” tax status in the UK, which exempts any of their non-UK income from taxation, to the extent that it is not repatriated in the UK. The second top destination is the US. In 2005 Sweden renegotiated the 1995 tax treaty with the US. The alteration was removing the tax on dividends paid from a US subsidiary to a Swedish parent company. Sweden had already unilaterally removed this tax. Two other noteworthy destination are Switzerland, and Austria. The latter has a beneficial bilateral tax treaty exempting taxation of capital gains for Swedish expatriates. Overall, destination countries of wealthy Swedish taxpayers are predominantly lower tax countries, suggesting that geographic migration patterns are in large part driven by tax concerns.

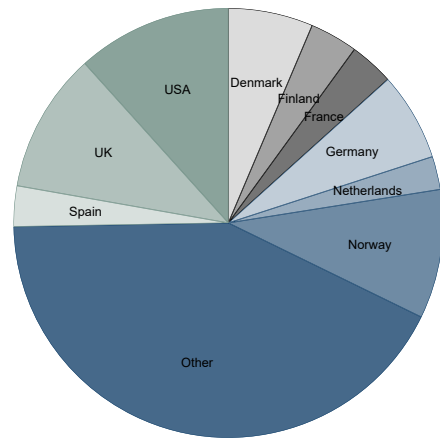
Figure III.2: Migration at the Top in Sweden: Main Origin and Destination Countries

A. Origin Countries

A1. Top 2%

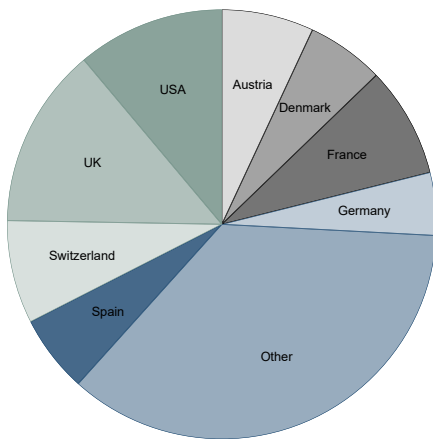


A2. Top 10-20%

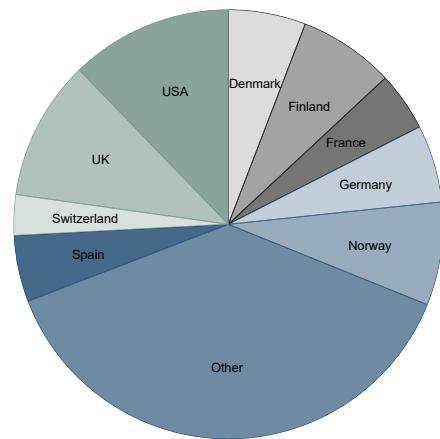


B. Destination Countries

B1. Top 2%



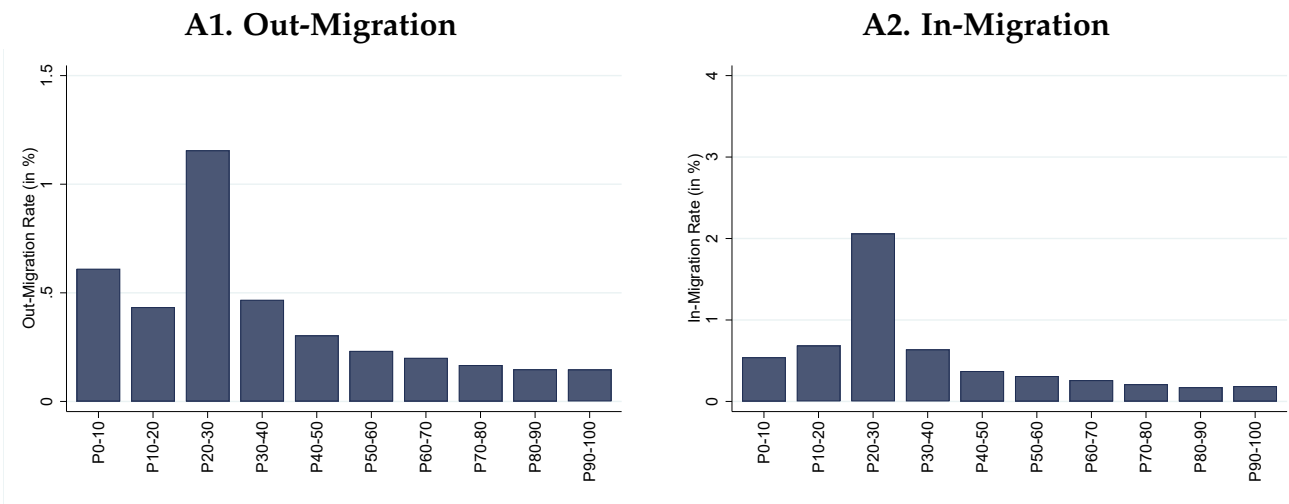
B2. Top 10-20%



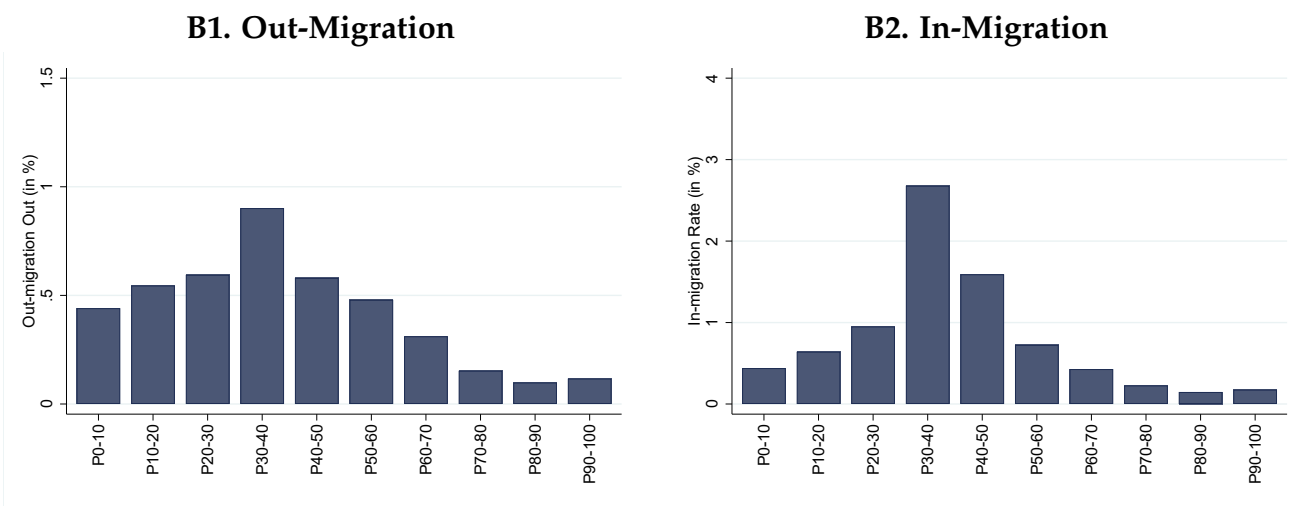
Notes: This figure describes the main origin countries of individuals in-migrating into Sweden (Panel A) and the main destination countries of individuals out-migrating from Sweden (Panel B) by wealth group (individuals whose household net wealth is either in the top 2% or in the top 10-20%).

Figure III.3: Wealth and International Migration Patterns in Scandinavia

A. Sweden (1999-2006)



B. Denmark (1989-1996)

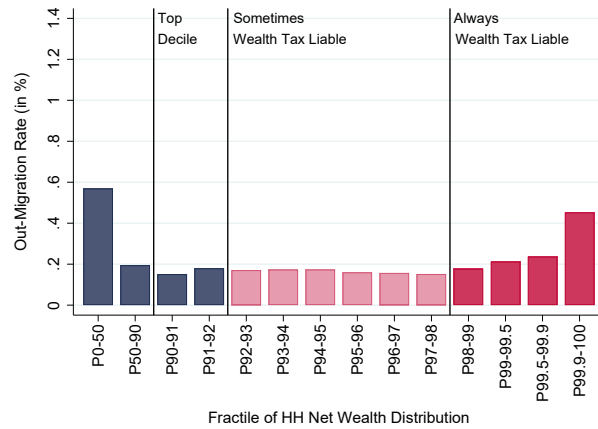


Notes: This figure shows in-migration and out-migration rates by decile of total net worth, in Sweden (Panel A) and Denmark (Panel B). We compute those statistics during the last years when the wealth tax was still in place in each country, which corresponds to the period 1999-2006 in Sweden and the period 1989-1996 in Denmark.

Figure III.4: Wealth and International Migration Patterns in Sweden Between 2000 and 2004

A. Migrations in 2000

A1. Out-Migration

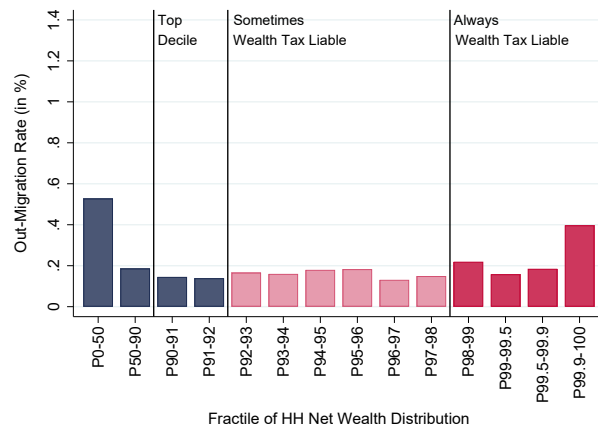


A2. In-Migration

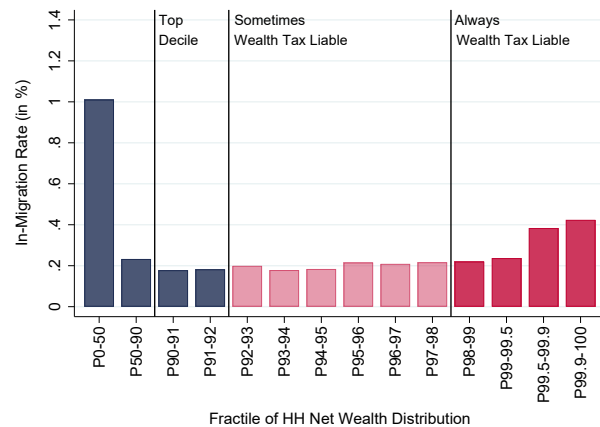


B. Migrations in 2001

B1. Out-Migration

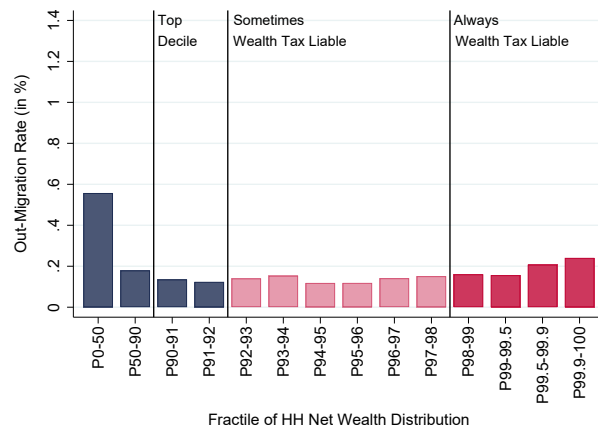


B2. In-Migration

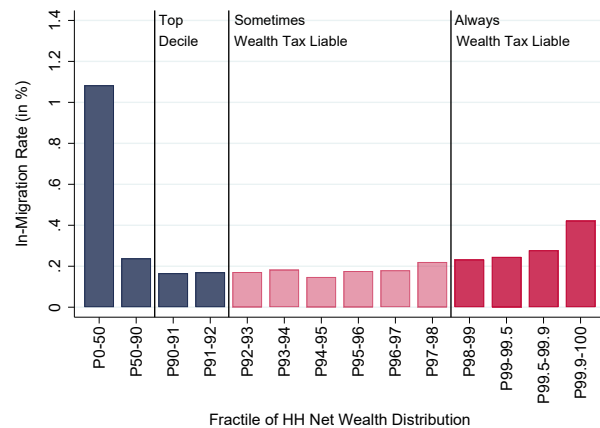


C. Migrations in 2002

C1. Out-Migration

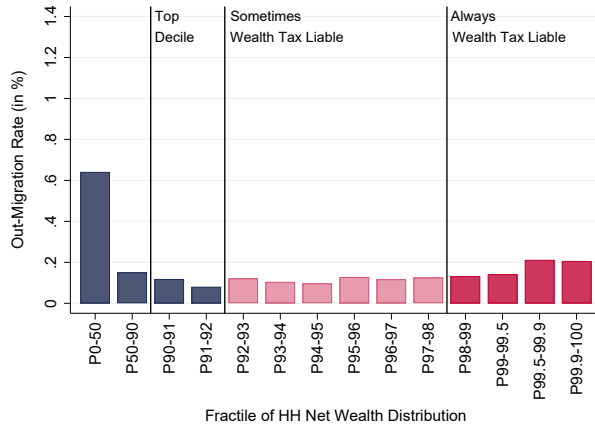


C2. In-Migration

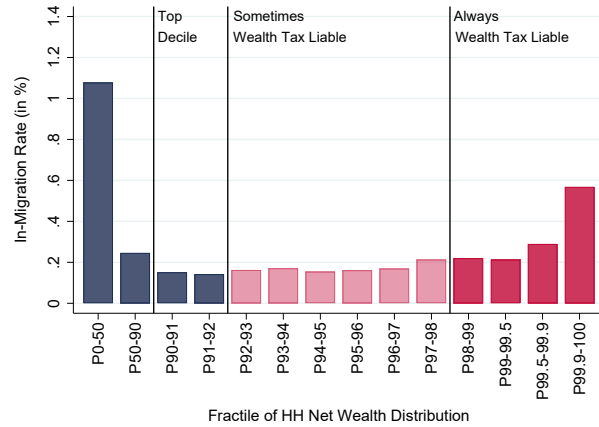


D. Migrations in 2003

D1. Out-Migration

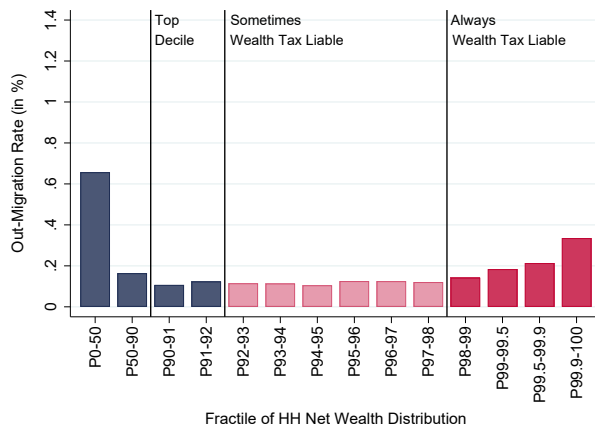


D2. In-Migration

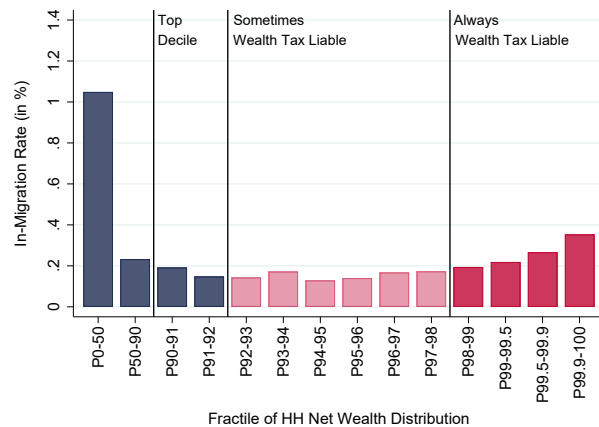


E. Migrations in 2004

E1. Out-Migration



E2. In-Migration



Notes: This figure shows in-migration and out-migration rates by level of total household net worth in Sweden for each year between 2000 and 2004. Each year, we rank individuals based on their household net worth, and show the out-migration and in-migration rates for each fractile of the household net wealth distribution.

III.3. Selection: Who Migrates In and Out?

We study selection into migration out of and into Sweden. We perform the analysis during the period 2000-2006 when the wealth tax was in place. We start by asking what individual characteristics can predict out-migration (respectively in-migration) in Sweden. We run a first regression model:

$$\mathbb{P}\{Y = 1\} = \beta' X_0 \quad (6)$$

where X_0 is vector of individual characteristics. The set of characteristics includes age bins, education bins, a dummy for being a closely-held active business owner, a dummy for being an independent contractor, a dummy for being foreign born, and cognitive skills quartiles.⁴ To ease computation, we run this specification on the full population of the wealthy (top 2% of wealth distribution) and on a 10% random sample of the rest of the population.⁵

The blue coefficients in Figure III.5 correspond to the estimated coefficients on each of those demographics estimated on the all Swedish individuals, regardless of their wealth. We rescale the coefficients by the average predicted probability of out-migration, in order to interpret our coefficients in relative terms. In the general population, we find that younger individuals are much more mobile than others. Education is also a strong predictor of mobility, with the more educated more likely to move out of, or move into Sweden, although PhDs are much more likely to out-migrate than in-migrate. The gradient of self-selection into mobility follows the same pattern for cognitive ability, with higher skills leading to more mobility out of and into the country. On the contrary, in the full population, entrepreneurs and business owners are less likely to migrate than others. This suggests that owning a firm could be associated with large mobility costs and frictions that make it harder to leave the country.

We then focus on self-selection into migration at the top of the wealth distribution. To understand if selection patterns into migration are different for the very wealthy, we estimate:

$$\mathbb{P}\{Y = 1\} = \mathbb{1}\{W < P_W\} \cdot \beta' X_0 + \mathbb{1}\{W \geq P_W\} \cdot \beta'_w X_0 \quad (7)$$

Where we use X_0 to denote vectors and we set $P_W = P_{98}$, i.e. the 98th percentile of the

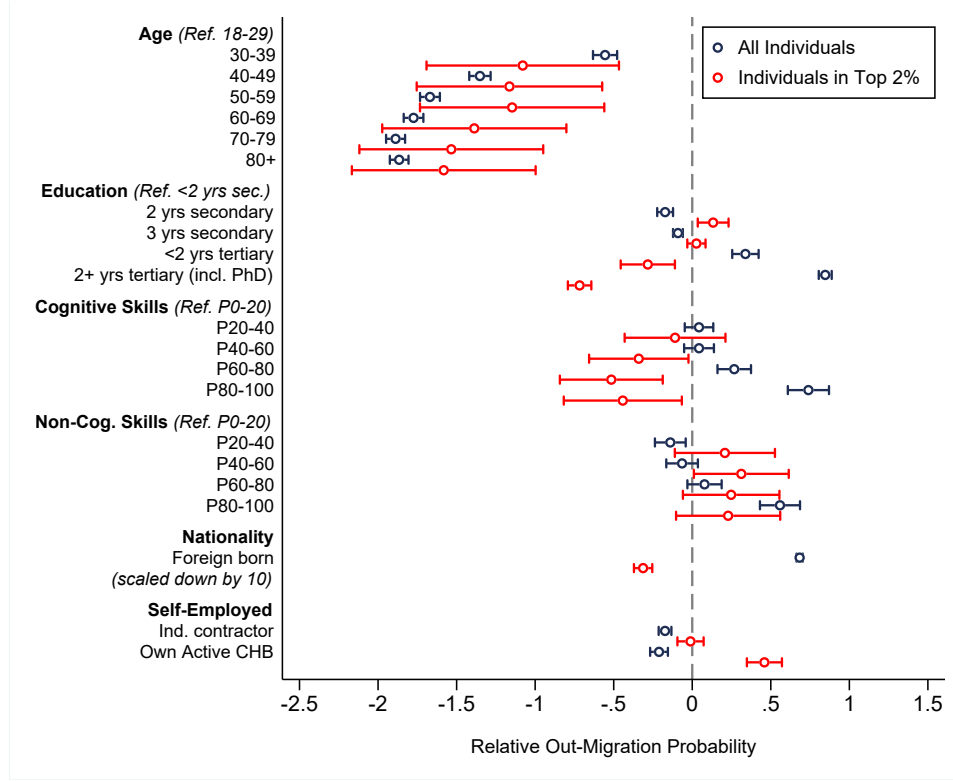
⁴Active closely-held businesses are firms that have at least one employee beyond the owner. For cognitive skills, we proceed by running a separate regression because cognitive skills come from enlistment data and are therefore available for a limited subset of people.

⁵When using a 10% random sample of the entire population of the wealthy, we found similar results. The baseline approach instead allows to exploit all variations in our migration events at the top, that we will study later.

wealth distribution. The red coefficients in Figure III.5 plot the estimated vector of coefficients β_w representing the effects of the various components of X_0 on the out-migration (respectively in-migration) probability for individuals in the top 2% of the wealth distribution.

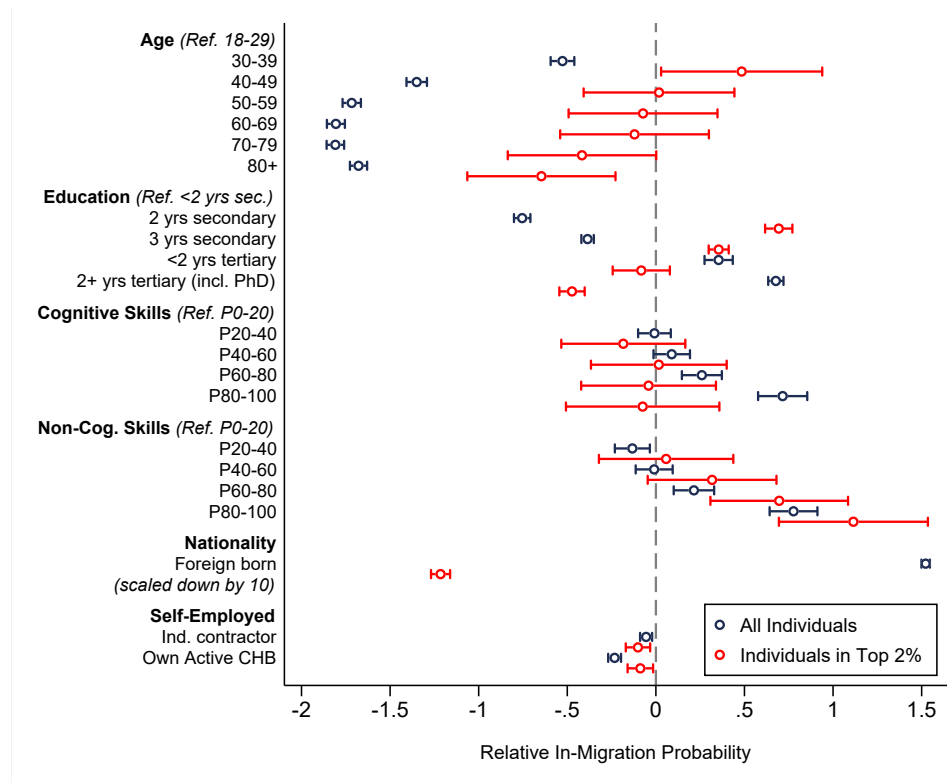
The results show that selection patterns differ among the very wealthy. For individuals in the top 2% of the wealth distribution, the education and cognitive ability gradient are inverted. In other words, while there is positive selection (“brain drain”) into migration for the population at large, the reverse is true at the top of the wealth distribution. We also find that wealthy business owners do appear to be more likely to out-migrate, but not to in-migrate. The estimated (rescaled) coefficient of 0.41 in the top panel means that all else equal, individuals in the top 2% who control an active firm in Sweden are 40% more likely than wealthy non-entrepreneurs to outmigrate. The effect was of the opposite sign when focusing on the entire population. On the contrary, being an entrepreneur is not associated with higher propensity to become a Swedish resident for wealthy individuals during that period, as shown in the bottom panel. In Appendix Table III.1, we also investigate whether owners of fast-growing firms are more likely to leave Sweden than others. We find that wealthy owners controlling firms with high value added growth rates are substantially more likely to out-migrate.

Figure III.5: Selection into Out-Migration from Sweden



Notes: This figure describes selection into out-migration from Sweden during the period 2000-2007. We estimate Equation 6 and plot the estimated vector of coefficients β and their confidence intervals in blue. We estimate Equation 7 and plot the estimated vector of coefficients β_w and their confidence intervals in red. All coefficients are rescaled by the average predicted probability of out-migration. The estimation sample includes all taxpayers with household net wealth in the top 2%, and a 10% random sample of all other Swedish individuals. The information on cognitive and non-cognitive skills is available only for a subsample of Swedes who passed enlistment tests.

Figure III.6: Selection into In-Migration to Sweden



Notes: This figure describes selection into in-migration in Sweden. We focus on all individuals moving to Sweden over the period 1999-2005, when the wealth tax was still in place. Blue coefficients come from estimates of Equation 6 detailed in the text, while red coefficients are estimated on the interaction term of Equation 7. All coefficients are rescaled by the average predicted probability of in-migration. The estimation sample includes all taxpayers with household net wealth in the top 2% and a 10% random sample of all other Swedish individuals.

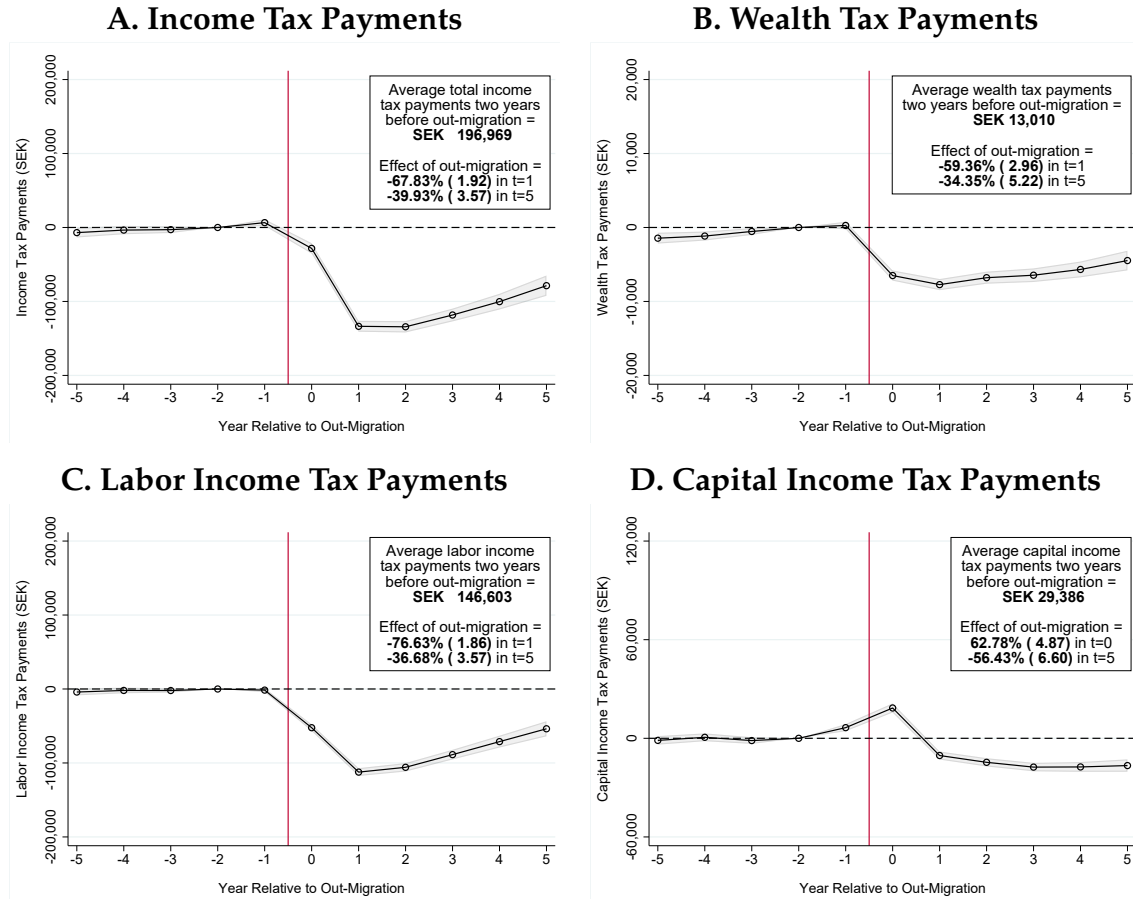
Table III.1: Selection into Out-Migration from Sweden for Owners of Fast-Growing Active Closely-Held Businesses

	All Individuals	Individuals in Top 2% of Net Wealth
<i>Value Added Growth</i>		
Year-on-Year, Top 25%	-20.01% (6.54)	65.95% (12.50)
Past 3 Years, Top 25%	-19.09% (7.13)	93.41% (13.88)
Past 5 Years, Top 25%	-32.10% (7.52)	85.97% (14.97)
Year-on-Year, Top 50%	-25.31% (4.10)	51.99% (7.87)
Past 3 Years, Top 50%	-27.90% (4.13)	54.01% (8.00)
Past 5 Years, Top 50%	-27.47% (4.77)	52.11% (9.36)
<i>Employment Growth</i>		
Year-on-Year, Top 25%	-29.86% (6.22)	40.60% (12.29)
Year-on-Year, Top 50%	-19.92% (4.43)	32.53% (8.82)

Notes: This table describes selection into out-migration for owners of fast-growing active closely-held businesses (“CHB”) in Sweden. Each row considers a different indicator of ownership of a fast-growing active CHB. We compute the growth rate of each firm based on either its value added or its number of employees over different intervals of time (year-on-year, past 3 years, past 5 years). We then create yearly ranks of active CHBs based on those growth rates. In any given year, we define an individual as the owner of a fast-growing active CHB if the highest growth rank among the active CHBs she owns in that year is in the top 25% or above the median, depending on the specification. We use this indicator to estimate Equation 6 and Equation 7, together with age bins, education bins, a dummy for being foreign-born and a dummy for being an independent contractor. We run one separate regression for each indicator of ownership of a fast-growing active CHB shown in each table row. The equations were estimated on a sample including all individual-year observations where the individual was in the top 2% of household net wealth and a random 10% of the remaining observations. The first column of the table displays the estimated coefficients for each ownership indicator from Equation 6. The second column displays the estimated coefficients for each ownership indicator interacted with a dummy for being in the top 2% of household net wealth from Equation 7. All the coefficients displayed in the table are rescaled by the average predicted probability of out-migration and multiplied by 100 to be interpreted in relative terms.

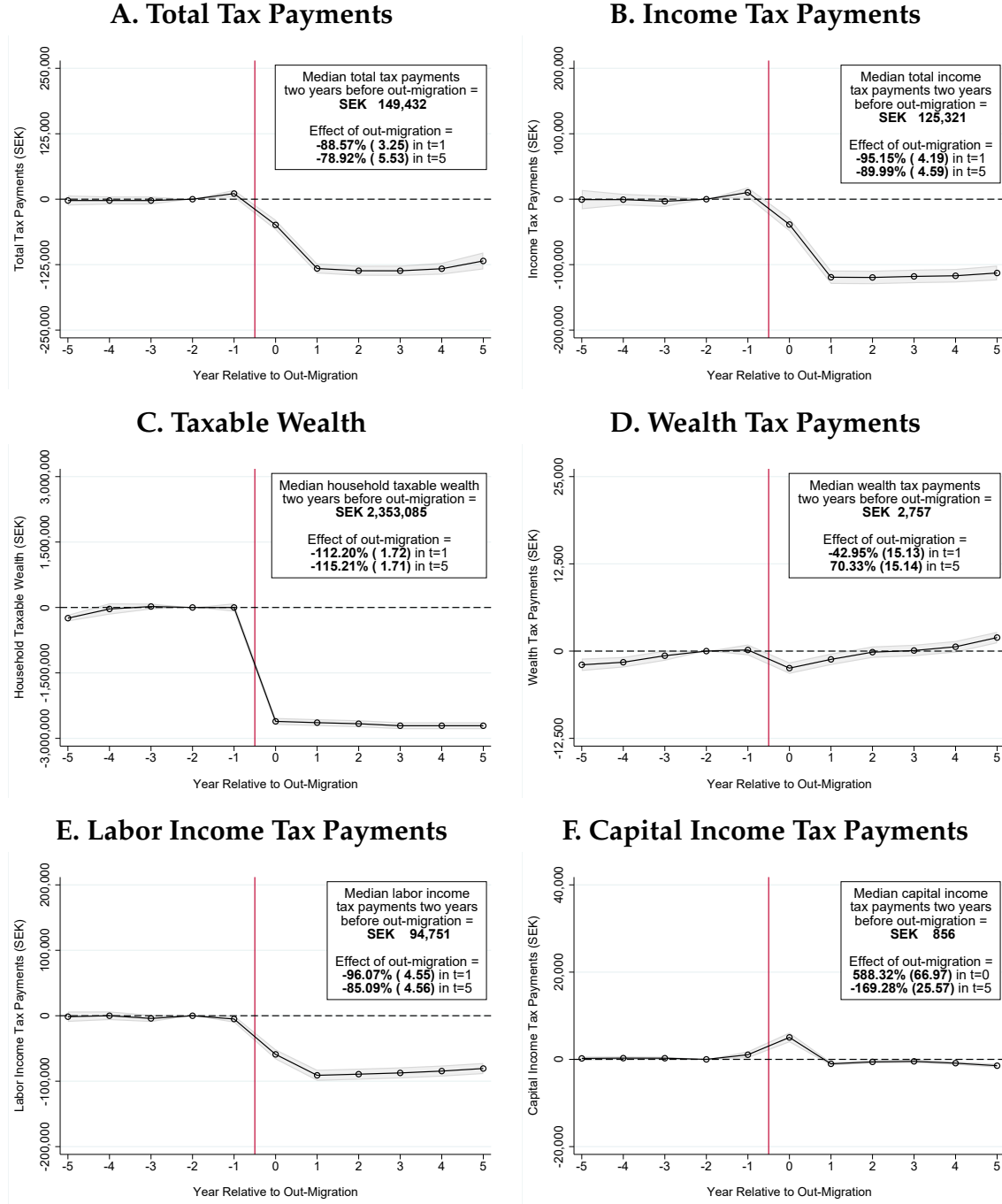
IV. Impact of Migration on Individual and Firm Outcomes: Additional Event Study Results

Figure IV.1: Effects of Wealthy Out-Migration on Tax Payments



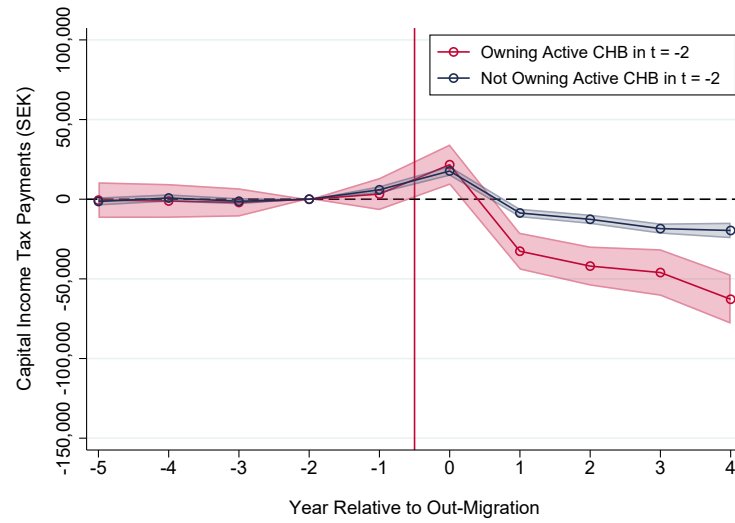
Notes: This figure describes the evolution of several additional tax outcomes of wealthy individuals' before and after they leave Sweden, compared to control wealthy individuals who do not move that same year. These include the components of total tax payments shown in Figure 6, Panel A. The sample includes individuals in the top 2% of the household net wealth distribution in Sweden for at least one year before their true or placebo out-migration date. We focus on out-migration events occurring between 2000 and 2007, with wealth ranks drawn from 1999-2006, when the wealth tax was still in place. We winsorized the bottom 1% and top 5% of all outcomes. We plot the estimates β_j from Equation 7 and their 95% confidence intervals. The estimates displayed in the text boxes are computed as the estimate of β_j when $t = 0$, $t = 1$ or $t = 5$ divided by the average outcome in the treatment group in $t = -2$, multiplied by 100. The standard errors are rescaled using the same approach.

Figure IV.2: Median Effects of Wealthy Out-Migration on Tax Payments, Top 2%



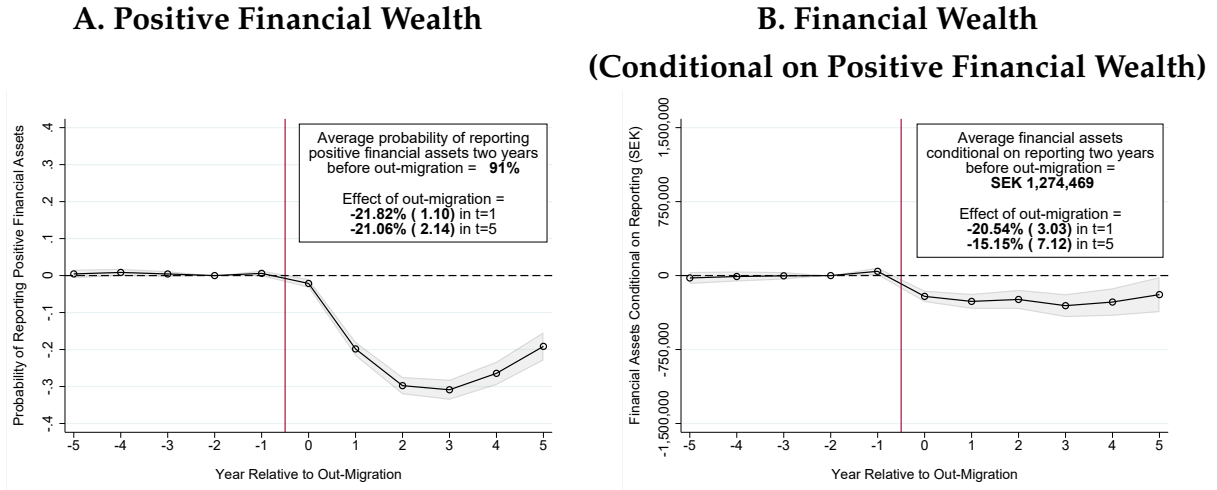
Notes: This figure describes the evolution of wealthy individuals' outcomes before and after they leave Sweden, compared to control wealthy individuals who do not move that same year. The sample includes individuals whose household net wealth was in the top 2% of the household net wealth distribution in Sweden for at least one year before their true or placebo out-migration date. We plot the estimates β_j from Equation 7 in the main text estimated via median regressions.

Figure IV.3: Effect of Wealthy Out-Migration on Capital Income Tax Payments, Top 2%



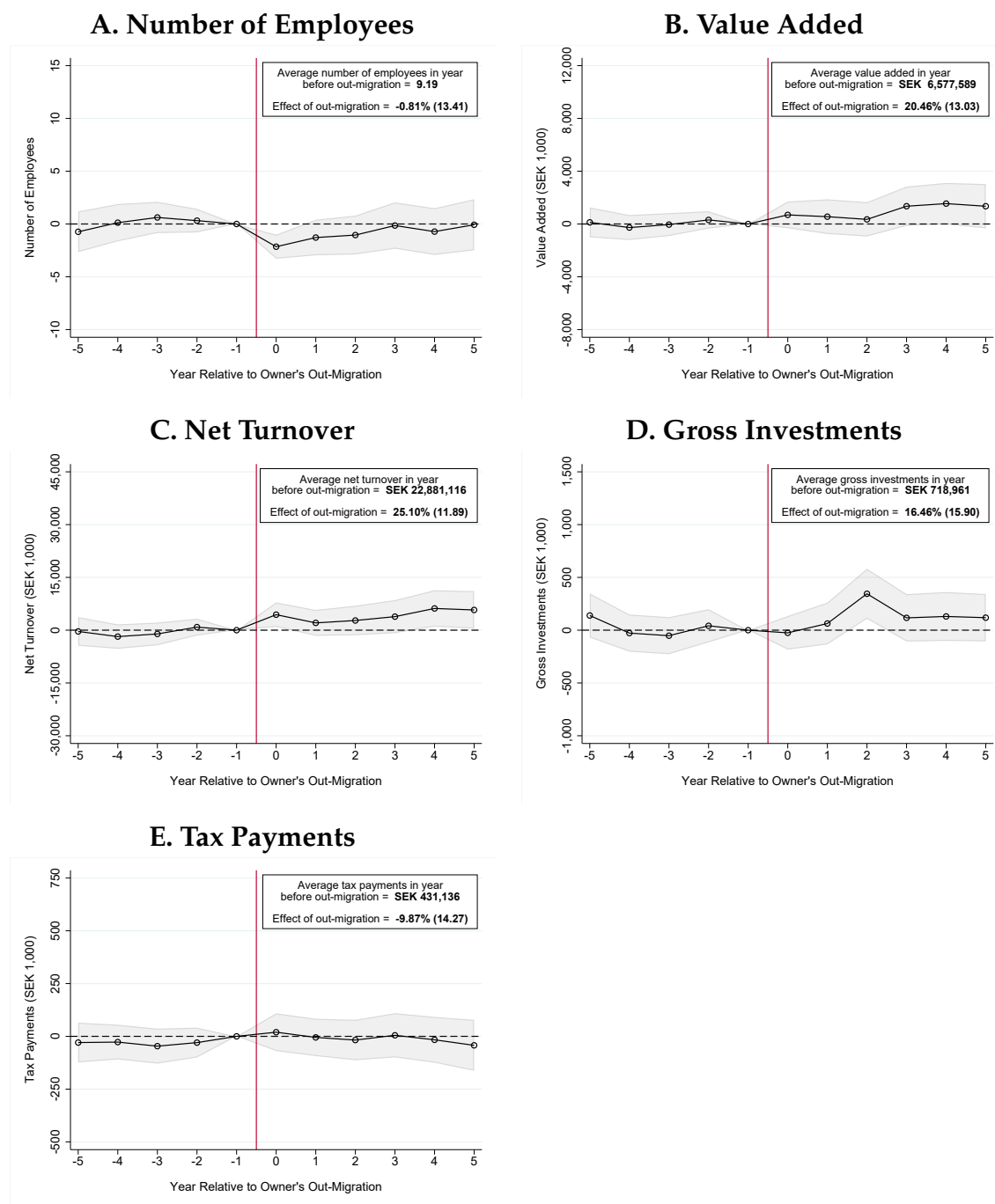
Notes: This figure describes the evolution of wealthy individuals' capital income tax payments before and after they leave Sweden, compared to control wealthy individuals who do not move that same year. We focus on individuals in the top 2% of the household net worth distribution in Sweden for at least one year before their true or placebo out-migration date. We focus on out-migration events occurring between 2000 and 2007, with wealth ranks drawn from 1999-2006, when the wealth tax was in place. We plot the estimates β_j from Equation 7 in the main text. We estimate the equation separately for wealthy individuals who own an active firm before out-migrating (red series) and wealthy individuals who do not control an active firm (blue series).

Figure IV.4: Effects of Wealthy Out-Migration on Portfolio Composition



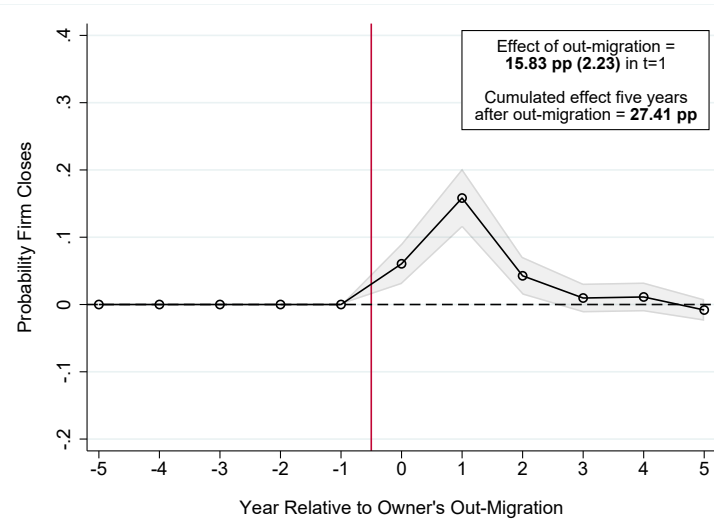
Notes: This figure describes the evolution of wealthy individuals' outcomes before and after they leave Sweden, compared to control wealthy individuals who do not move that same year. The sample includes individuals who were in the top 2% of the household net worth distribution in Sweden for at least one year before their true or placebo out-migration date. We focus on out-migration events occurring between 2000 and 2007, with wealth ranks drawn from 1999-2006, when the wealth tax was in place. We plot the estimates β_j from Equation 7 in the main text and their 95 percent confidence intervals. The estimates displayed in the text boxes are computed as the estimate of β_j when $t = 1$ or $t = 5$ divided by the average outcome in the treatment group in $t = -2$, multiplied by 100. The standard errors are rescaled using the same approach.

Figure IV.5: Effects of Wealthy Out-Migration on Closely-Held Businesses Outcomes Conditional on Existence of Firm, Top 2%



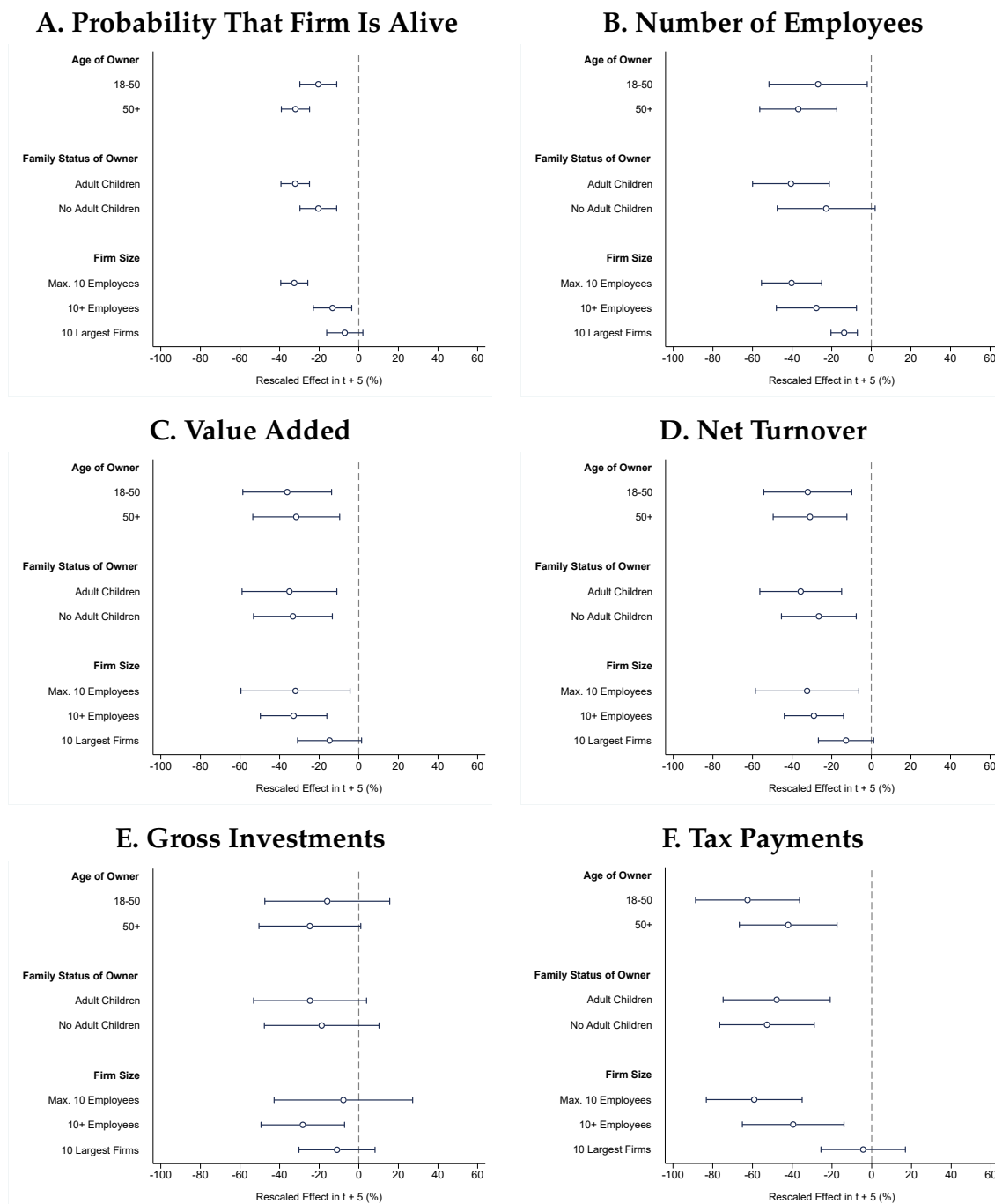
Notes: This figure shows the effects of wealthy owners' out-migration events on active firms controlled by individuals who were in the top 2% of the household net worth distribution in Sweden for at least one year before their migration event. We focus on out-migration events occurring between 2001 and 2007, with wealth ranks drawn from 2000-2006. We plot the estimates of β_j from Equation 7 in the main text. We restrict the analysis to firms that keep being active in Sweden throughout the period.

Figure IV.6: **Effect of Wealthy Out-Migration on Closure of Closely-Held Businesses, Top 2%**



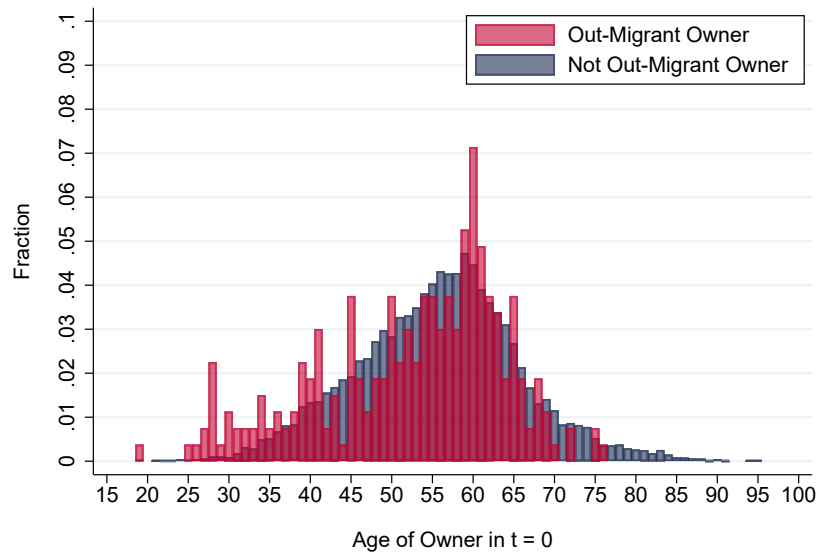
Notes: This figure shows the effects of wealthy owners' out-migration events on the probability that the firm they control before the out-migration event closes. We focus on out-migration events occurring between 2001 and 2007. The sample includes firms that in $t = -1$ were active closely-held businesses owned by individuals whose real or placebo out-migration happened in year t and who were in the top 2% of the household net worth distribution in Sweden for at least one year before t . We plot the estimates of β_j from Equation 7 in the main text.

Figure IV.7: Effects of Wealthy Out-Migration on Closely-Held Businesses Outcomes by Wealthy Owners' Characteristics



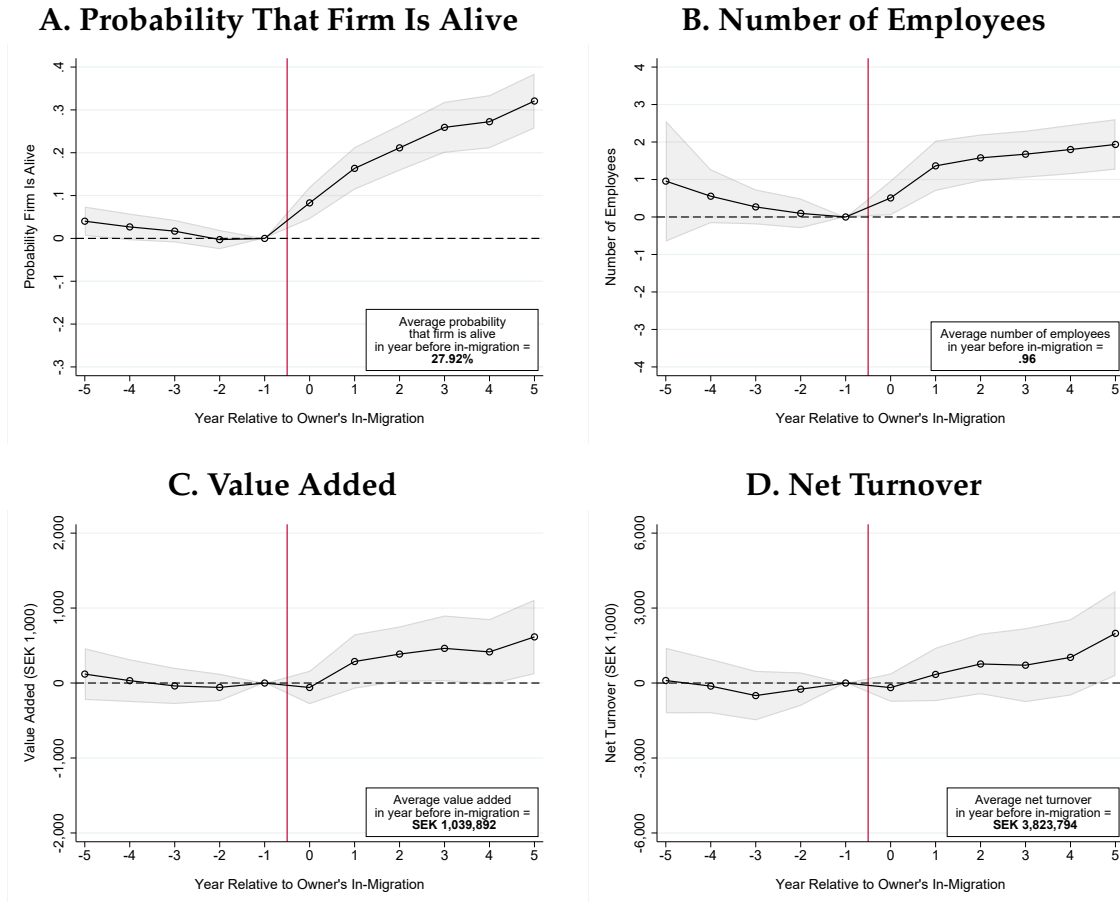
Notes: This figure shows the effects of wealthy owners' out-migration events on firm-level outcomes by wealthy owners' characteristics. Each coefficient and its confidence intervals refer to one separate regression. We focus on out-migration events occurring between 2001 and 2007. The sample includes active closely-held businesses controlled by wealthy individuals in the year $t = -1$, with (real or placebo) out-migration events occurring in the subsequent year $t = 0$. We plot the estimates of β_j and their confidence intervals estimated from Equation 7 in the main text. The effects displayed in the figures are computed as the estimate of β_5 divided by the average outcome in the treatment group in $t = -1$, multiplied by 100. The standard errors are rescaled using the same approach.

Figure IV.8: Age Distribution of Closely-Held Business Owners, Top 2%



Notes: This figure shows the age distribution of the wealthy owners of the closely-held businesses in our firm-level event-study sample in $t = 0$. These entrepreneurs own one of the closely-held businesses in the sample in $t = -1$ and are in the top 2% of the household net wealth distribution for at least one year before their observed or placebo out-migration event.

Figure IV.9: Effects of Wealthy In-Migration on Firm Outcomes, Top 2%



Notes: This figure shows the effects of wealthy owners' in-migration events on firm-level outcomes. The sample includes firms that were held by individuals who were in the top 2% of the household net worth distribution in Sweden for at least one year since their true or placebo in-migration date and before their potential following out-migration. The sample is further restricted to firms who were active closely-held businesses in at least one of the years in which they were held by the wealthy true or placebo in-migrant owner (between her true or placebo in-migration date and her potential following out-migration). We study in-migration events during the period 1999-2006, when a wealth tax was still in place in Sweden. We winsorize the bottom 1% and top 5% of all outcomes except for the number of employees, for which we winsorize only the top 5%. We plot the estimates of β_j from the in-migration counterpart of Equation 7 in the main text.

Table IV.1: Test of Equality of Out-Migration and In-Migration Effects in Firm-Level Event Studies

<i>Sample of CHBs with owner in the top 2% of net worth</i>			
Outcome	Effects In Levels		
	Out-Migration $t_{out} = +5$	In-Migration $t_{in} = +5$	T-Statistic $t_{out} = -t_{in}$
Prob. Firm Is Alive (pp)	-27.41 (2.91)	32.07 (3.27)	1.06
Number of Employees	-2.87 (0.67)	1.94 (0.34)	-1.25
Value Added (SEK 1,000)	-2,121.31 (490.43)	615.58 (253.93)	-2.73
Net Turnover (SEK 1,000)	-6,745.50 (1,528.62)	1,987.76 (868.31)	-2.71
Gross Investments (SEK 1,000)	-139.37 (64.17)	20.36 (41.38)	-1.56
Tax Payments (SEK 1,000)	-197.26 (36.18)	11.22 (16.96)	-4.66

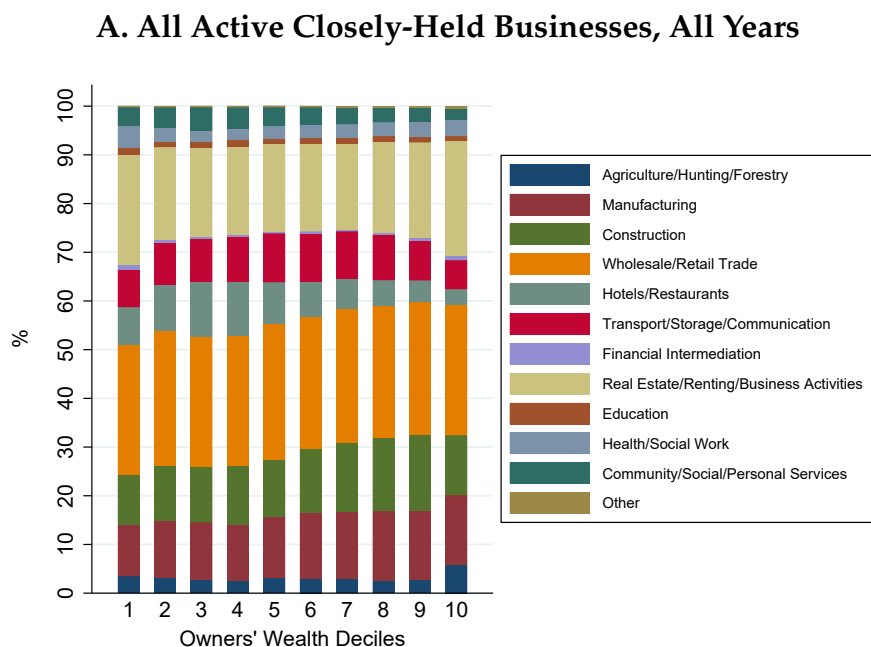
Notes: This table compares the estimates of firm-level responses to out-migration events and in-migration events of wealthy entrepreneurs (in the top 2%) in Sweden. For each firm-level outcome in each row, we present β_5 estimated from Equation 7 in the main text, focusing on out-migration events (Column (1)) and in-migration events (Column (2)). Column (3) reports the t-statistic of a test of equality for the estimates of out-migration and in-migration effects.

Table IV.2: Descriptive Statistics on the Characteristics of Firms in the Event Study Samples

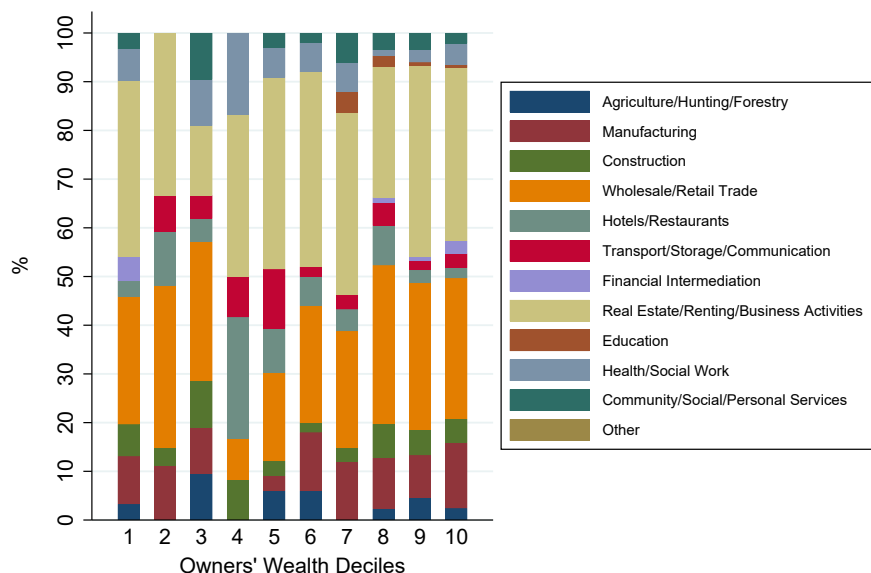
Variable	Control Firms				Treated Firms			
	Mean	Median	Std. Dev.	Obs.	Mean	Median	Std. Dev.	Obs.
<i>Panel A. Sample of active CHBs with owner in the top 2% of net worth</i>								
No. of Owners	2.20	2	3.26	14,664	5.03	2	19.10	235
No. of Employees	9	4	12	14,664	10	4	13	235
Value Added	4,398	2,038	5,872	13,445	6,260	2,985	7,567	203
Net Turnover	14,353	5,389	21,004	13,445	21,462	6,328	26,991	203
Gross Investments	560	92	1,039	13,445	689	77	1,218	203
Tax Payments	202	44	362	13,445	376	87	521	203
<i>Panel B. Sample of active firms with direct or indirect owner in the top 2% of net worth</i>								
No. of Owners	3.64	2	8.64	22,433	24.22	4	111.72	717
No. of Employees	15	5	21	22,433	23	10	27	717
Value Added	6,567	2,557	9,743	20,705	11,698	5,208	14,202	655
Net Turnover	23,229	7,300	36,845	20,705	41,877	17,625	51,043	655
Gross Investments	732	108	1,416	20,705	1,033	114	1,819	655
Tax Payments	245	42	570	20,705	435	56	946	655

Notes: This table reports descriptive statistics for the sample of firms in our event-study sample. We study active firms controlled by individuals in the top 2% of the household net wealth distribution in Sweden at least one year before their true or placebo first out-migration event, as defined by our predicted wealth measure. In Panel A, we restrict our attention to closely-held businesses ("CHBs"), while in Panel B, we additionally look at indirectly held firms, identified using the registry of ownership links. We focus on out-migration events occurring between 2000 and 2007, with wealth ranks drawn from 1999-2006. The unit of measure for value added, net turnover, gross investments, and tax payments is SEK 1,000. All outcomes are measured in year $t - 1$ relative to the true or placebo out-migration event. We winsorize the top 5% and bottom 1% of outcomes, except for number of employees, which we winsorize only the top 5% of.

Figure IV.10: **Distribution of Closely-Held Businesses Across Sectors by Owners' Net Worth**



B. Only Active Closely-Held Businesses with Out-Migrant Owners



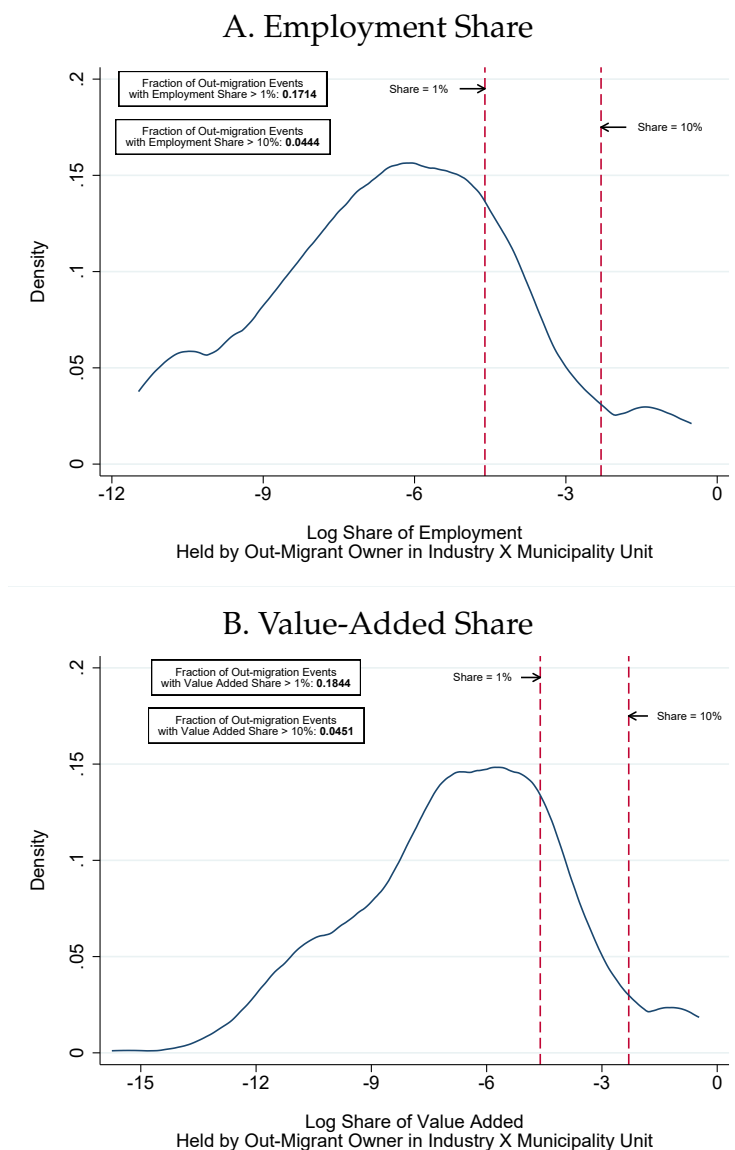
Notes: This figure shows the composition of sectors for firms controlled by Swedish individuals with different levels of household net worth during the period 2000-2006. The category “Other” includes fishing, mining/quarrying, electricity/water/gas supply, public administration/defence, private households employing domestic staff, and extra-territorial organisations. A closely-held business is defined as active if it employs at least one worker beyond the owners. For each year, closely-held businesses are assigned the highest wealth rank among their owners in a given year. Panel B considers only firms with out-migrant owners in the year before their out-migration (our “treated CHBs” sample).

Table IV.3: Effects of Wealthy Out-Migration on Closely-Held Business Outcomes in t+5 by Winsorization Method

<i>Sample of CHBs with owner in the top 2% of net worth</i>					
Winsorization	Number of Employees	Value Added	Net Turnover	Tax Payments	Gross Investments
None	-29.1% (27.41)	58.0% (49.46)	42.6% (39.85)	1.7% (45.97)	70.1% (48.48)
Top 1%	-33.7% (10.87)	-12.9% (15.62)	-22.2% (11.37)	-75.7% (38.35)	1.6% (17.49)
Top 5%	-33.3% (7.72)	-24.0% (16.27)	-31.7% (7.19)	-133.7% (70.68)	-21.9% (10.08)
Top 10%	-32.3% (6.12)	-21.6% (20.41)	-30.6% (6.34)	-208.4% (129.84)	-21.5% (8.79)
Bottom 1% + Top 1%		-21.8% (11.37)	-22.2% (11.37)	-35.8% (11.89)	1.6% (17.49)
Bottom 1% + Top 5%		-34.2% (7.91)	-31.7% (7.19)	-50.5% (9.27)	-21.9% (10.08)
Bottom 1% + Top 10%		-35.6% (6.99)	-30.6% (6.34)	-51.3% (8.07)	-21.5% (8.79)

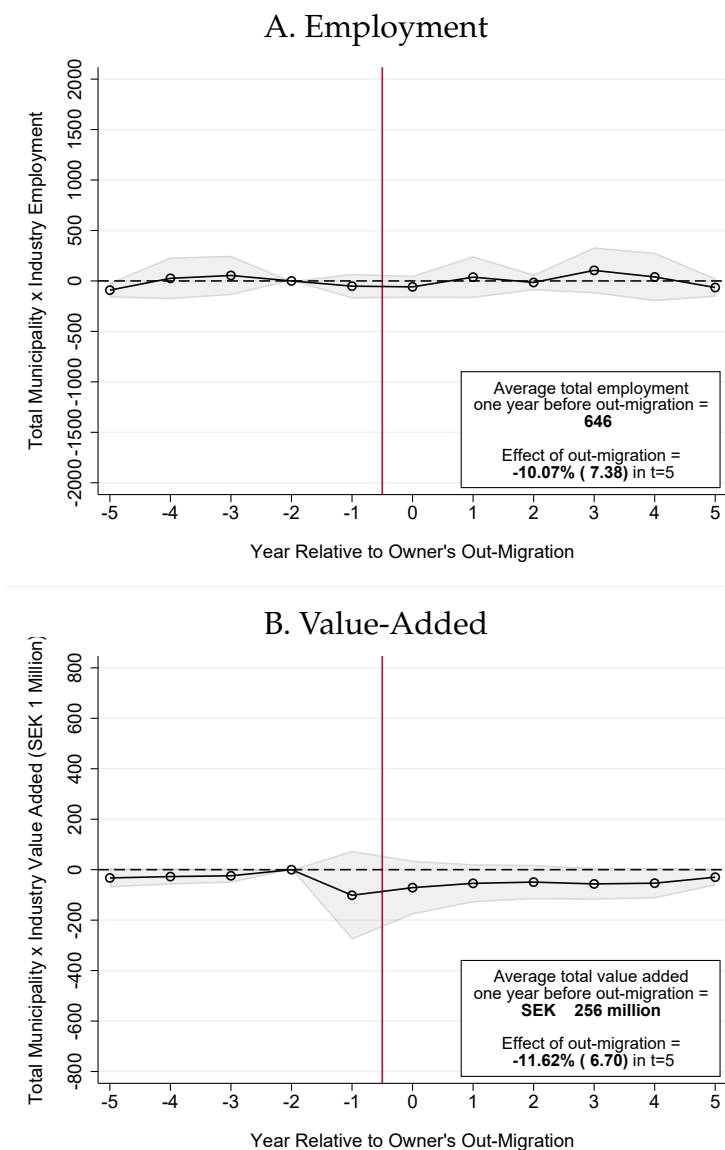
Notes: This table reports the estimated β_j in $t = 5$ from Equation 7 using various strategies to handle outliers. For each estimation procedure, we rescaled the estimated β_5 by the average outcome of the treatment group in $t = -1$ and multiplied it by 100. The standard errors are rescaled following the same approach. Each row considers a different winsorization methodology. Each column refers to a separate regression that corresponds to a specific outcome.

Figure IV.11: **Distribution of Employment and Value-Added Market Shares for Businesses of Owners Migrating Out of Sweden**



Notes: This figure plots the distribution of the log share of total employment (panel A) and total value-added (panel B) that closely-held businesses of wealthy out-migrants represent within local sectoral markets, in the year prior to migration. We define these markets as municipality times one-digit sector cells, and, for this exercise, restrict these cells to those with a total number of employees larger than 50.

Figure IV.12: **Effect of Migration of Owner of Granular Business (i.e. With Market Share > 10%) on Market-Level Employment and Value-Added**

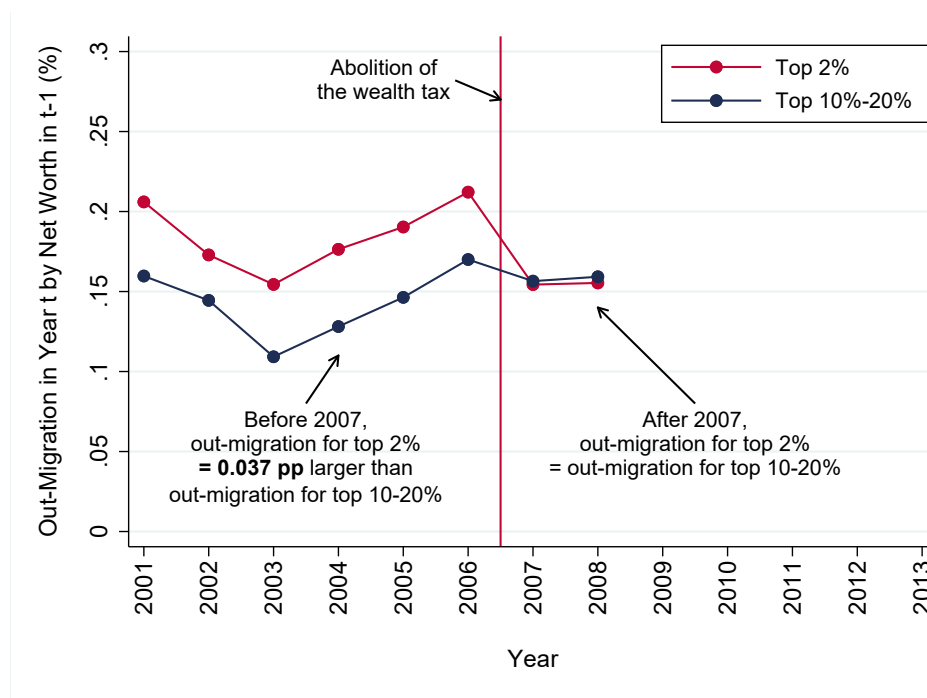


Notes: In this figure, we define local sectoral markets as municipality times one-digit sector cells, and, for this exercise, restrict cells to those with a total number of employees larger than 50. We focus on a treatment group of “granular” businesses that have an employment share of more than 10% in the year following out-migration of their owner. We run an event study on market-level outcomes (total employment in panel A and total value-added in panel B) and include a set of control markets, following the same strategy as in our baseline event study designs. The coefficients correspond to β_j from Equation 7.

V. Migration Elasticity Estimates

V.1. Additional Evidence: Sweden

Figure V.1: Out-Migration Rates



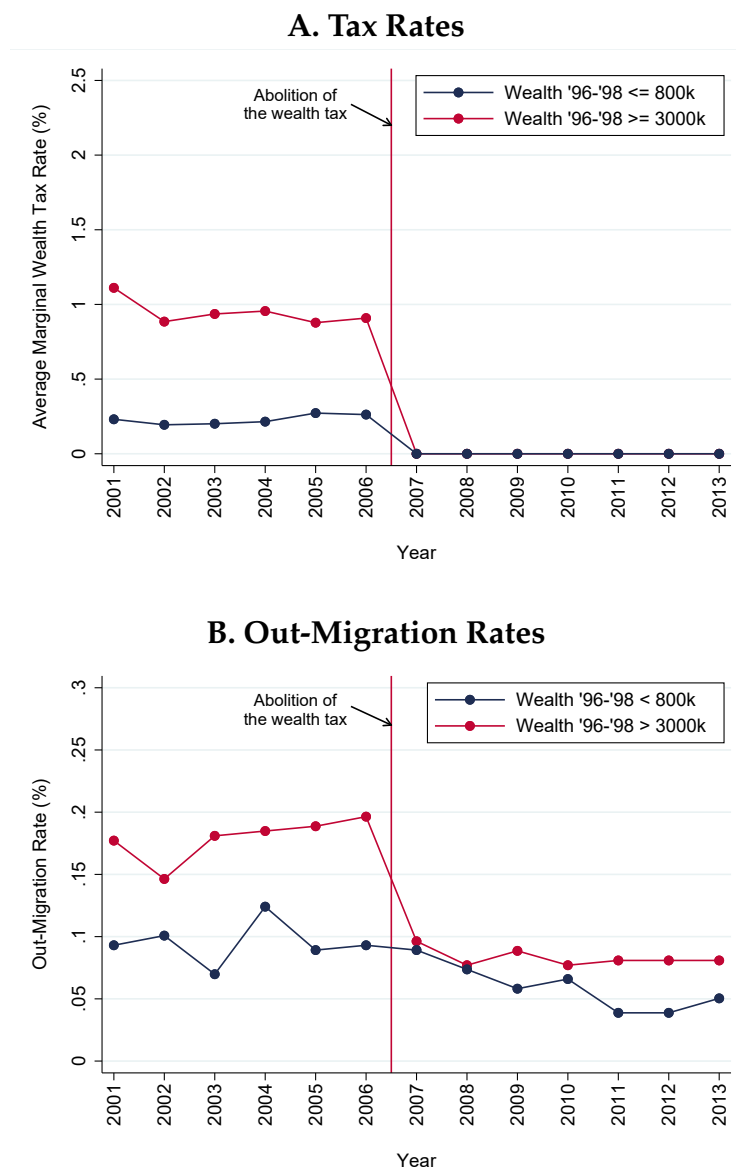
Notes: This figure shows the evolution of out-migration rates among the wealthy around the time of the wealth tax repeal. We show this for both our treated group in the top 2% of household net wealth (red series) and control group in the top 10-20% of household net wealth (blue series).

One challenge in comparing out-migration rates of wealthy individuals subject or not subject to the wealth tax is that we do not observe their total net wealth after 2007. We can circumvent this challenge by predicting individuals' wealth levels after 2007. The obtained measure of predicted wealth enables us to tackle two issues. First, we can compare out-migration patterns of treated and control individuals in the entire post-reform period, allowing to better gauge the credibility of our design and estimates. Second, using predicted instead of actual level of wealth avoids capturing potential endogenous responses to the reform through other channels than migration, for instance through savings responses following the change in wealth taxes (Jakobsen et al., 2020).

Figure V.2 illustrates this empirical strategy: we use 1996-1998 wealth levels as the simplest predictors of wealth levels at the time of the reform. We define individuals with household taxable wealth above SEK 3,000K in 1996-1998 as high exposure to the reform, since they were already above the highest wealth tax exemption threshold then. Individuals with taxable wealth below the minimum wealth tax exemption threshold of SEK 800K in 1996-1998 are the low exposure group. To verify that those categories of wealth levels in 1996-1998 translate to differences in exposure to the 2007 reform, Panel A shows the changes in effective wealth tax rates faced by those taxpayers. The figure confirms that past levels of wealth (measured 10 years before the reform) predict differential exposure to the wealth tax reform of 2007. Panel B shows the corresponding out-migration patterns for the same taxpayers. Out-migration rates of taxpayers with high exposure to the reform dropped suddenly in 2007, compared to out-migration rates of taxpayers in the control group. Compared to our previous specification in Figure 3, we can see that those patterns last after 2008, and up to 2013. This confirms that the drop in out-migration rates for the very wealthy persisted several years after the repeal of the wealth tax.

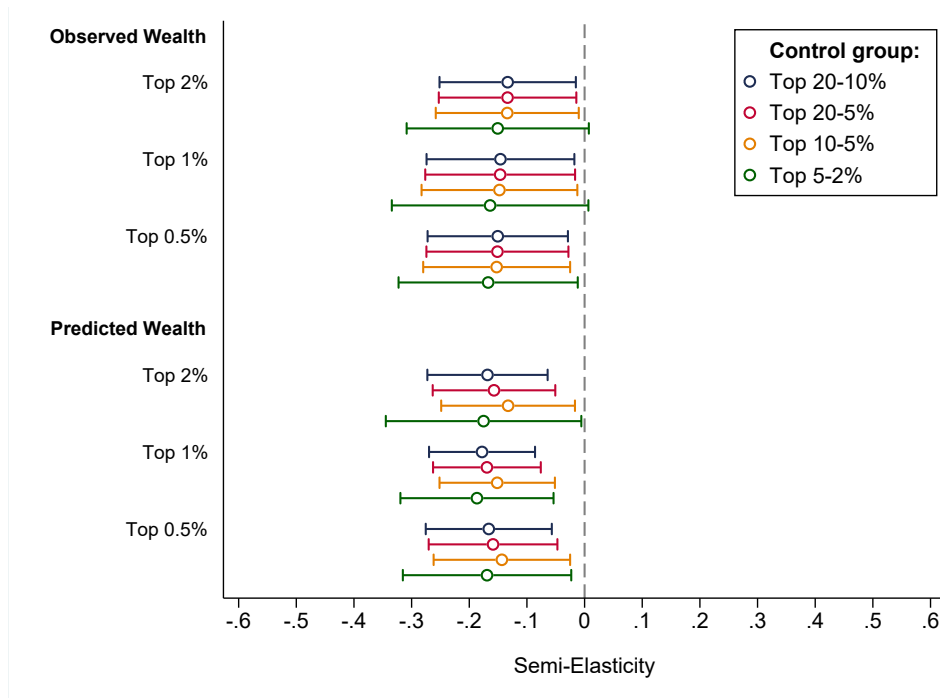
Our empirical approach in the paper (section 5) consists in generalizing this strategy, relying on the more elaborate prediction model of wealth described in Appendix II.1. Individuals are allocated to the treatment group if their predicted wealth belongs to the top 2%, and to the control group if their predicted wealth falls between the top 20% and top 10% of the distribution. Assigning treatment status according to predicted wealth (based on pre-reform variables) rather than actual wealth has an important empirical advantage: we avoid assigning treatment status based on a wealth variable that is endogenous to the current wealth tax level.

Figure V.2: Out-Migration Rates by Predicted Exposure to the Wealth Tax Reform



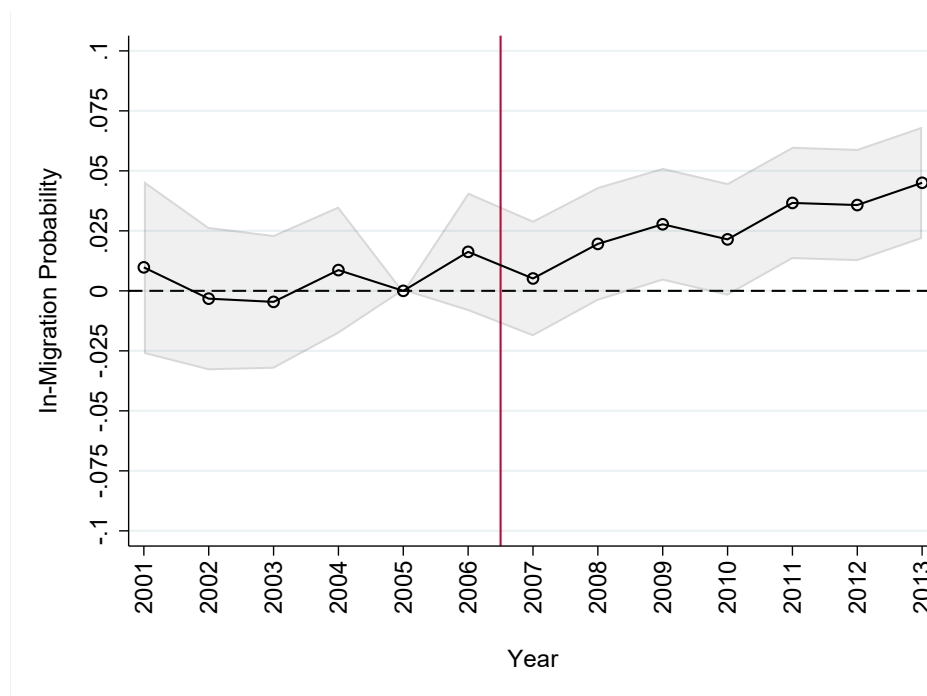
Notes: This figure shows the evolution of wealth tax rates (Panel A) and out-migration rates (Panel B) by exposure to the repeal of the wealth tax in Sweden, denoted by the vertical red line. We predict exposure to the reform by level of wealth in 1996-1998, comparing individuals above the highest exemption threshold during that period (3,000K, red) and individuals below the lowest exemption threshold (800K, blue).

Figure V.3: Sensitivity of Migration Semi-Elasticities To Different Control Groups



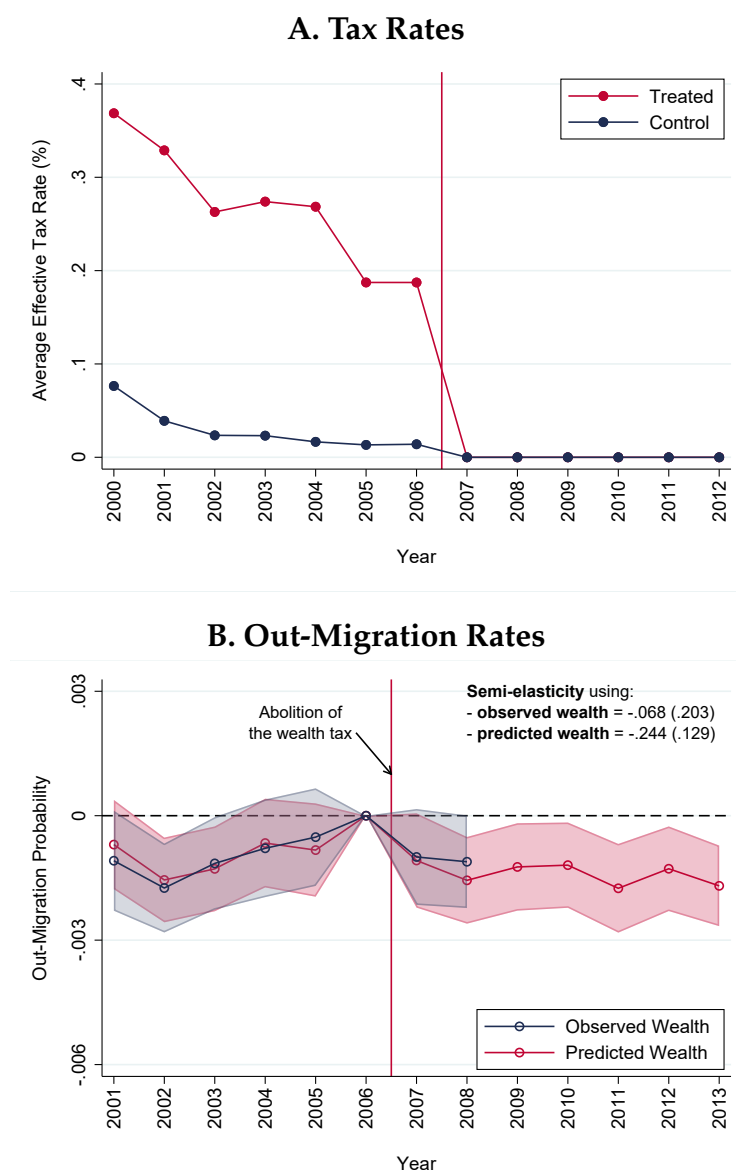
Notes: This figure tests the sensitivity of our migration semi-elasticities estimates to our choice of control group. The semi-elasticities plotted in the figure correspond to ε estimated from Equation 4 in the main text. We define treatment and control groups using observed level of wealth (top panel) or predicted level of wealth based on pre-reform assets and income flows (bottom panel). Each coefficient and its confidence intervals refer to one separate regression using the same treatment group (defined in left axis), but different wealth cut-offs for the control group. Our baseline and more conservative approach using the top 20-10% as the control group is shown in blue.

Figure V.4: Return Migration Rate Response To The Abolition Of The Wealth Tax: Difference-In-Differences Estimates



Notes: This figure shows the differential effects of the repeal of the wealth tax on the in-migration probability of treated (top 2% of the household net wealth distribution, subject to the wealth tax) and control (top 10-20%, not subject to the wealth tax) individuals. We define the treatment and control group in 1999. We regress the yearly probability to return in Sweden on an interaction between year fixed effects and a dummy variable equal to one if the individual was subject to the wealth tax in 1999. We plot the estimated coefficients β_j from Equation 3 in the main text and their 95 percent confidence intervals.

Figure V.5: **Out-Migration Rates by Predicted Exposure to the Wealth Tax Reform (Entrepreneurs Only)**



Notes: This figure shows the evolution of average effective tax rates (Panel A) and out-migration rates (Panel B) of entrepreneurs by exposure to the repeal of the wealth tax in Sweden, denoted by the vertical red line. In Panel A, the treatment and control groups are defined by entrepreneurs with predicted household wealth in the top 2% and 20-10%, respectively. Panel B uses the same groups, defined separately using predicted household wealth (red) and observed household wealth (blue). We plot the estimated coefficients β_j from Equation 3 in the main text and their 95 percent confidence intervals. The semi-elasticities reported in the figure are estimated from Equation 4 in the main text.

V.2. Out-of-Sample Validation: Migration Elasticity Estimates in Denmark

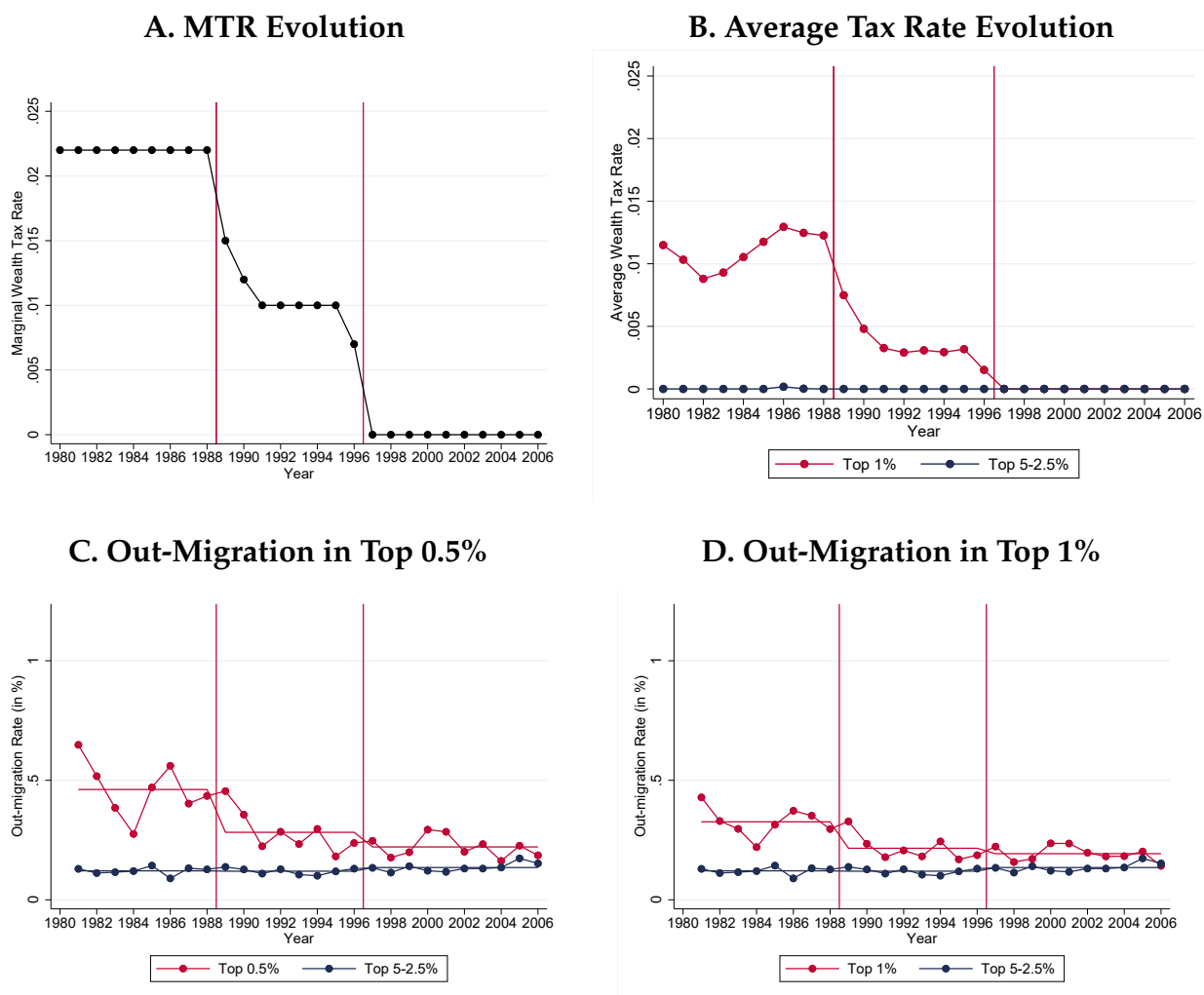
Net taxable wealth in Denmark was taxed every year at 2.2% until two major reforms. In 1988, the marginal tax rate on wealth above the exemption threshold was reduced to 1%. While the change in tax rate was large, this change was also gradual, and scattered over three years, as shown in Figure A.16, Panel A and B. The wealth tax was then entirely abolished between 1996 and 1997. Those two reforms provide potentially compelling identifying variation to study migration responses to the wealth tax.

We use the same identification strategy as with Sweden and compare out-migration patterns of wealthy tax payers just above the wealth tax exemption threshold, to those of taxpayers just below that threshold. The changes in marginal and effective wealth tax rates induced by the two Danish reforms are shown in Panel A and B of Figure V.6. The evolution of out-migration patterns around those changes for treated (red series) and control (blue series) individuals are shown in Panel C and D. Consistent with our main findings in Sweden, we find that reduced wealth tax rates are associated with a decrease in out-migration rates for taxpayers subject to the tax, with no corresponding changes for wealthy taxpayers not subject to the tax.

We note that the graphical evidence (from Figure V.6) is noisier in Denmark than in Sweden, mostly for two reasons. First, we have greater statistical power in Sweden because it is a larger country and it imposed wealth taxes on a larger share of taxpayers. Second, the Swedish wealth tax repeal offers a stronger research design for estimating migration responses: the tax changes were large, immediate, and unexpected. The Danish wealth tax reforms, on the other hand, were phased in gradually and across two reforms in close proximity to one another.

But as we show in the main text, the migration elasticities we obtain exploiting the Danish and Swedish reforms are very close, and not statistically different.

Figure V.6: Effect of the Wealth Tax Reform on Out-Migration in Denmark



Notes: This figure shows the out-migration patterns of wealthy taxpayers in Denmark around two large reforms in the wealth tax. The reforms generated two successive drops in marginal tax rates on wealth for wealthy taxpayers (Panel A) and drops in effective tax rates on wealth (Panel B). Panel C plots out-migration rates of the very wealthy (top 0.5%, red series) treated by the reforms and the wealthy not treated by the reforms (top 5-2.5%, blue series), before and after the two reforms, depicted by the two vertical red lines. Panel D plots out-migration rates of the wealthy (top 1%, red series) treated by the reforms and the wealthy not treated by the reforms (top 5-2.5%, blue series), before and after the two reforms, depicted by the two vertical red lines.

VI. Quantification of Aggregate Implications

In section 4 of our paper, we produce well-identified estimates of the semi-elasticities of migration flows with respect to the net-of-tax rate on wealth. Our evidence shows clear and significant effects on migration flows that appear immediately and are persistent. To draw policy implications, we need to translate these effects on migration flows into effects on the stock of the population of the very wealthy, and on the stock of taxable wealth.

Indeed, even though the estimated effects on migration flows are very small, one may wonder how large the effect on the stock may be if these small flow effects were to cumulate over time. This begs the question: how exactly and for how long should one cumulate these flow effects in order to properly measure the stock effect? It is clear that one needs to cumulate these flow effects somehow. But it is also clear that one cannot cumulate these flow effects infinitely. The latter would imply that the composition of the population of wealthy individuals can be considered as fixed: if a rich individual leaves, he is never replaced. Which would quite trivially imply that any significant constant negative effects on migration flows would inevitably drive the population of wealthy individuals to shrink to zero in the long run. In practice, a fraction of individuals constantly disappear from the population of wealthy individuals, through death, or because of wealth destruction, while a fraction of new individuals appear through wealth creation or inheritances. These constant inflows and outflows through birth and death, and through the creation, destruction, and transmission of wealth means that the composition of the population of the wealthy is not fixed, but constantly changing. The impact of migration flows on the stock of the population in the long run will therefore depend on the the relative magnitude of these different flows in and out of the population of wealthy individuals, which we can measure in the data.

Here, we explain how we use our data and the structure of a simple model to understand and measure precisely how these forces play out, in order to obtain a proper quantification of the stock elasticity.

VI.1. A Simple Framework Accounting for the Dynamics of the Population of Wealthy Individuals

Our framework accounts for the various forces that shape the dynamics of the local population of wealthy individuals. First, individuals can enter the population of wealthy individuals due to increases in their wealth. For now, we do not separate whether these increases come from wealth accumulation or from inheritances, but we come back to this

important distinction later. Second, individuals can exit the population of wealthy individuals due to wealth destruction or death. Finally, individuals can enter or exit the local population of wealthy individuals through migration.

We start with an OLG framework that provides the simplest formalization of these different forces. In each time period, indexed by t , a new group of individuals B_t is “born” into the population of wealthy individuals in Sweden. They “die” after T periods. In between their birth and their death, individuals can move in or out of the country. We define as the “age” of an individual the number of periods k since that individual was born in the wealthy population. At each time t , the total number of individuals of age k (N_t^k) present in Sweden corresponds to the number of individuals who are in Sweden and who were born in the wealthy population k periods ago.

$$\begin{aligned}
\text{Population of age 0 in } t & N_t^0 = B_t \\
\text{Population of age 1 in } t & N_t^1 = B_{t-1}(1 - \alpha_t^1) \\
\text{Population of age 2 in } t & N_t^2 = B_{t-2}(1 - \alpha_t^2)(1 - \alpha_{t-1}^1) \\
& \dots \\
\text{Population of age } k \text{ in } t & N_t^k = B_{t-k} \prod_{j=0}^{k-1} (1 - \alpha_{t-j}^{k-j}) = B_{t-k} S_{t-k}(k)
\end{aligned}$$

where α_t^k is the net out-migration rate of individuals of age k in period t and $S_{t-k}(k)$ is the “survival rate” in Sweden at age k of individuals born in $t - k$.

At each period t , the total population N_t is simply the sum over all ages k of the number N_t^k of individuals who have been in the wealthy population for k number of periods. The total local stock of population of wealthy individuals is therefore $N_t = \sum_{k=0}^T N_t^k$.

VI.2. Microfoundations of Migration Responses

We are interested in how wealth taxation affects migration flow rates α_t^k of the very wealthy, and as a consequence, the local population stock of wealthy households. To get a microfoundation of migration behaviors, we rely on a random utility model. Individuals form utility over being in location $l \in \{S; O\}$, where S stands for Sweden, and O is the individual’s next best location alternative. The utility of an individual of age k in location l in period t depends on her consumption opportunities in l , as well as some idiosyncratic location taste shock $\mu_{l,t}^k$. Consumption opportunities are simply her net-of-tax

labor income and net-of-tax wealth in location l .

$$\mathbb{U}_{l,t} = \underbrace{u((1 - \mathbf{t}_{l,t})y_{l,t}^k + (1 - \tau_{l,t})W_{l,t}^k)}_{u_{l,t}^k} + \mu_{l,t}^k$$

where $y_{l,t}^k$ and $\mathbf{t}_{l,t}$ are respectively earnings and taxes on earnings in location l ; W is wealth, and τ is the effective tax rate on wealth. Individuals of age k for whom $\Delta\mu_t^k = \mu_{O,t}^k - \mu_{S,t}^k < -\Delta u_t^k = u_{O,t}^k - u_{S,t}^k$ leave Sweden (or stay out of Sweden), while individuals move into Sweden (or stay in Sweden) when the opposite is true. Wealth taxation affects location decisions by affecting the threshold level of utility Δu^k below which individuals decide to be in or out of Sweden. We assume that idiosyncratic shocks $\mu_{l,t}^k$ are independent, and taken from the same distribution $F(\cdot)$ every period. These shocks generate constant expected migration flow rates in and out of Sweden every period when earnings and taxes are stationary (i.e. when $\Delta u_t^k = \Delta u_{t-1}^k, \forall t$). We define the net expected out-migration flow rate α_t^k as the expected rate of out-migration minus the expected rate of in-migration into Sweden of individuals of age k . Based on our assumptions, α_t^k are constant over time ($\alpha_t^k = \alpha_{t-1}^k, \forall t$) when earnings and taxes are stationary.

VI.3. Impact of Wealth Tax Changes on the Steady State Population Stock

Our framework focuses on the quantification of the extensive margin effects of wealth taxation. We therefore abstract away from effects of the wealth tax on the birth rate into the wealthy population through intensive margin effects. We consider a simple version of the steady state population where birth rates B_t are constant over time ($B_t = B, \forall t$).

At the steady state, the population stock N is therefore given by:

$$N = B \cdot \sum_{k=0}^T \prod_{j=0}^k (1 - \alpha^{k-j}) = B \cdot \sum_{k=0}^T S(k) = B \cdot D \quad (8)$$

where D is the average time that a wealthy individual spends living in Sweden over her lifespan.

We are now considering the impact of change in the net-of-tax rate on wealth in Sweden $d(1 - \tau_S)$. For convenience, we will be dropping the subscript S for the tax rate going forward, and always refer to the tax rate on wealth *in Sweden* when writing τ .

The elasticity of the population stock N with respect to the net-of-tax rate on wealth in Sweden is:

$$\frac{dN/N}{d(1-\tau)/(1-\tau)} = \frac{B}{N} \cdot \sum_{k=0}^T \frac{\partial(\prod_{j=0}^k (1-\alpha^{k-j}))}{\partial(1-\tau)/(1-\tau)} \quad (9)$$

We define $\xi^j = \frac{\partial(1-\alpha^j)}{\partial(1-\tau)} \cdot \frac{1-\tau}{1-\alpha^j}$. This is the elasticity of the instantaneous probability of remaining in Sweden w.r.t the net-of-tax rate on wealth.

Note that because α^j is small, $1-\alpha^j \approx 1$. This means that $\xi^j \approx \frac{-\partial\alpha^j}{\partial(1-\tau)/(1-\tau)} = -\varepsilon^j$.

In other words, an empirical analog of ξ^j is the opposite of the semi-elasticity of the net-migration rate w.r.t the net-of-tax rate on wealth ε^j . The latter is the semi-elasticity that we estimate in specification (4) in section 5 of the paper.

We now decompose $\frac{dN/N}{d(1-\tau)/(1-\tau)}$ into the effect of the tax change on the population of each age k :

Percentage effect on population of age 0	$\frac{dN^0/N^0}{d(1-\tau)/(1-\tau)} = 0$
Percentage effect on population of age 1	$\frac{dN^1/N^1}{d(1-\tau)/(1-\tau)} = \frac{\partial(1-\alpha^1)}{\partial(1-\tau)} \cdot \frac{1-\tau}{1-\alpha^1} = \xi^1$
Percentage effect on population of age 2	$\frac{dN^2/N^2}{d(1-\tau)/(1-\tau)} = \xi^1 + \xi^2$
	...
Percentage effect on population of age k	$\frac{dN^k/N^k}{d(1-\tau)/(1-\tau)} = \sum_{j=1}^k \xi^j$

This allows us to rewrite equation (9) as:

$$\frac{dN/N}{d(1-\tau)/(1-\tau)} = \sum_{k=0}^T \frac{S(k)}{D} \cdot \sum_{j=1}^k \xi^j \quad (10)$$

Discussion Equation (10) shows that the effects of a change in the wealth tax on the net flows of migration do cumulate over time. The percentage effect on the population of age 1 is $\xi^1 \approx -\varepsilon^1$, the percentage effect on the population of age 2 is $\xi^1 + \xi^2 \approx -\varepsilon^1 - \varepsilon^2$, etc. This is quite intuitive: if individuals leave at age k , they are not available in the

population of wealthy individuals in Sweden for the subsequent periods of their lives. As a consequence, it matters if the effects on out-migration are concentrated on younger or older individuals. If younger individuals move out of Sweden, this will tend to have a larger effect on the population stock in the steady-state, because they are going to be missing for all the subsequent periods when they could have been alive and in Sweden. If older individuals tend to move instead, this will have a smaller effect on the population stock.

The interest of formula (10) is that it connects easily to our data and empirical setting. Indeed, we can use the fact that $\xi^j \approx -\varepsilon^j$ and rewrite:

$$\frac{dN/N}{d(1-\tau)/(1-\tau)} \approx \sum_{k=0}^T \frac{S(k)}{D} \cdot \sum_{j=1}^k -\varepsilon^j \quad (11)$$

We can thus provide a direct estimate of the elasticity of the population stock based solely on measures of D and S which we can observe in the data, and based on our estimates of the effects of the wealth tax on the net out-migration flows at age k , ε^j . These are the estimates that we get from a regression like specification (4) in section 5 of the paper. In Figure 5 of the paper, we show results where we implemented this specification for the net-migration rate α of the population on average, and also results where we look at heterogeneity in the migration rate effect ε^k by age k .

Simplified Formula and Intuition Results in section 5 of the paper suggest that semi-elasticities ε^k are approximately constant. That is, the semi-elasticity of the net-migration rate α w.r.t to the net-of-tax rate does not seem to vary with age k . Using this, and the fact that $S/D \approx 1/T$ we can get a simplified formula for the effect of a change in the wealth tax rate on the size of the population of wealthy individual:

$$\frac{dN/N}{d(1-\tau)/(1-\tau)} \approx -\varepsilon \cdot \frac{(T+1)}{2} \quad (12)$$

This formula can be implemented simply with our estimate of the average semi-elasticity of net out-migration flows $\hat{\varepsilon}$ from specification (4) in section 5 and a measure of T , the average “lifespan” of a wealthy individual. Note that T in the steady-state is directly related to the birth rate of individuals into the population of wealthy individuals: $B = 1/T$. The longer individuals’ lifespan in the wealthy population, the lower the net birth rate in

the population of wealthy individuals, i.e. the lower the rate at which the population of wealthy individuals regenerates.

Formula (12) has a simple interpretation: to get an estimate of the effect on the population stock N of wealthy individuals, we simply need to cumulate the flow effect ε for the half-life that individuals spend in the wealthy population. The larger the average lifespan T , the larger the effect on the stock. This is because a larger T implies a lower regeneration rate of the wealthy population absent migration. So when we lose a wealthy individual to migration, it is harder to replace her.

Calibrations We start by implementing our simplified formula. To calibrate T , we measure in our data the average time spent in the top 2% of the wealth distribution. We find that on average wealthy individuals spend 15.1 years of their lives in the top 2% of the wealth distribution. We check empirically that this estimate of T matches the hazard rates in and out of the top 2% of the wealth distribution that we observe every year. These estimates imply that around 30% of the population of the top 2% is replaced every five years. This is due to the steep lifecycle profile of wealth. For ε (i.e. the average semi-elasticity of net out-migration flows), we take the sum of the out-migration and in-migration semi-elasticities from Figure 4 in the main text.⁶

Simplified Formula: Calibration 1 *The estimated elasticity of the stock of the population of wealthy individuals with respect to the net-of-tax rate on wealth is:*

$$\frac{dN/N}{d(1-\tau)/(1-\tau)} = 1.77 \quad (0.47)$$

Note that because τ is small, $1 - \tau \approx 1$, and so $d(1 - \tau)/(1 - \tau) \approx -d\tau$. As a result, the elasticity can also be interpreted as the percentage change in the size of the wealthy population when the effective tax rate on wealth is increased by 1 percentage point.

This estimate confirms that, even when we properly cumulate the very small flow effects to account for the steady state effect on the stock of the population, migration effects remain extremely modest. In the long run, a 1 percentage point increase in the effective tax rate on wealth would only decrease the size of the population of the very wealthy in Sweden by 1.77 % due to migration responses. One should note that a 1 percentage point increase in the effective tax rate on wealth is a large increase. The abolition of the wealth tax in Sweden, which is one of the largest wealth tax reforms ever studied amounted to a

⁶Note that for Sweden, we do not have an estimate of the semi-elasticity of in-migration flows. Given the estimates for out-migration flows are very similar for Sweden and Denmark, we therefore take the Danish estimate of the semi-elasticity of in-migration flows.

0.5 percentage point decrease in the effective tax rate. In other words, the abolition of the Swedish wealth tax only increased the long run population of the wealthy in Sweden by about 1%, which is small.

We can also calibrate our exact formula (10) using estimates of semi-elasticity of migration flows ε^k by age. We find an estimated elasticity of population size of 1.88 (1.26), which is almost identical to our simplified formula.

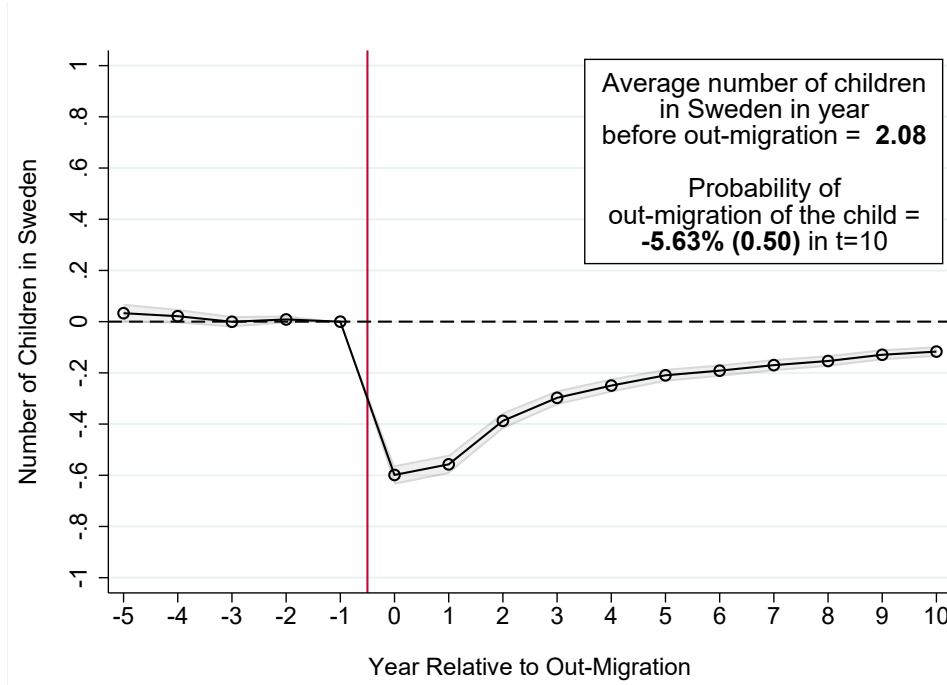
Dynastic effects Our approach so far assumes that there is no effect of migration on the birth rate into the wealthy population. In practice though, individuals who have had wealthy parents are more likely to become wealthy. This creates a dynastic connection between wealthy individuals today and wealthy individuals tomorrow. And if children of wealthy parents are also out-migrating with their parents, out-migration responses today may also affect the birth rate into the wealthy population, potentially increasing the effect on population size. Accounting for these dynastic connections, the formula for the semi-elasticity becomes:

$$\frac{dN/N}{d(1-\tau)/(1-\tau)} = \frac{\partial B}{\partial(1-\tau)} \cdot \frac{1-\tau}{B} + \sum_{k=0}^T \frac{S(k)}{D} \cdot \sum_{j=1}^k -\varepsilon^j \quad (13)$$

where, compared to our baseline formula, we now need to add the elasticity of the birth rate into the wealthy population with respect to the net-of-tax rate on wealth. The size of this elasticity will depend on the importance of these dynastic effects: if all new wealthy individuals are heirs of wealthy parents, and if heirs systematically outmigrate with their parents, then this could significantly increase the impact of wealth taxes on the population size.

Importantly, we can directly calibrate formula (13) and assess how these dynastic effects affect our baseline results. First, we start by estimating the probability that children of wealthy parents out-migrate when their parents do so. In Figure VI.1 below, we report the results of an event study similar to specification (7) in the main text. The event is out-migration of a wealthy individual in the top 2% of the wealth distribution. The outcome is a dummy equal to one if one of her children is living in Sweden. The graph shows precisely estimated but small effects of parental migration parents on the out-migration probability of their children. Ten years after her wealthy parent has out-migrated, the probability that a child also remains out of Sweden increases by just 5%.

Figure VI.1: Event Study: Probability for Children of Wealthy Parents to Live in Sweden Around the Event of Their Parents' Out-Migration



Notes: This figure shows the probability for children of wealthy parents (in the top 2% of household net wealth distribution) to live in Sweden around the event of their parents' out-migration. We plot the coefficients β_j (and their confidence intervals) estimated from Equation 7 in the main text, where the outcome is the number of children residing in Sweden in a given year. The effect displayed in the text boxes is the estimate of β_{10} divided by the average outcome in the treatment group in $t = -1$, multiplied by 100.

To get an estimate of the impact of these dynastic effects on birth rates, we now simply need to multiply the migration elasticity for the population of parents by the estimate of the impact of parental migration on children migration.

We find a total effect on birth rates into the wealthy population of $\frac{\partial B}{\partial(1-\tau)} \cdot \frac{1-\tau}{B} = .001$. And calibrating formula (13), we find an estimated elasticity of population size of 1.87 (.50), almost identical to our original estimate. In other words, although we can clearly detect the presence of dynastic migration effects, these effects are very small and do not affect significantly our baseline estimate of the elasticity of population size with-respect-to wealth taxation.

Robustness to Tax Evasion One potential concern with our estimates of migration responses to the wealth tax is that we may have mismeasurement in the denominator $d\tau$ due to tax evasion. For instance, offshoring wealth in tax havens has been shown to be

a significant driver of tax evasion by the very wealthy (e.g. [Boas et al. \(2024\)](#)). In the presence of tax evasion, the actual net wealth of top taxpayers is underestimated in the administrative data by a factor $(1 - e)$, where e is the fraction of wealth that is evaded. And as a result, our measure of tax rates τ is an overestimate of their effective tax rates on wealth $\tilde{\tau} = (1 - e)\tau$. A corollary is that our estimated elasticities may overestimate the true elasticity with respect to the effective net-of-tax rate $(1 - \tilde{\tau})$:

$$\frac{dN/N}{d(1 - \tilde{\tau})/(1 - \tilde{\tau})} = \frac{dN/N}{d(1 - \tau)/(1 - \tau)} \cdot \frac{1}{(1 - e)}$$

We can nevertheless easily explore the sensitivity of our estimates to the extent of tax evasion by using direct estimates of the fraction of wealth evaded by top wealth groups in Sweden from [Boas et al. \(2024\)](#). In their paper, they provide an upper bound and a lower bound on the fraction of wealth e evaded by each top fractile of wealth.⁷ Using these estimates, we compute a lower bound and an upper bound on the total fraction of wealth evaded by the top 2% of wealthy taxpayers, and provide in panel A of Figure 5 two bounds for our estimates of the elasticity of the stock of the population of the wealthy accounting for tax evasion. The upper bound elasticity is 1.92 and the lower bound elasticity is 1.85, indicating that accounting for the presence of tax evasion does not affect the main qualitative message of our baseline results, namely that the impact of wealth taxes on the size of the population of the wealthy is small.⁸

Elasticity w.r.t Net-of-Tax Rate on Capital Income While our estimated elasticities of population with respect to the net-of-tax rate on wealth have an intuitive interpretation, their magnitude can be hard to compare to existing estimates of migration elasticities, which are typically expressed with respect to the net-of-tax rate on income. It is however easy to convert our estimate into an elasticity of population with respect to the net-of-tax rate on capital income. For this, we simply compute the change in capital income taxation induced by the wealth tax. To do this, we need to compute the implied capital income that wealth taxpayers receive each year out of their total net wealth and add it to the tax rate t^K they pay on their other capital income. We define $t \approx \tau/r + t^K$ the average tax rate on capital income. Over our period of study, we observe in our data $r = .042$ and $\tau \approx .006$ which translates into $t = 14.3\% + 20\%$. Therefore, we obtain:

⁷We use Table J3 Sweden of their online appendix for the upper bound scenario, Table J3.B Sweden for their average scenario, and Table J3.C Sweden for their lower bound scenario.

⁸As we pointed in [Kleven et al. \(2020\)](#), the availability of tax evasion opportunities through wealth offshoring may actually structurally affect migration elasticities downwards, as it makes it cheaper to move one's wealth rather than residence to avoid the burden of taxation.

$$\frac{dN/N}{d(1-t)/(1-t)} = \frac{dN/N}{d(1-\tau)/(1-\tau)} \cdot \frac{d \ln(1-\tau)}{d \ln(1-t)} \approx .049 \quad (.013) \quad (14)$$

When appropriately rescaled, the implied migration elasticity is thus small. Panel B of Figure 5 compares our estimates to migration elasticities available in the literature, which come from two strands of papers. The first focuses on migration elasticities of top income earners, the second, much less developed, investigates migration responses to capital taxation, but relies exclusively on intra-national variation across local jurisdictions. Two insights emerge from the comparison. First, our estimates accord in magnitude to cross-border migration elasticities of top incomes. These elasticities are typically found to be quite small, around .1, except when focusing on specific subsegments of the labor force such as foreign nationals and expatriates. Second, our elasticity is substantially smaller than migration elasticities to capital taxation obtained by [Agrawal et al. \(2023\)](#) or [Brülhart et al. \(2022\)](#), who investigate intra-national migration of wealthy taxpayers in Spain and Switzerland respectively. This is consistent with the fact that we focus on international migration responses, rather than within-country moves, and that we measure real relocation responses, rather than avoidance through artificial changes in reporting, for instance.

VI.4. Aggregate External Effects on Employment and Economic Activity

We explain here how we combine our estimates of migration responses to the wealth tax with our estimates of the persistent impact of migration on individual and firm-level outcomes to quantify the aggregate effects of wealth taxation on employment, investment and business dynamism.

VI.4.1. Formula & Heterogeneous Treatment Effects of Migration

Our approach is based on a simple decomposition of the effect of wealth tax-induced migration on any aggregate outcome Y into three terms:

$$\frac{dY/Y}{d(1-\tau)/(1-\tau)} = \underbrace{\frac{\partial N}{\partial(1-\tau)} \cdot \frac{1-\tau}{N}}_{\text{Migration Elasticity}} \times \underbrace{\frac{\partial Y}{\partial N} \cdot \frac{1}{Y^w}}_{\text{Migration Impact}} \times \underbrace{\frac{NY^w}{Y}}_{\text{Top Wealth Share of } Y} \quad (15)$$

where Y^w is the average outcome (e.g., employment, investment or tax payments)

generated by wealthy entrepreneurs and N is the population of wealthy individuals. The first term in the decomposition (the *migration elasticity*), captures the impact of the change in the average wealth tax-rate on the overall stock of wealthy individuals via migration responses. The first step of our methodology, presented in Section 5, consists in estimating this migration elasticity using wealth tax reforms.

The second step consists in estimating the impacts of migration events on economic outcomes, the second term of the formula. Importantly, this approach, presented in Section 6, does not rely directly on any wealth tax experiment. Specifically, using an event study design, we estimate the impacts of out-migration on individual-level, firm-level, and market-level outcomes. This approach is statistically precise because it allows us to pool a large number of individual migration events over a relatively long time period. The event study design also relies on transparent identification assumptions, some of which can be easily tested in the data.

We can then measure the aggregate economic implications of tax-induced migration by the wealthy by combining our estimates of (i) the migration elasticity (step 1), and (ii) the impact of migration on various economic outcomes (step 2). When combining these estimates, we simply need to account for the share that the wealthy represent in the aggregate outcome Y (the third term in decomposition (1)). Looking at a range of key outcomes, our approach allows to shed light on the potential trickle-down effects of taxing high-wealth individuals. We do so in Section 7 of the paper.

Our approach combines estimates from two different designs: tax reforms for the migration elasticity and event studies for the impact of migration. To quantify the aggregate economic effects of wealth tax-induced migration events, we need a LATE in the latter design that corresponds to the same target population as the LATE from the former design. Our two-step procedure thus requires that the LATE for migration impacts are based on the same population as the LATE for migration elasticities. That is, we need to identify the effects of migration for the group of individuals m who are *at the margin* of deciding to migrate when the wealth tax is abolished:

$$\frac{\partial Y}{\partial N} \cdot \frac{1}{Y^w} = \underbrace{\frac{\partial Y}{\partial N} \Big|_m \cdot \frac{1}{Y^m}}_{\text{Migration Impact for Marginals}} \times \underbrace{\frac{Y^m}{Y^w}}_{\text{Selection into Migration}} \quad (16)$$

In the absence of treatment effect heterogeneity, things are straightforward. But in the presence of heterogeneity, event studies based on all migration events may not give the relevant LATE for estimating the economic effects of migration responses to the wealth tax repeal.

For this reason, we now probe into the presence of heterogeneous treatment effects.

Selection Into Tax-Induced Migration We apply standard methods used in the literature on selection in insurance markets (e.g. [Hendren et al., 2021](#)). We define marginal individuals as the group of compliers in the wealth tax reform, i.e. individuals who would have moved before the reform, and stopped moving after reform.

We can first look at the observable characteristics of these compliers. To do so, in [Table VI.1](#), we start by comparing the average characteristics of wealthy migrants before (column (2)) vs after the reform (column (3)). Results reveal that these two groups of migrants are extremely similar (panel A) and that the firms they own are also remarkably comparable (panel B). In column (4) we formally test for the null of equality of characteristics between pre- and post-abolition migrants. t -statistics suggest that pre-abolition migrants are a bit younger, a bit wealthier and slightly more likely to own a closely-held business. But apart from that, we cannot reject that they are the same as post-abolition migrants, neither can we reject that the characteristics of their firms are the same.

From these comparisons, we can then formally retrieve the average characteristics Y^m of compliers. The methodology is simple. The average characteristics of pre-abolition wealthy migrants is a weighted average of the characteristics of the compliers and of infra-marginal migrants: $\omega Y^m + (1 - \omega)Y^i$. While the average characteristics of infra-marginal migrants Y^i corresponds to the post-abolition average. By plugging the fraction of compliers $\omega = .3$ in pre-abolition migration flows (estimated in [section 5](#)), we can then easily estimate Y^m , which we report in column (5) along with standard errors computed with the Delta-method.⁹

The average migrant induced to move because of the wealth tax is 46 years old, born in Sweden, has tertiary education, and is around the 65-th to 70-th percentile of the distribution of cognitive and non-cognitive skills. To gauge the extent of selection into tax-induced migration, it is useful to compare compliers to the average characteristics of wealthy individuals Y^w from column (1). We find compliers to be a bit younger, a bit more likely to be entrepreneurs, and their net wealth to be slightly larger.¹⁰ But as far as firms are concerned, we find no evidence of selection into migration. In other words, firms owned by compliers are not different from firms owned by all wealthy individuals.

Treatment Effects of Migration for Tax Reform Compliers The analysis of compliers suggests that movers are very similar before vs after the repeal of the wealth tax. While

⁹For firm outcomes we use the fraction of compliers among entrepreneurs, 0.51.

¹⁰One implication is that the elasticity of the wealth tax base (i.e. the elasticity of migration flows weighted by wealth level) is -.30 (.07), a bit larger than the flow migration elasticity.

this should alleviate concerns that treatment effects of migration are strongly different among compliers, we now test for treatment effect heterogeneity more formally.

In the main text, in Table 2, we compare migration impacts estimated on events happening before and after the repeal of the wealth tax. The table shows that the event study estimates of out-migration are similar before and after the repeal of the wealth tax. For almost all outcomes, the effects of out-migration are slightly larger, but not statistically different. This implies that the out-migration effects for those who move because of wealth taxes are not statistically different from the effects of other moves. Following the same strategy as in Table 2, we then estimate the treatment effect of migration for compliers $\frac{\partial Y}{\partial N}\big|_m \cdot \frac{1}{Y^m}$, which we report in column (4) of Appendix Table 2.

To compute this LATE, we proceed as follows. To start, we know that the estimate in the pre-abolition period β^{pre} is a weighted average of β^m , the treatment effect of migration for individuals who migrate because of the wealth tax (the “marginals”), and β^i , the effect for individuals who would migrate irrespective of the wealth tax (“infra-marginals”). The respective weights are simply ω and $(1-\omega)$, where ω is the fraction of migrants who migrate because of the wealth tax, so

$$\beta^{pre} = \omega \beta^m + (1 - \omega) \beta^i$$

After the repeal of the wealth tax, the estimate β^{post} captures β^i directly, the effect for the “infra-marginals” only, i.e.

$$\beta^{post} = \beta^i$$

Note that we have identified ω in section 5 of the paper: estimating the causal impact of the abolition of the wealth tax on migration flows, we found that one third of flows disappeared after the repeal of the wealth tax (i.e. $\omega = .30$)¹¹. Our estimates of ω , β^{pre} and β^{post} allow us to back-out migration effects for “marginals”, as re-arranging the above gives

$$\beta^m = (\beta^{pre} - (1 - \omega) \beta^{post}) * \omega^{-1} \quad (17)$$

So far, our approach assumes that β^m and β^i are the same in the pre-abolition and post-abolition periods. We can relax this assumption, and compute an alternative LATE, which allows for aggregate dynamics in the treatment effects. To do so, we additionally estimate a β^{pre} and β^{post} for our control group (individuals in the top 20-10% of net wealth).

¹¹Replicating this analysis focusing on wealthy entrepreneurs only, we find $\omega = .51$. We use this estimate when computing the LATE for the ‘Firm-level’ outcomes and characteristics in Table 2 and Table VI.1.

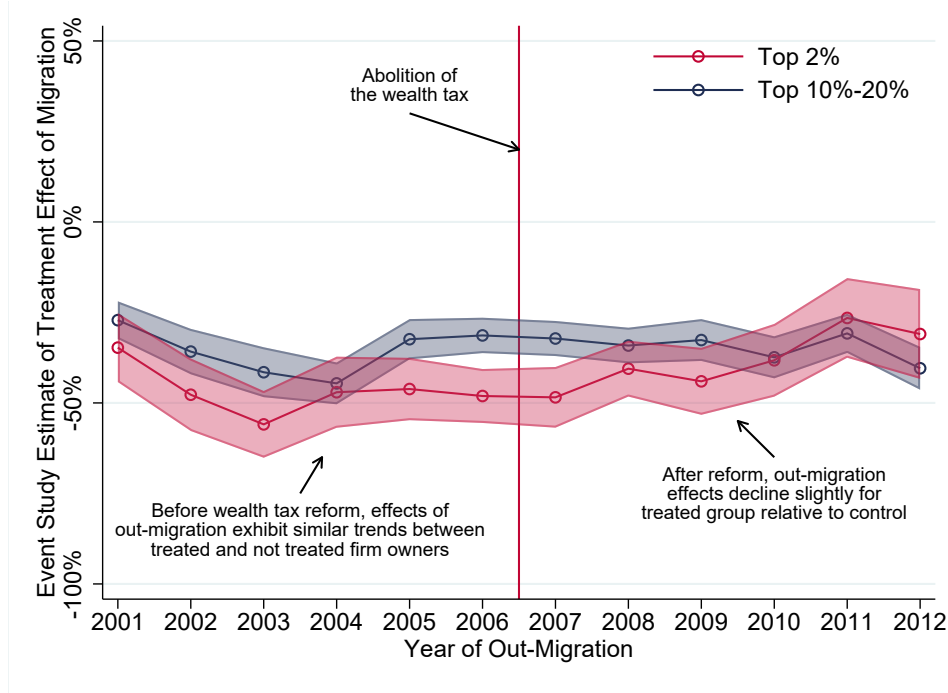
Since these out-migrants are not affected by the abolition, there are no “marginals”. If there are aggregate dynamics in treatment effects, they will be captured by these estimates. As such, we can estimate an event-study version of a Difference-in-Difference-in-Difference-in-Differences specification: the difference-in-differences in treatment effects pre-abolition and post-abolition, for the treatment group versus the control group. Specifically, we estimate the following DDDD version of specification (7):

$$\begin{aligned}
Y_{it} = & \sum_{j \neq -1} \beta_j^{DDDD} \cdot \mathbb{1}[M_i = 1] \cdot \mathbb{1}[t = j] \cdot \mathbb{1}[P_i = 1] \cdot \mathbb{1}[T_i = 1] + \\
& \sum_{j \neq -1} \gamma_j^1 \cdot \mathbb{1}[M_i = 1] \cdot \mathbb{1}[t = j] \cdot \mathbb{1}[P_i = 1] + \dots + \\
& \sum_{j \neq -1} \delta_j^1 \cdot \mathbb{1}[M_i = 1] \cdot \mathbb{1}[t = j] + \dots + \\
& \sum_{j \neq -1} \theta_j^1 \cdot \mathbb{1}[t = j] + \dots + \varepsilon_{it}
\end{aligned} \tag{18}$$

where P_i is a dummy indicating if an out-migration happens post-abolition and T_i is a dummy indicating if an out-migrant is in the treatment group. All the lower-order interaction terms are included in the regression. Here, β^{DDDD} gives us the difference in treatment effects post-abolition, after subtracting overall dynamics in treatment effects.

Appendix Figure VI.2 provides a simple graphical illustration of this double difference approach, focusing on one outcome, total tax payments. This graphical illustration enables us to test for the validity of the parallel trend assumptions, at the heart of identification. For each year, we plot the coefficient estimates of the event study for all migrations happening in that year. We do this separately for individuals belonging to the treatment group (i.e. in to the top 2% of wealth, in red) and for individuals belonging to the control group (i.e. the top 20 to top 10% of wealth, in blue). We find that the effects of migration are quite stable over time in the pre-abolition period, and evolve similarly for the treated and control group of migrants. After the abolition, we observe a small decline in the magnitude of the effects of migration among the treated group, while the estimates remain unchanged for control migrants. The *double difference* gives therefore very comparable results to the simple difference approach. The LATE estimates for compliers derived from the double difference approach are reported in column (5) of Table 2, and are very similar to the estimates from column (4). In the next section, we use column (5) as our preferred estimates of migration impacts among the wealth tax compliers.

Figure VI.2: Event-study estimates of the impact of migration on total tax payments, by year of migration, for the treatment and control group in the wealth tax reform natural experiment



Notes: This figure reports the coefficient estimates of the event study for all migrations happening in that year. We focus on individual total tax payments. We do this separately for individuals belonging to the treatment group (i.e. in to the top 2% of wealth, in red) and for individuals belonging to the control group (i.e. the top 20 to top 10% of wealth, in blue).

We can then compute the LATE as

$$\beta^{\tilde{m}} = (\beta^{pre} - (1 - \omega) (\beta^{pre} + \beta^{DDDD})) * \omega^{-1} \quad (19)$$

This is the second LATE that we report in column (5) of Table 2 in the main text.

We find slightly larger LATE for our population of compliers than for the whole population of migrants pre-2007, although we cannot rule out homogenous treatment effects.

Table 2 conveys three important insights:

- (i) For almost all outcomes, we cannot detect statistically significant differences between the effects pre and post abolition. For three out of the 9 outcomes the difference is statistically significant at the 5% level.

- (ii) This in turn implies that for almost outcomes, we cannot reject that migration effects are the same for “marginals” and “non-marginals”.
- (iii) For most outcomes however, the point estimates for pre-abolition effects tend to be slightly larger than for post-abolition. As a result, the point estimates of the LATE for our marginals tend to be somewhat larger than the estimates for the pre-abolition period. Even though these LATE estimates are a bit imprecise, we use them as our baseline to calibrate equation 15.

Table VI.1: Test of Equality of Out-Migration Characteristics Before and After the Repeal of the Wealth Tax

Characteristic	Top 2% Out-Migrants			T-Statistic	Compliers (Marginal Migrants)
	Top 2%	Pre-Abolition	Post-Abolition		
	(1)	(2)	(3)	(4)	(5)
A. Mean Individual Characteristics					
<i>Sample of individuals in the top 2% of net worth</i>					
Age	61.89 (0.02)	48.98 (0.24)	50.23 (0.25)	-3.65***	46.06 (0.98)
Higher Education (%)	43.60 (0.09)	61.90 (0.94)	65.86 (0.96)	-2.95***	52.65 (3.84)
Cognitive Skills Percentile	50.17 (0.14)	63.11 (1.09)	60.34 (1.02)	1.86*	69.58 (4.34)
Non-Cog. Skills Percentile	50.16 (0.15)	61.49 (1.16)	60.42 (1.09)	0.67	63.97 (4.64)
Foreign Born (%)	1.62 (0.02)	10.20 (0.58)	11.38 (0.64)	-1.35	7.46 (2.46)
Owns Active CHB (%)	7.08 (0.05)	8.98 (0.55)	6.39 (0.49)	3.50***	15.02 (2.17)
Wealth Rank Within Top 2%	50.00 (0.05)	50.28 (0.68)	48.35 (0.67)	2.02**	54.78 (2.75)
B. Mean Firm Characteristics					
<i>Sample of firms with owner in the top 2% of net worth</i>					
Number of Employees	22.57 (0.31)	22.81 (1.00)	22.55 (1.26)	0.16	23.06 (2.28)
Value Added (SEK 1,000)	10,340.92 (162.71)	11,698.29 (554.90)	11,185.00 (707.21)	0.57	12,183.62 (1,269.89)
Net Turnover (SEK 1,000)	38,690.96 (711.61)	41,877.01 (1,994.41)	40,315.60 (2,490.15)	0.49	43,353.36 (4,538.66)
Tax Payments (SEK 1,000)	501.58 (14.38)	434.88 (36.98)	375.01 (46.57)	1.01	491.48 (84.35)
Gross Investments (SEK 1,000)	1,646.18 (45.78)	1,032.77 (71.07)	877.61 (80.33)	1.45	1,179.48 (157.76)

Notes: This table compares the characteristics of wealthy out-migrants (in the top 2% of household net wealth), and firms owned by wealthy out-migrants, in Sweden, who migrate before (Column (2)) and after (Column (3)) the abolition of the wealth tax (from 2001-2006 and 2007-2012, respectively). We present the group average of these characteristics in $t = -1$, i.e. the year before out-migration. Column (4) reports a t-statistic of a test of equality of these averages. For the firm characteristics, we focus on active firms which are either directly or indirectly held by wealthy out-migrants. We winsorized the top 5% and bottom 1% of these firm characteristics (except for number of employees, which we winsorized only the top 5% of). As we need to measure characteristics post-abolition, we use predicted wealth to define exposure to the reform here. For comparison, Column (1) shows the average pre-reform characteristics within the entire top 2% (i.e. without restricting to out-migrants) and the firms owned by them. In Column (5), we compute the average characteristics of marginal out-migrants using the methodology detailed in Appendix VI.4. We use the post-reform characteristics to infer the average characteristics of the infra-marginal out-migrants. Specifically, Column (5) reports the $\bar{Z}_{t-1}^m$ from the following equation: $\bar{Z}_{t-1}^m = (\bar{Z}_{t-1}^{pre} - (1 - \omega) \bar{Z}_{t-1}^{post}) * \omega^{-1}$, where $\bar{Z}_{t-1}^{pre}$ and $\bar{Z}_{t-1}^{post}$ are the average characteristics in $t - 1$ of wealthy out-migrants and firms owned by wealthy out-migrants pre-abolition and post-abolition, respectively, and ω is the fraction of migrants who migrate because of the wealth tax.

VI.4.2. Results & Robustness

Results of our calibrations of equation 15 are reported in Figure 10. The Figure shows a baseline calibration, and alternative calibrations for robustness. We explain the details of these calibrations below.

- Baseline calibration: For $\frac{\partial N}{\partial(1-\tau)} \cdot \frac{1-\tau}{N}$, this baseline calibration uses the elasticity of the population stock estimated for active entrepreneurs. This elasticity is computed using the semi-elasticity of migration flows for entrepreneurs in Figure 4 and then translating it into an elasticity of the population stock using formula 12, with $T = 15.1$. For the impact of migration, we use estimated out-migration effects for treated entrepreneurs coming from the LATE estimates in Table 2 and we do similarly for in-migration effects. Out-migration and in-migration effects are then weighted to give an average migration effect, where weights are the relative size of the respective in and out-migration elasticities, i.e. $w_{in} = \frac{\varepsilon_{in}}{\varepsilon_{in} + \varepsilon_{out}}$ and $w_{out} = \frac{\varepsilon_{out}}{\varepsilon_{in} + \varepsilon_{out}}$. Finally, the estimates for the top wealth share (i.e. the fraction of total outcome due to firms with owners in the top 2%) come from Table 1, Panel C. The standard errors of the final estimated aggregate effect are again computed using the delta method, where we assume zero covariance between the migration elasticity and migration effects.
- Alternative calibration 1: The migration effects estimated for the specific subset of “marginals” using Table 2 are not incredibly precise. Besides, results suggest that migration effects are not statistically different for marginals and infra-marginals. In light of this, and to gain in precision, in this first alternative calibration, we estimate the average migration impacts over the whole sample of treated entrepreneurs, rather than retrieving the LATE for marginal migrants. As such, we replace the estimates in the baseline calibration with those from Figure 8 (for directly and indirectly held firms).
- Alternative calibration 2: In the same spirit of gaining in terms of precision, in this second alternative calibration, we use the out-migration elasticity estimated over the whole population which is much more precisely estimated than the elasticity for entrepreneurs only. These estimates come from the ‘Top 2%’ coefficient in Figure 4 and are for the entire treated group, i.e. individuals with predicted household wealth in the top 2%.
- Alternative calibration 3: In the third and final alternative calibration, we account for firms’ buy-outs after closure. The out-migration treatment effects are given by

taking the estimates accounting for directly and indirectly held firms and buyouts from Figure 8.

For each of these alternative calibrations, all parameters remain the same as in the baseline unless explicitly specified.