The international platform of Ludwig-Maximilians University's Center for Economic Studies and the Ifo Institute

CESifo Working Papers
www.cesifo.org/wp

Behavioral Responses to Local Tax Rates: Quasi-Experimental Evidence from a Foreigners' Tax Scheme in Switzerland

Kurt Schmidheiny
Michaela Slotwinski

CESIFO WORKING PAPER NO. 5518
CATEGORY 1: PUBLIC FINANCE
SEPTEMBER 2015

ISSN 2364-1428

CESifo
Center for Economic Studies & Ifo Institute

CESifo *Working Paper No. 5518*

Behavioral Responses to Local Tax Rates: Quasi-Experimental Evidence from a Foreigners' Tax Scheme in Switzerland

Abstract

We study behavioral responses to local income taxes exploiting a special tax regime which applies to foreign employees residing in Switzerland. The used institutional setting generates two thresholds through which locally heterogeneous taxation is assigned: An income threshold at 120,000 Swiss francs and a duration threshold at 5 years of stay in Switzerland. We exploit these thresholds by applying a discontinuity in density design and a fuzzy RDD to administrative income data. We find causal evidence for strategic income bunching for wage earners and tax induced intra-national mobility. Several pieces of evidence suggest that individuals have to "learn the tax code" and that knowledge and information transmission through local networks plays a major role in the behavioral response to tax incentives.

JEL-Code: H240, H310, J610.

Keywords: income bunching, tax induced mobility, income taxes, regression discontinuity design.

Kurt Schmidheiny
University of Basel
Faculty of Business and Economics
Peter Merian-Weg 6
Switzerland – 4002 Basel
kurt.schmidheiny@unibas.ch

Michaela Slotwinski
University of Basel
Faculty of Business and Economics
Peter Merian-Weg 6
Switzerland – 4002 Basel
michaela.slotwinski@unibas.ch

September 8, 2015
We gratefully acknowledge financial support from the Swiss National Science Foundation (Sinergia Grants 130648 and 147668). We would further like to thank Taisuke Otsu, Ee-Li Xu, and Yukitoshi Matsushita for providing the basic code for their methodology.

1 Introduction

The effect of income tax reforms depends fundamentally on taxpayers' responses to these reforms. Individual responses affect taxable income and hence the change in tax revenue. Precise estimates of the overall elasticity of taxable income (ETI) with respect to tax rates as well as the heterogeneous responses of different groups of taxpayers are key in designing optimal tax policies. For a long time, the theoretical and empirical literature has almost entirely focused on labor supply responses such as the extensive and intensive adjustment in hours worked or effort (Saez et al., 2012). More recently other types of behavioral responses to evade or avoid taxes have gained attention (Feldstein, 1999, 1995; Saez et al., 2012).

This paper provides evidence on the behavioral responses of high income foreigners during their first years in Switzerland. We analyze individual data on foreign workers in Switzerland from a large labor market survey that is merged with administrative data on wage income. We find systematic earning responses at tax notches as well as systematic relocation decisions based on local income tax rates. However, these responses are not fully optimal. Several pieces of evidence show that deviations from fully optimal behavior are consistent with slow acquisition of local information through learning.

We exploit a special institutional regulation in Swiss tax law that affects high income foreigners during their first years in the country. Foreigners whose yearly gross income is below 120,000 Swiss francs (around 130,000 US-Dollars in 2015) are subject to a special tax regime (*Quellenbesteuerung*) until they get a permanent residence permit, while foreigners with gross income above 120,000 Swiss francs are taxed like Swiss citizens (*ordinary taxation*). For taxpayers in the ordinary tax regime, income tax rates differ across individual municipalities within cantons (states), while taxpayers in the special tax regime pay a single rate within each canton. This results in two types of municipalities: high-tax municipalities where the ordinary tax rate is higher than the special tax rate, and low-tax municipalities where the ordinary tax rate is lower than the special tax rate. Foreigners can apply for permanent residency after five years in Switzerland, which shifts them into the ordinary tax regime. This institutional arrangement produces two thresholds which allow identifying the causal effects of local income tax rates using a local randomization around them: an *income threshold* at 120,000 Swiss francs and a *duration threshold* after 5 years of stay. The income threshold gives rise to a sharp regression discontinuity design (sharp RD or SRD) and the duration threshold to a fuzzy regression discontinuity design (fuzzy RD or FRD).

The two tax regimes generate differential incentives for newly arrived foreigners around the income threshold. Foreigners with income below the threshold have an incentive to systematically locate in high-tax municipalities because they are not subject to the high tax rate but benefit from the relatively lower housing prices in these municipalities. Foreigners with income above the threshold have an incentive to systematically locate in

low-tax municipalities because they benefit from the low ordinary tax rates. However, we find that foreign taxpayers seem to be ignorant about local tax rates in their initial location choice. Taking the place of residence as given, the two tax regimes generate a tax notch, i.e. a discontinuous change in the tax rate at the income threshold. Using a discontinuity in density design, we find that newly arrived foreign taxpayers significantly and substantially adjust their labor income at these tax notches. Foreigners located in high-tax municipalities push their income just below the threshold to stay in the special tax scheme. Foreigners located in low-tax municipalities lift their income just above the income threshold to get into the ordinary tax scheme. This form of income adjustment is often referred to as bunching. We find more pronounced lifting of reported income than lowering. Furthermore, sub-populations differ considerably in their sensitivity to tax incentives. We trace this heterogeneity to differences in bargaining power and learning about the tax system through social networks.

Incentives also change fundamentally for foreigners in the special tax regime around the duration threshold of 5 years. Foreigners living in a high-tax municipality face a rise in income tax rates while foreigners in low-tax municipalities benefit from a drop in income tax rates. Hence, the former should be more likely to reassess their location decision while the latter should do nothing. This is exactly what we find using a fuzzy RD. This finding constitutes well identified evidence of tax induced migration for our specific subpopulation. We further find that reactions are stronger for those who would experience a higher tax rise.

The ultimate aim in the tax notch and tax kink literature is the estimation of structural elasticities. Unfortunately, the number of observations in our sample is far too low for a reliable elasticity estimation. We therefore concentrate on identifying the direction of different response channels rather than the magnitude of the response. The rich information about demographics allows identifying those sub-populations which are most sensitive to tax incentives. Furthermore, the setting used allows us to investigate bunching behavior not only at positive but also at negative notches, i.e. drops in the average tax rate. This has been studied only once before, but in a different setting by Kleven et al. (2014).

The remainder of the paper is organized as follows: The next section discusses related literature. Section 3 gives a detailed description of the institutional setting. Section 4 derives the hypotheses about individuals' behavior in the given institutional setting. The methodology and empirical strategy applied are discussed in section 5. The data is described in Section 6 and the empirical results shown in Section 7. Section 8 concludes.

2 Related Literature

Our study is related to two strands of literature: Income bunching at tax notches and tax induced migration.

Most studies investigating bunching behavior have concentrated on the analysis of tax *kinks*, i.e. discontinuous changes in the marginal tax rate (Slemrod, 2010). The majority find that bunching around kink points is modest (Chetty et al., 2011; Saez, 2010; Saez et al., 2012; Spencer and Selin, 2014). Strong responses are only found for large and salient kink points (Saez et al., 2012; Saez, 2010; Chetty et al., 2011). Systematic bunching behavior is dominantly found for self-employed workers, which could be explained by easier underreporting of self-employed income and hence may represent tax evasion rather than real income responses (Saez, 2010; Spencer and Selin, 2014; Chetty et al., 2011). The strength of this strand of empirical literature is its identification strategy, as treatment and control groups are almost identical in income and initial tax rate.

Few studies investigate bunching behavior at tax *notches*, i.e. discontinuous changes in the average tax rate. Tax notches produce stronger bunching incentives than kinks but share the identification strategy. However, tax notches are rarely observed in developed countries exactly because of their strong distortive effect on incentives (Slemrod, 2010). Kleven and Waseem (2013) find modest responses to tax notches in the Pakistan tax system, but strong bunching reactions. They further find that most taxpayers in the dominated range, generated by a positive notch, are insensitive to the arising incentives and thus the density in this range is much higher than theory would predict. They propose a methodology to use tax notches to estimate structural elasticities and optimization frictions. Numerous studies conclude that bunching is less sharp than one would expect and explain this by missing information, as do, e.g. Kleven and Waseem (2013), who explain it by the low degree of tax literacy in developing countries. So far sharp and distinct income bunching was mostly documented for the self-employed, who have more discretionary power to underreport income (see, e.g., Chetty et al., 2013).

The second important response to local taxation is location choice. Tax-induced migration is difficult to empirically identify due to the inherent reverse causality between income taxes and their tax base. The effect of taxation on location choice is not yet clearly identified and results are mixed.

Most empirical studies therefore either use instrumental variables when working with aggregate data or look at individual choices when using micro data. For example, Feld (2000) and Liebig and Sousa-Poza (2006) find no significant effect of tax rates on aggregate migration flows across Swiss municipalities. Feld and Kirchgässner (2001) and Feld (2000) study Swiss municipality-level data and find that mainly high income groups choose their location based on the resulting tax burden. Schaltegger et al. (2011) find that the proportion of high income residents in Swiss municipalities is higher if the neighboring municipalities have higher tax burdens. Schmidheiny (2006a) investigates individual-level administrative data on moving households from the metropolitan area of Basel, Switzerland, and finds that rich households are systematically more likely to move to low-tax municipalities. The most recent study on intra-national migration in Switzerland, Morger (2013), finds that local tax rates matter in location decisions but to a lesser extent than

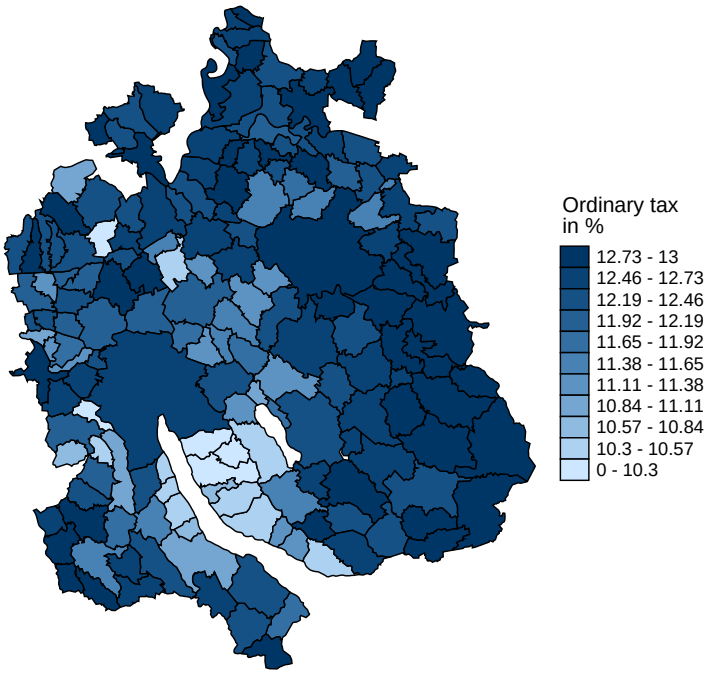


Figure 1: Ordinary average income tax rates in the canton of Zurich. Combined cantonal and municipal rate for a single household without children and a gross income of 120,000 Swiss francs in the year 2010.

other location characteristics. More recent studies exploit quasi-experimental variation to identify the causal impact of taxes. Young and Varner (2011) study the effect of an introduced millionaire tax on the migration of the affected population in New Jersey with a difference-in-differences approach and report that there is no substantial outflow that could be associated with this change. Nevertheless, they find that the 0.1 percent of the richest high income earners react more sensitively to the tax change. Kleven et al. (2013), in contrast, find a strong effect of tax burdens on the international mobility of top earners in the European football market. This finding may likely represent the upper limit, since football players are naturally a very mobile group. Additionally, and consistent with the preceding study, Kleven et al. (2014) find international migration reactions by evaluating the effect of a preferential tax scheme for high income foreigners using Danish administrative tax data. They find that preferential tax rates for foreigners increased the inflow of high income individuals.

3 Institutional Background

Switzerland is a natural laboratory for research on local taxation, due to the huge variation in tax rates down to the municipal level. The Swiss Confederation consists of three organizational layers: the federal level at the top, the cantonal layer of 26 cantons, and around 3000 municipalities that form the municipal layer. All three layers have considerable autonomy in both setting tax rates and deciding on expenditures.

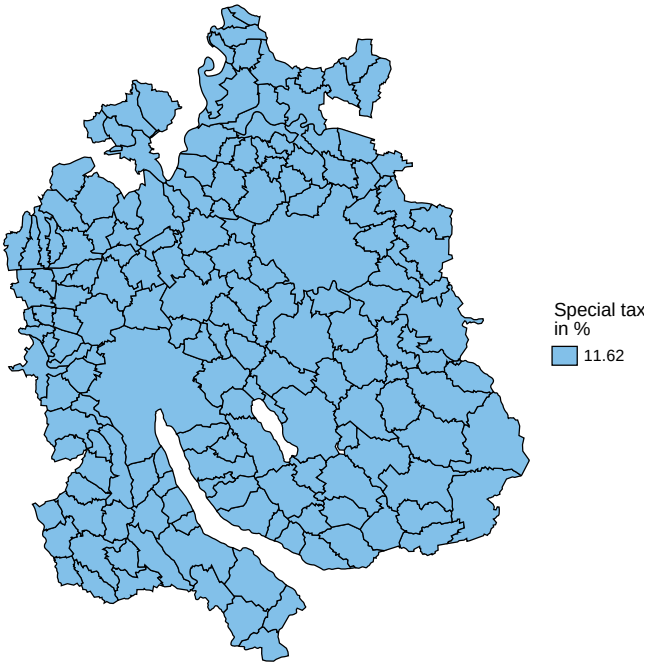


Figure 2: Special average income tax rates in the canton of Zurich. Combined cantonal and municipal rate for a single household without children and a gross income of 119,999 Swiss francs in the year 2010.

The income tax liability for a person with a tax domicile in Switzerland is composed of three parts:[1] The federal income tax, the cantonal income tax and the municipal income tax. The progressivity of tax rates is determined by the federal and the cantonal tax schedules. The municipal tax rate is calculated as the cantonal tax rate times a tax multiplier (*Steuerfuss*) set by the municipality. The total tax liability from all three parts therefore depends on household type and income class. Taxpayers file an annual tax declaration with this information and pay their tax retrospectively. The described *ordinary* tax regime results in substantial differences in income tax rates across Swiss municipalities. Figure 1 shows the ordinary tax rates for a high-income individual in the canton of Zurich as an example.

Besides the ordinary tax regime, there is a *special* tax scheme ("Quellensteuer") which has been applied nationwide since 2001. Tax payments in the special regime are directly subtracted from the salary and collected by the employer. As a result, all labor income of foreigners within this tax regime is third-party reported. Household characteristics such as marital status and children, which affect deductibles, are taken into account such that taxpayers in the special regime do not have to, and are not meant to, file a tax declaration.

The special tax regime applies to incomes from employed *foreign* persons who have their tax domicile in Switzerland, have not received the Swiss permanent residence permit (type C) yet, and are not married to a Swiss citizen or to a foreign person who holds a

[1]In some cantons the church tax is also added. Depending on the canton this can be mandatory or voluntary.

Table 1: Thresholds through which tax regime is assigned

	temporary residence permit (B)	*permanent* residence permit (C)
gross income $< 120,000$	*special* tax regime (constant tax rate)	*ordinary* tax regime (local tax rate)
gross income $\geq 120,000$	*ordinary* tax regime (local tax rate)	*ordinary* tax regime (local tax rate)

permanent residential permit.[2] Foreigners holding a temporary residence permit (type B) can apply for a permanent residence permit (C) after a five year stay in Switzerland.[3] All foreigners with yearly taxable income above 120,000 Swiss francs are subject to the ordinary tax regime.[4]

In the special tax regime, there is a canton-wide municipal and cantonal tax rate. This canton-wide tax rate is calculated as a weighted average of the municipal rates within the canton for a given income and household type. Foreigners subject to this special tax regime therefore face no variation in tax rates across municipalities within a canton. Figure 2 shows the special tax rates for a high-income individual in the canton of Zurich as an example. The average income tax rate for an unmarried foreign worker without children and a gross income of about 120,000 Swiss francs amounts to 11.62 percent in the special tax regime and between 10 and 13 percent in the ordinary tax regime.

The two tax regimes produce two thresholds: an income threshold and a duration threshold. The two thresholds are summarized in Table 1.

The *income* threshold occurs if the total gross yearly income of a foreign employee holding the temporary residence permit (B) crosses the threshold of 120,000 Swiss francs. This threshold leads to a sharp assignment of the treatment, in this case locally heterogeneous taxation, as soon as the assignment variable (total yearly gross income) exceeds

[2]Foreigners who cross the border daily are also taxed at a special tax rate, but this is not of interest in this study, thus an explanation will be neglected in this work.

[3]The five year limit applies to foreigners from Belgium, Germany, Denmark, France, Liechtenstein, Greece, Italy, the Netherlands, Austria, Portugal, Spain, Finland, Great Britain, Ireland, Iceland, Luxembourg, Norway, Sweden, USA, Canada, Andorra, Monaco, San Marino and Vatican City. For foreigners from all other countries the normal year limit is ten years, but they are also allowed to apply for a permanent residence permit after five years if they can meet certain integration conditions, which seems to be the case for a great number.

[4]In the canton of Geneva the limit is 500,000 Swiss francs. Geneva is therefore ignored in the empirical analysis.

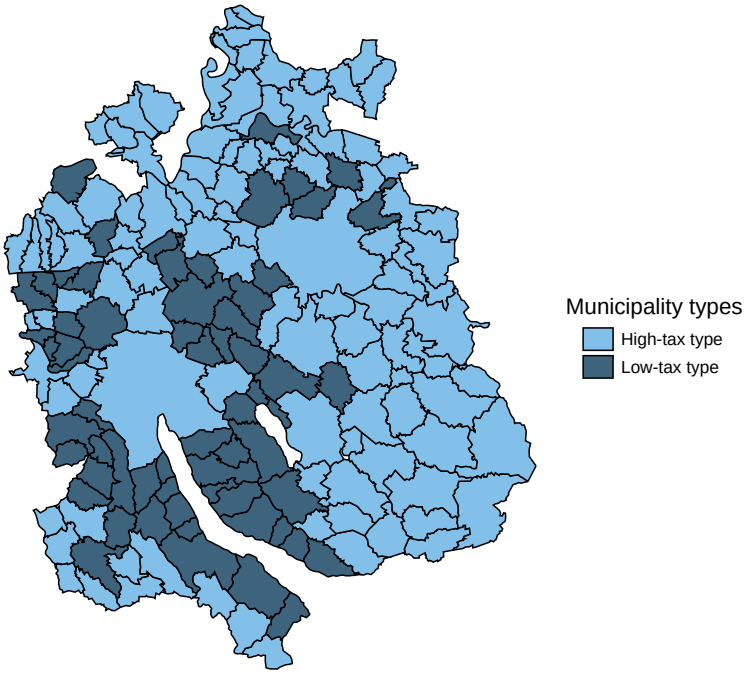


Figure 3: High- and low-tax type municipalities in the canton of Zurich.

a fixed threshold (120,000 Swiss francs). So the treatment probability jumps from zero to one at this threshold. [5]

The *duration* threshold occurs if an individual with an income below 120,000 Swiss francs applies for the permanent residence permit (C). Application for a permit C is a legal right after five years of residence and not an obligation. The *duration* threshold therefore leads to a fuzzy regression discontinuity design (fuzzy RD or FRD) in which the treatment (ordinary tax regime) can only be taken up if the assignment variable (duration of stay) crosses the threshold (five years).

As described above, the special tax rate is calculated as the population weighted average of the local tax rates across all municipalities within a canton. The local ordinary tax rate will therefore be higher than the canton-wide special tax rate in some municipalities and lower in others. We will address the former municipalities as *high-tax* municipalities and the latter as *low-tax* municipalities. Figure 3 shows the resulting grouping exemplified for the canton of Zurich.

4 Theoretical Considerations

For the majority of taxpayers, i.e. those in the ordinary tax regime, low-tax municipalities are ceteris paribus attractive. In spatial equilibrium, we would therefore expect that tax advantages are capitalized in higher housing prices or offset by lower public goods

[5] This setting seems to give rise to a sharp regression discontinuity design (sharp RDD or SRD). However, as we will document in our results, bunching behavior of the taxpayer around the threshold invalidates the RDD assumptions but allows for evidence of income bunching instead.

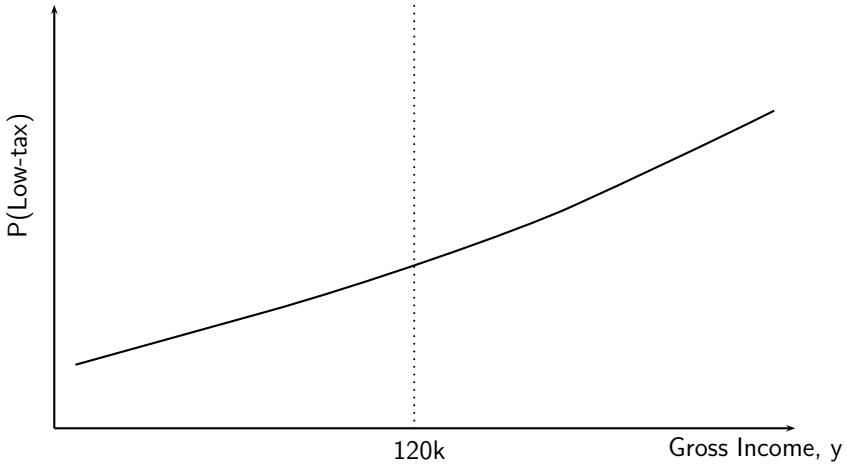


Figure 4: Location for residents in ordinary tax scheme.

provision. Tax capitalization has been shown for Switzerland by Feld (2000) and recently Stadelmann and Billon (2012). Public goods provision in Swiss municipalities, however, is on average even higher in low-tax municipalities (see, e.g., Schmidheiny, 2006b). This is consistent with local budget balance and income sorting: Low-tax municipalities attract relatively more high income households than high-tax municipalities. The relatively higher per capita tax base allows low-tax municipalities to finance higher per capita public expenditures with lower linear tax rates. Low-tax municipalities are therefore typically characterized by low taxes, high housing prices and high public goods provision while high-tax municipalities are characterized by high taxes, low housing prices and low public goods provision. Schmidheiny (2006b) theoretically shows that this configuration is an equilibrium in a multi-community model. The systematic sorting of high-income households into low-tax municipalities is theoretically driven either by non-homothetic preferences, i.e. the decreasing housing expenditure shares with respect to income (Schmidheiny, 2006b) or by progressive income tax schedules (Schmidheiny, 2006a). Basten et al. (2014) provide recent empirical evidence for sorting of high-income households in Switzerland. Figure 4 shows the theoretical prediction for the probability that a taxpayer in the ordinary tax regime locates in a low-tax municipality.

We will now outline two orthogonal hypotheses for the behavior of foreigners not holding the permanent residence permit (C) with respect to the income threshold. In our first hypothesis, we assume that foreign households take their income as given and strategically choose their place of residence. Foreigners with income above 120,000 Swiss francs are in the ordinary tax regime. Foreigners with income below 120,000 Swiss francs are in the special tax regime. High-tax municipalities are particularly attractive for the latter because they pay low housing prices but do not have to pay the high tax rates. Hypothesis 1 is therefore:

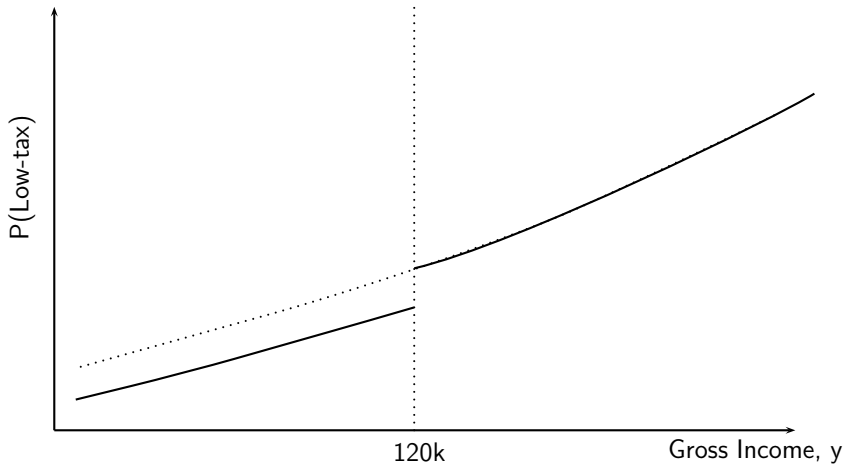


Figure 5: Location of foreign residents under Hypothesis 1.

Hypothesis 1 (mobility response at income threshold)**.**
Foreigners with an income below 120k and taxed at the special tax rate are more likely to locate in a high-tax municipality than foreigners with an income above 120k in ordinary taxation who are more likely to locate in a low-tax municipality.

Under Hypothesis 1, we would expect a positive jump in the probability of observing a foreign taxpayer in a low-tax municipality at the income threshold as shown in Figure 5.

In our second hypothesis, we assume that foreigners are ignorant as to tax rates when they arrive in Switzerland. This could be motivated by the fact that in most other countries there is no such local tax heterogeneity and individuals might not be aware of it when they arrive in Switzerland. The initial location choice is therefore not systematically related to tax rates. Once settled, they learn that local tax rates may depend on their income. In this situation, the shift between the two tax regimes at the income threshold creates a discontinuity in the average tax rate. Kleven and Waseem (2013) call this a pure *tax notch*. Figure 6 shows this notch for an exemplary high tax municipality and a low-tax municipality. For simplicity, we assume a linear tax here.[6]

We assume a smooth tax schedule in the ordinary regime and a smooth distribution of ability around the income threshold. This would theoretically translate into a smooth income distribution around the threshold. Given location, a household faces a tax notch at the threshold income y^* such that the tax rate is t below and $t_{low} \neq t$ or $t_{high} \neq t$ above the threshold. This generates a discontinuity in the budget set at y^*, which would theoretically translate into a discontinuity in the income distribution (Kleven and Waseem, 2013).

Given that a foreign household locates in a high-tax municipality, there is an income region, Δy, above the threshold of 120,000 Swiss francs, in which the household would

[6]Incomes tax rates are highly progressive at all three state levels. The derivation of the dominated regions holds as a local linear approximation.

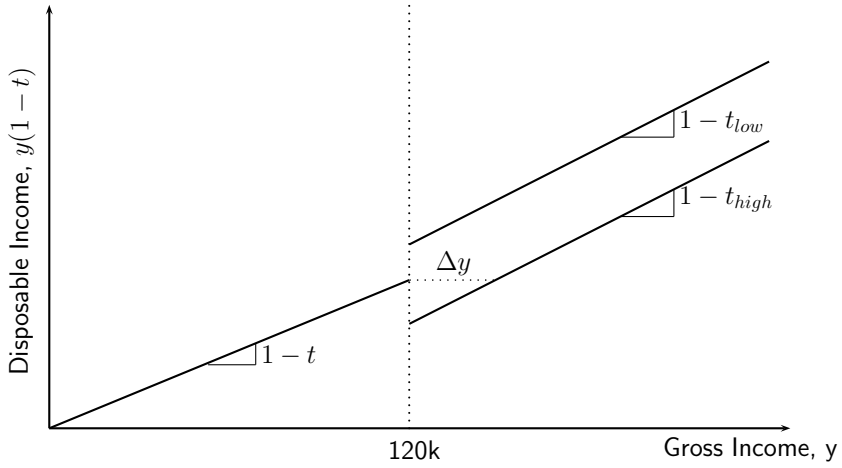


Figure 6: Disposable income for foreign residents conditional on location.

have more disposable income by simply "throwing away" all income above the threshold, as illustrated in Figure 6. This generates a strictly dominated income region above the threshold given by

$$120,000 \leq y^* < 120,000\frac{(1-t)}{(1-t_{high})}.$$

Consequently, we would expect households in high-tax municipalities with incomes in the dominated region to strategically adjust their income to just below the threshold of 120,000 Swiss francs. Figure 7 visualizes the resulting income distribution *conditional* on locating in a high-tax municipality. Assuming rational, well informed agents, there should be no mass in the dominated region above the income threshold and a mass point (excess bunching) just below it. Furthermore, households with income slightly above the dominated region, $120,000 + \Delta y$, could almost *keep* their disposable income when lowering their income below the threshold. We would expect some households above the dominated region also to reduce their income below the threshold. Households just above the dominated region will be more likely to reduce their income than households further above the threshold. The income distribution in high-tax municipalities should therefore be characterized by a negative discontinuity at the threshold income, excess mass below the threshold and a hole above the threshold, as visualized in Figure 7.

Given that a foreign household locates in a low-tax municipality, a household with income just below 120,000 Swiss francs could *substantially reduce* its tax payments by *slightly increasing* its income above the threshold. There is no strictly dominated region below the threshold as income adjustments in this direction are always costly. Assuming exponential costs (effort) to increase income, some households with income slightly below the threshold will marginally increase effort to achieve a substantial discrete increase in consumption possibilities (Kleven and Waseem, 2013). We would therefore expect some individuals below the threshold to bunch above it. Households just below the

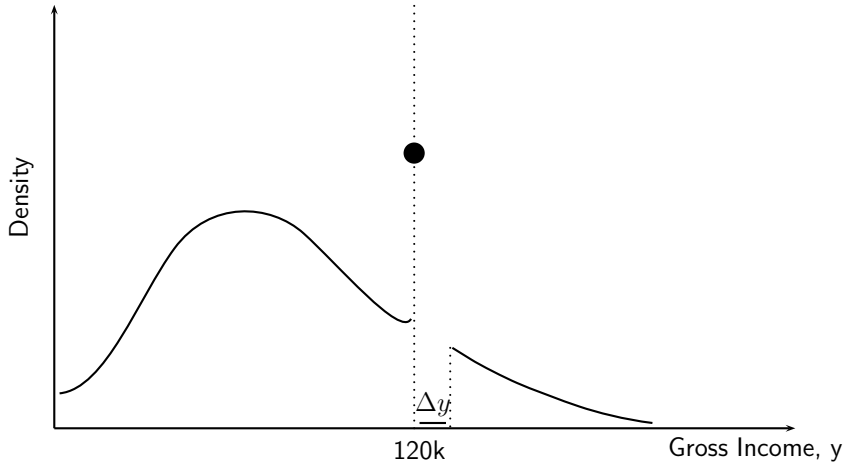


Figure 7: Income distribution conditional on locating in a high-tax municipality.

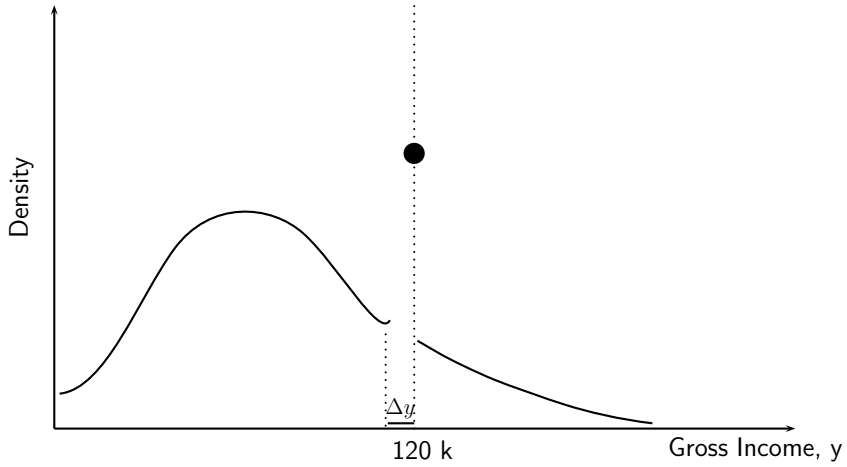


Figure 8: Income distribution conditional on locating in a low-tax municipality.

threshold will be more likely to increase their income than will households further below the threshold. This would again result in a discontinuous income distribution with excess mass above the threshold and missing mass below it. The income distribution in low-tax municipalities should therefore feature a positive discontinuity at the income threshold, excess bunching above the threshold, and a hole below it, as visualized in Figure 8.

Kleven and Waseem (2013) provide a formal framework for studying such income adjustments. Income can potentially be adjusted by lowering hours worked, lowering effort, by renegotiating declared wages or by negotiating the division between fixed and flexible income components. The income responses we find are unlikely due to tax evasion, as taxable income is third-party reported in our sample, i.e. for individuals in the special tax scheme. Kleven et al. (2011) e.g. show that enforcement is strong under third party reporting.

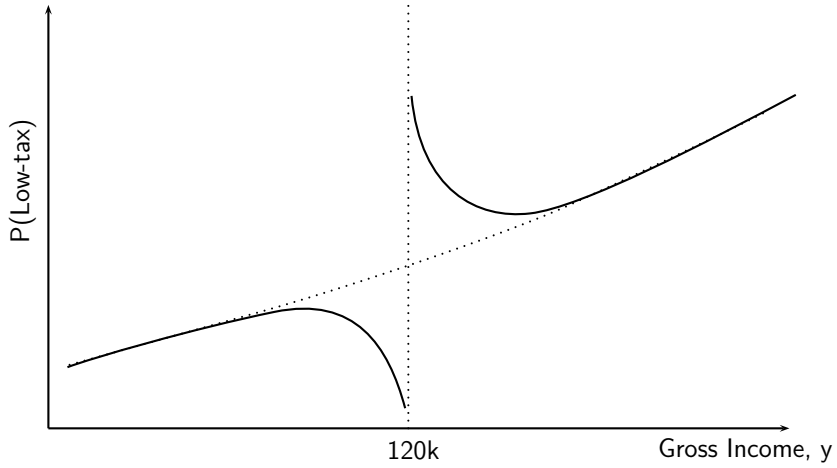


Figure 9: Location of foreign residents under Hypothesis 2.

The above outlined bunching behavior when taking location as given is summarized in Hypothesis 2:

Hypothesis 2 (wage response at income threshold)**.**
Individuals locating in a low-tax municipality seek to increase their income to be above 120k. Those locating in a high-tax municipality seek to decrease their income to be below 120k.

Hypothesis 2 translates into a distinct pattern of *observed* locations of individuals around the income threshold, as visualized in Figure 9. This form of visualizing the testable consequences of Hypothesis 2 allows us to discriminate between Hypothesis 1 (Figure 5) and Hypothesis 2 (Figure 9).

The third hypothesis considers the behavioral response of foreigners after 5 years of residence. As discussed in Section 3, foreigners in the special tax regime who located in high-tax municipalities will experience a substantial increase in tax rates while foreigners who located in low-tax municipalities will experience a substantial drop in tax rates after receiving the permanent residence permit (C). We therefore postulate the following testable Hypothesis 3:

Hypothesis 3 (mobility response at duration threshold)**.**
After 5 years, foreigners with an income below 120k who apply for the permanent residence permit reconsider their location choice. Those located in a low-tax municipality are likely to stay. Those located in a high-tax municipality are more likely to move. The probability of moving in high-tax type municipalities increases with income.

The following section outlines the empirical strategy applied in order to test the hypotheses derived in this section.

5 Empirical Strategy

5.1 Income Threshold

To disentangle Hypothesis 1 and 2 we estimate the probability of a foreign individual locating in a low-tax municipality as a function of his/her income. We are not only interested in the point estimate at the income threshold but also in the shape of the curve around the threshold. To further investigate Hypothesis 2, we then analyze the income distributions in high-tax and low-tax municipalities. We verify strategic income adjustment by testing whether the income densities exhibit systematic discontinuities at the income threshold. A discontinuity in the density at the threshold would not only be evidence for bunching but also for very precise income adjustment around the threshold.

We first estimate the probability of a foreign individual locating in a low-tax municipality in a classical sharp regression discontinuity design (RDD). We nonparametrically estimate the location probability using a local linear regression (LLR), separately on both sides of the threshold. See Porter (2003) or Hahn et al. (2001) for an introduction to RDD, and Fan and Gijbels (1996) to LLR. Our dependent variable is an indicator variable which takes the value 1 if the individual lives in a low-tax municipality and 0 otherwise. Therefore we apply the following estimation equation to every estimation point separately to both sides of the threshold (c).

$$\min_{\alpha\beta} \sum_{i=1}^{N} (y_i - \alpha - \beta(x_i - x))^2 K_h(x_i - x) \tag{1}$$

Where y is the dependent variable of interest, x_i represents the commonly labeled assignment variable, h represents the suitable bandwidth $h > 0$, $K_h(x_i - c)$ represents the kernel weighting function at the estimation point x and is defined as $K_h(x_i - c) = \frac{1}{h} K\left(\frac{x_i - c}{h}\right)$. We use a triangular kernel as it is the optimal one in RDD settings because of its properties at boundary points (for more details see Hahn et al. (2001) or Fan and Gijbels (1996)) and it should not make a huge difference for the shape of the curve estimated far from the threshold.

For nonparametric graphic evidence on bunching behavior and for the sake of consistency with our analytical estimates we estimate the income distributions using local likelihood density estimation separately from both sides of the threshold (see, Loader, 1996; Otsu et al., 2013).

The density to the left and the right of the threshold is estimated by maximizing the following maximization problems at each estimation point:

$$\max_{a,b} \left\{ \frac{1}{n} \sum_{i:x_i<c} K\left(\frac{x_i - x}{h}\right)(a + b(x_i - x)) - \int_{u<c} K\left(\frac{u - x}{h}\right) exp(a + b(u - x))du \right\};$$

$$\max_{a,b} \left\{ \frac{1}{n} \sum_{i:x_i\geq c} K\left(\frac{x_i - x}{h}\right)(a + b(x_i - x)) - \int_{u\geq c} K\left(\frac{u - x}{h}\right) exp(a + b(u - x))du \right\}$$

where x is the estimation point in x_i and c stands for the threshold value in x_i. $K(\cdot)$ is a symmetric kernel weighting function and h the bandwidth used. We use the triangular kernel in our application. The estimators of f are then defined as $\hat{f} = exp(\hat{a})$.[7]

We then nonparametrically estimate discontinuities in the income distribution to detect bunching around the income threshold.[8] Bunching invalidates the classical RDD estimates and RDD studies therefore typically test for the presence of bunching. We extend the bunching analysis used so far by nonparametric inference on the discontinuity in the density. The resulting discontinuity measure provides evidence for the presence of strategic income adjustment and for this adjustment being targeted very precisely around the threshold. Combined with the usual bunching approach, it could be further set into relation with the excess mass to render information about individuals' ability to precisely adjust their earnings. Additionally, we propose to use an estimation procedure which allows the analysis of strategic bunching behavior even in situations where data is scarce. The most widely used test for discontinuities in densities by McCrary (2008) is based on a local linear smooth of binned counts. Recently, Otsu et al. (2013) proposed an alternative approach for the estimation of discontinuities in densities based on the empirical likelihood approach, using a local likelihood density estimate (Loader, 1996). This approach has several advantages over McCrary (2008) in our application.[9] First, the approach by Otsu et al. (2013) shares the good boundary properties of the local linear estimate.[10] This is in our case crucial for the separate estimates at the income threshold from the left and from the right. Second, the Otsu et al. (2013) estimator is nonnegative by construction, while the McCrary (2008) estimator can produce negative density estimates. This property is particularly important in our application because the theoretical density in dominated regions is zero and the empirical density possibly very low. Third, the McCrary (2008) approach is sensitive to the chosen binwidth for the binned counts in small samples. The Otsu et al. (2013) approach only requires choosing the bandwidth. Further it provides a general framework for inference of discontinuities in densities, drawing on the idea of empirical likelihood, and proposes a test and confidence sets which are invariant to the formulation of the parameter of interest and are well defined even if the local linear binning estimate turned out to be negative.

[7] More explicitly we use the local likelihood implementation (locfit) in the Chronux software package for Matlab including the weighting (Bokil et al., 2010).

[8] Most recent studies investigating income responses around tax kinks and notches seek to estimate the underlying elasticities of taxable income. See e.g. Spencer and Selin (2014), Saez et al. (2012), Saez (2010), Chetty et al. (2011) and Kleven and Waseem (2013). These studies need very large administrative data sets in order to estimate the difference (excess mass) between the observed density and the counterfactual density in the absence of bunching. Unfortunately, the number of observations in our survey data is far too small to follow this approach and produce reliable estimates.

[9] For a more general introduction to empirical likelihood methods and their computational issues see Owen (2001) or Kitamura (2008).

[10] Compared to other nonparametric estimators a local linear specification exhibits a superior bias and convergence behavior in estimations at boundary points. For more details see Fan and Gijbels (1996).

Otsu et al. (2013) propose to estimate the density separately from both sides of the threshold using a local (linear) likelihood density estimation. The point estimates of the density from the left ($\hat{f}_l$) and the right ($\hat{f}_r$) of the threshold are estimated by maximizing the following maximization problems:

$$\max_{a_l,b_l}\left\{\frac{1}{n}\sum_{i:x_i<c}K\left(\frac{x_i-c}{h}\right)(a_l+b_l(x_i-c))-\int_{u<c}K\left(\frac{u-c}{h}\right)exp(a_l+b_l(u-c))du\right\}; \quad (2)$$

$$\max_{a_r,b_r}\left\{\frac{1}{n}\sum_{i:x_i\geq c}K\left(\frac{x_i-c}{h}\right)(a_r+b_r(x_i-c))-\int_{u\geq c}K\left(\frac{u-c}{h}\right)exp(a_r+b_r(u-c))du\right\}$$

where c is the threshold value in x_i and, thus in our case, the income threshold of 120,000 Swiss francs. $K(\cdot)$ is a symmetric kernel weighting function and h the bandwidth used.

There are different procedures for an automatic optimal bandwidth choice in non-parametric estimation. We choose the bandwidth to be as small as possible such that the resulting curve still looks reasonably smooth. Although the optimal bandwidth would vary in sample size, we choose to interpret the result for a fixed set of bandwidths to keep results comparable. To validate that our results are not driven by the bandwidth choice, we report the analytical results for different bandwidths, where the last row in the result tables reports the results for optimal bandwidths in discontinuity in density estimates following the bandwidth choice criterion in McCrary (2008). As argued above, we use the triangular kernel in our application. We multiply the kernel weight by the inverse sampling weight d_i in the summation terms of (2), but not the integral term, to account for the survey weights in our data. d_i is defined as $d_i = (w_i/\sum_{i=1}^{n} w_i)\cdot N$, where the sampling weight $w_i = 1/\Pi$ is the inverse of the inclusion probability Π and N is the sample size.

The estimators of f_l and f_r are then defined as $\hat{f}_l = exp(\hat{a}_l)$ and $\hat{f}_r = exp(\hat{a}_r)$. The discontinuity parameter θ_0 is estimated by their difference $\hat{\theta} = \hat{f}_r - \hat{f}_l$.

Inference on the parameter of interest is based on the empirical likelihood function for the parameter of interest θ_0 which is constructed as follows:

$$L(a_l,a_r,b_l,b_r) = \sup_{\{p_i\}_{i=1}^n}\prod_{i=1}^n p_i,$$
$$s.t. \quad 0\leq p_i\leq 1, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i g_i(a_l,a_r,b_l,b_r) = 0,$$

with

$$g_i(a_l,a_r,b_l,b_r) = \begin{pmatrix}(1,x_{i,h})'(1-I_i)K(x_{i,h}) - & \int_{x<c}\left(1,\frac{x-c}{h}\right)'K\left(\frac{x-c}{h}\right)exp(a_l+b_l(x-c))dx,\\ (1,x_{i,h})'I_iK(x_{i,h}) - & \int_{x\geq c}\left(1,\frac{x-c}{h}\right)'K\left(\frac{x-c}{h}\right)exp(a_r+b_r(x-c))dx\end{pmatrix}$$

Where g results from the estimating equations (or sample moment conditions), the FOC of (2), I_i is defined as $I_i = \mathbb{1}\{x_i\geq c\}$ and $x_{i,h} = \frac{x_i-c}{h}$. The weights p_i can be seen as probability mass allocated to the observed values of x_i. The problem can be represented

in its dual form of the log empirical likelihood ratio function

$$lr(a_l, a_r, b_l, b_r) = 2 \sup_{\lambda \in \mathbb{R}^4} \sum_{i=1}^{n} log(1 + \lambda' g_i(a_l, a_r, b_l, b_r)).$$

This considerably reduces the dimensions of the maximization problem. The concentrated log local likelihood ratio function ((profile) empirical likelihood ratio) of the parameter of interest ($\hat{\theta} = \hat{f}_r - \hat{f}_l$) is

$$lr(\theta) = \min_{a_l, a_r, b_l, b_r : \theta = exp(a_r) - exp(a_l)} lr(a_l, a_r, b_l, b_r)$$

Under the null hypothesis $H_0 : \theta = \theta_0$ the empirical likelihood function follows asymptotically a chi-square distribution

$$lr(\theta_0) \xrightarrow{d} \chi^2(1).$$

The null hypothesis $H_0 : \theta_0 = \theta$ for some θ can be tested by $lr(\theta)$ using $\chi^2(1)$ critical values.[11] For a more in-depth discussion and derivations see the original paper of Otsu et al. (2013). As before, we multiply the kernel weights by d_i in the summation part of the FOC or g, but not in the integral, part to adjust the estimates for the sampling weights of our data.

Hypothesis 2 predicts $f_r/f_l < 1$ for a positive notch (high-tax municipalities) and $f_r/f_l > 1$ for a negative notch (low-tax municipalities). The empirical likelihood ratio is invariant to the formulation of the null hypothesis. The test statistics for $H_0 : f_r - f_l = 0$ and $H_0 : f_r/f_l = 1$ are therefore identical.

5.2 Duration Threshold

As described in Section 3, in our setting foreign workers can apply for a permanent residence permit (C) after 5 years in Switzerland. For foreigners with incomes below 120,000 Swiss francs, the permanent residence permit (C) implies a switch from the special to the ordinary tax regime. The effect of the change in the tax regime on the mobility of foreign individuals can be estimated with a fuzzy regression discontinuity design (fuzzy RD).

The fuzzy RD identifies average treatment effects for compliers at the threshold. In our application these are foreign workers applying and receiving the permanent residence permit (C) at the threshold of five years. The estimated effects are consistent under the standard assumptions of the fuzzy RD: local randomization, a clear jump in the probability of treatment, continuity of conditional expectations, excludability and monotonicity.[12]

[11] More precisely for matters of implementation, to test a particular value of θ g_i is set to $g_i(a_l, log(\theta_0 + exp(a_l)), b_l, b_r)$.

[12] These identifying assumptions are formally stated in Hahn et al. (2001) and well described in Lee and Lemieux (2010) or Imbens and Lemieux (2008).

Our estimates are standard nonparametric (local linear) fuzzy RD estimates following Porter (2003). We include sampling weights from the survey design.[13] The treatment effect (τ) is consistently estimated as[14]

$$\widehat{E(\tau_i|x_i = c)} = \frac{\widehat{\lim_{x\to c^+} E(y_i|x_i = x)} - \widehat{\lim_{x\to c^-} E(y_i|x_i = x)}}{\widehat{\lim_{x\to c^+} E(w_i|x_i = x)} - \widehat{\lim_{x\to c^-} E(w_i|x_i = x)}} \tag{3}$$

where y stands for the dependent variable of interest, w represents the treatment indicator, x represents the assignment variable (duration of stay in Switzerland), c represents the threshold value (5 years), $\lim_{x\to c^+}$ stands for the limit approaching the threshold from the right hand side and $\lim_{x\to c^-}$ that coming from the left hand side. In our application, w is set to one if an individual holds permit C and to zero if not. $E(w_i|x_i = x)$ stands for the probability of applying for the permanent residence permit C after five years of stay. $E(y_i|x_i = x)$ stands for the probability of moving after five years of stay. These limits are estimated by local linear regression using a triangular kernel, which is the optimal weighting function in RDD settings. The bandwidth is again chosen to be as narrow as possible but still rendering a smooth function, and as before we choose to interpret the results for a fixed set of bandwidths to keep the results comparable. We report results using several bandwidths to validate the robustness to its choice, where the last row in the results tables reports regression discontinuity estimates based on the optimal bandwidth proposed in Calonico et al. (2013). Fuzzy RD is often linked to instrumental variables (see Imbens and Angrist, 1994), in this context our *first stage* would be the probability that an individual applies for the permanent residence permit after five years of stay, while the *second stage* would be the probability of moving (Imbens and Angrist, 1994; Imbens and Lemieux, 2008).

The local randomization assumption is likely valid in our setting. Every foreign individual living in Switzerland and holding a temporary residence permit (B) will automatically cross the threshold after five years. The only possibility of manipulation would be to emigrate shortly before crossing the threshold and then immigrating again, which would carry the risk that a new permit could be denied. But this manipulation is rather unrealistic because the permanent residence permit (C) is very beneficial to workers. In contrast to permit B, permit C is unconditional, does not expire, does not depend on employment status or whether a ward depends on welfare. It further equalizes foreigners to natives on the labor market, so e.g. foreigners holding the permanent residence

[13] We use the results of Harms and Duchesne (2010) to incorporate and correct for sampling weights in the nonparametric LLR estimates. The correction term is estimated by

$$\widehat{\Theta + s} = \frac{N}{\sum_{i=1}^{N} \frac{1}{\Pi_i}} \cdot \frac{\sum_{i=1}^{N} \frac{1}{\Pi_i} \cdot \frac{1}{\Pi_i}}{\sum_{i=1}^{N} \frac{1}{\Pi_i}}$$

and results in 1,624 in our application.

[14] For details of LLR estimation in the RDD framework see Porter (2003) or Fan and Gijbels (1996) for LLR estimation in general. Imbens and Lemieux (2008) and Lee and Lemieux (2010) provide a very practical guide on the estimation and validation of a RDD.

permit are allowed to start their own business. Many cantons even automatically assign the permanent residence permit to foreigners as soon as they are eligible and individuals cannot legally refuse to accept it. These advantages and regulations very likely prevent a strategic self selection into permit B status and thus the special tax regime for longer; they further speak very much in favor of the monotonicity assumption. The long-run proportion of those who hold permit C is consequently very high, so in our estimation sample about 95 percent of individuals residing in Switzerland for longer than 10 years hold the permanent residence permit.

Another important assumption in the FRD setting is a clear jump in the probability of treatment ($P(w = 1)$). It will be empirically validated in Section 7.2.

The continuity of conditional expectations assumption, i.e. that the effect found can be attributed to the treatment received, is generally difficult to defend. We are not aware of any other change taking effect at the threshold of 5 years of stay. Of course if an individual decides to stay in a country, and thus applies for a permanent residence, it is possible that other influencing factors beyond taxation affect location choice. The restrictions concerning the possibility to move within Switzerland do not differ greatly between permits B and C. Thus the choice set remains more or less the same. Since individuals under special taxation can choose their location within a canton freely, without concerns about the applicable tax rate, it can be assumed that they choose the municipality which fits their preferences best. Assuming that their preferences do not change dramatically when they cross the threshold of five years, they would choose the same one except if taxation begins to make a difference. Consequently, they should choose the same municipality if taxes do not matter for their location choice. We propose a quasi-placebo test to validate the continuity of conditional expectations and excludability assumption in our setting. We repeat the analysis for those foreigners whose income exceeds the income threshold of 120,000 Swiss francs, and who are thus already taxed under the ordinary regime when receiving the permanent residence permit. If in fact it is the change in the taxation scheme that produces the effect, we would not expect a mobility reaction for this placebo group, while this group should react like the other one if it is something different. We include the results for this particular quasi-placebo group in each analysis.

6 Data

We use individual data from the largest Swiss labor market survey SAKE ("Schweizerische Arbeitskräfteerhebung") augmented by individual income data from the social security administration. This linked dataset is called SESAM ("Syntheseerhebung soziale Sicherheit und Arbeitsmark"). It combines detailed individual information with very accurate income measures. SAKE is a household survey, commissioned by the Swiss Federal Statistical office (BFS). Its main goal is to record the Swiss permanent resident labor force's structure and employment behavior. It has been carried out in computer assisted tele-

phone interviews since the year 1991. The survey is organized as a rotating panel study. Each person is interviewed annually over a duration of five years and then replaced (BFS, 2004).

The administrative income data of SESAM is essential for our analysis as we need to observe the income threshold very precisely. See Moore et al. (2000); Saez (2010); Kleven et al. (2013) on problems of measurement error in reported income measures. We use the pooled SESAM dataset for the years 2001 to 2012[15].

The main variables in the empirical analysis are:

- *Gross income*: This variable contains the gross income of an individual, including 13th or14th monthly salary and premiums or bonuses, within the last 12 months.
- *Duration*: This variable contains information about how long a foreign target person has already lived in Switzerland. Respondents are first asked to indicate the year and month they arrived in Switzerland. This information is converted to days using approximations for missing information on the month (for details see BFS, 2012).
- *Move*: This variable is an indicator variable with value one if an individual indicates having moved within the last 12 months and zero otherwise.

Additionally, we use several demographic variables for the definition of the sample and the treatment groups, where needed.[16] Descriptive statistics on the main variables can be found in Table A.1.

Our basic estimation sample comprises individuals who meet all requirements for the special tax scheme, except for the income, the duration of stay, and the permit type restriction. Our sample includes foreign employees whose employment status is known (every individual indicating self-employed income is excluded), who are at least 16 years old, and whose location of residence as well as nationality is known.[17] Swiss citizens and dual citizens are excluded for the basic sample. The sample is further restricted to those holding either permit B or C, such that exceptional cases like asylum seekers, etc. are excluded. We also exclude those who are married to a Swiss citizen or a dual citizen since they would be taxed ordinarily.[18] We further restrict our sample to foreigners residing in Switzerland for at least one year, so that we can observe their yearly income.

[15]We use only data up to 2012 as the practical implementation of the special tax scheme changed in several cantons due to a court judgment that this tax scheme is discriminatory in 2010.

[16]These are in particular: age, employment status, nationality, permit type, marital status, nationality of spouse, information on real estate property, canton and municipality of residence.

[17]We exclude individuals who could not respond correctly when asked where their location of residence is. Further, proxi-interviews are excluded, i.e. observations of individuals who were not interviewed in person.

[18]One exception applies for the canton Ticino. Individuals married to a Swiss citizen are not excluded if they reside in Ticino, because there this does not lead to ordinary taxation.

Individuals who own real estate in cantons where this leads to ordinary taxation are excluded.[19] Citizens of The United States and Eritrea are ignored, as theses states tax the worldwide income of their citizens not residing in the country. Incentives arising from the two tax regimes are consequently very different for these individuals. We use only observations for which the indicated duration of stay is trustworthy, i.e. we exclude observations for which the duration of stay in the same municipality exceeds the duration of stay in Switzerland for more than half a year. Finally we exclude very small municipalities with a population size below 500. This constitutes our basic sample which is then further adjusted for the analysis at hand. The sample of Swiss nationals used as a placebo group is generated by the same criteria except for nationality.

The foreign individuals included are drawn from a central index of all foreigners living in Switzerland ("Zentrales Ausländer Register", ZAR). The sampling strategy with respect to foreign persons is not only geographically stratified but also adjusted for demographic proportions, groups of foreigners, gender and duration of stay. Sample weights are therefore very important in the estimation. The SAKE dataset provides a weight variable which is meant to correct for the stratification and the complex survey design.[20] It represents the inverse of the probability of being included in the sample ($w_i = 1/\pi_i$). It not only corrects for the consequences of the stratification, but is additionally adjusted for the probability of a drop out before being sampled out and further factors which are not discussed in detail here (for more details with respect to the creation of this weight see BFS, 2004).

We also use data on average tax rates in the ordinary tax regime for the years 2001 to 2012. These tax rates are calculated by the Swiss Federal Tax Administration for the 813 largest Swiss municipalities.[21] The calculation considers cantonal statutory tax rates as well as canton-specific deductions for e.g. children. The cantonal tax rate depends on income and in most cantons also on marital status. The municipal tax rate is added as cantonal tax rate times the local tax multiplier. These data have been completed by Parchet (2014) to cover all municipalities in Switzerland.

We mainly use this information to define the low- and high-tax type municipalities. As the tax variation across municipalities within a canton is purely driven by the applicable tax multiplier this can be defined by comparing the municipal tax multiplier to the weighted average of tax multipliers that constitutes the tax multiplier used to calculate the special tax rate. Therefore we calculate the weighted tax multiplier as the average

[19] This regulation applies for the cantons Argau, Vaud, Geneva, Jura, Lucerne, Uri, Schwyz, Nidwalden, Glarus, Freiburg, Solothurn, Basel, Basel District and Schaffhausen.

[20] For the years 2001 to 2009, we use the weighting variable *IXPXH*, and for the years after 2010 *IXPXHJ*, since the annual dataset is used.

[21] Downloadable at http://www.estv.admin.ch/dokumentation/00075/00076/00720/

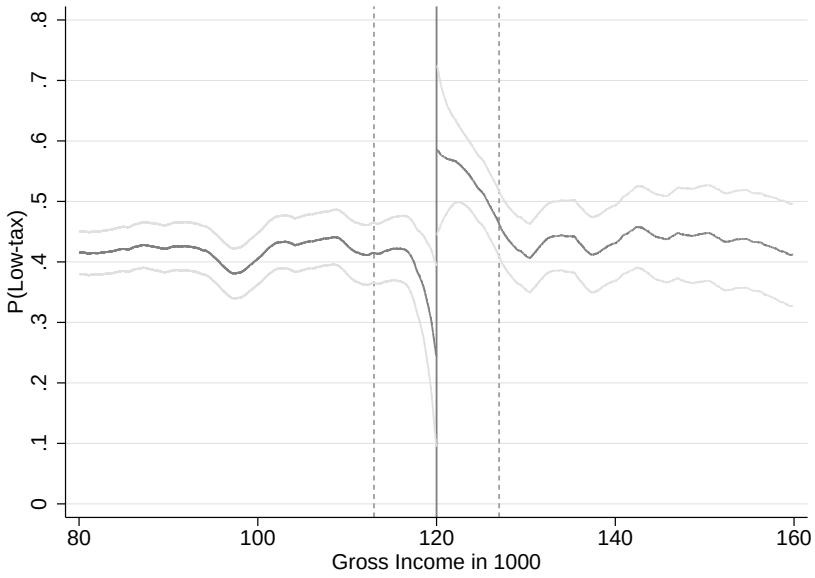


Figure 10: Probability of locating in a low-tax municipality. Foreigners not holding the permanent residence permit (C) during their first five years in Switzerland. Local linear regression, bandwidth $5k$ and taking sampling weights into account.

tax multiplier within a canton, weighted by the population share of the municipality to the cantonal population.[22]

- *Low-tax*: This is an indicator variable set to one if the tax multiplier in the ordinary tax scheme of the respective municipality is lower than the tax multiplier in the special tax scheme. These are thus those municipalities in which the special tax rate is higher than the ordinary one.

7 Results

7.1 Income Threshold

In this section, we study the behavior of individuals at the income threshold of 120,000 Swiss francs. To discriminate between Hypothesis 1 and 2 from Section 4, we first estimate the probability of observing a foreigner not holding the permanent residence permit in a low-tax municipality in the first five years of stay.

We estimate this relationship using a local linear regression (LLR), as described in the previous section. Figure 10 shows the resulting probability pattern for foreigners with temporary residence permit (B) during the first five years in Switzerland. There is a significant jump at the threshold ($p < 0.05$). However, it is not a jump above the income threshold of 120,000 Swiss francs, as postulated in the *mobility* response in

[22]The original tax multipliers for the special tax rate are not available for all cantons. Based on the reported definitions of local tax authorities, we compared this measure to the original one, where available, and it seems to be a valid approximation.

Table 2: RDD estimates for the probability of observing an individual in a low-tax municipality

p.	h	τ	se.	upper 95%	lower 95%	p-value	N_l	N_r
Basic sample								
x	5k	0.34	0.14	0.57	0.12	0.01	210	219
x	2.5k	0.40	0.19	0.72	0.08	0.04	93	111
x	10k	0.25	0.10	0.41	0.10	0.01	474	394
x	9.83k	0.26	0.10	0.42	0.10	0.01	469	385
Excluding income range of 7k around the threshold								
x	5k	0.15	0.13	0.37	−0.06	0.24	293	180
x	2.5k	0.22	0.18	0.53	−0.08	0.23	137	92
x	10k	0.03	0.09	0.18	−0.12	0.76	609	354
x	9.45k	0.03	0.09	0.19	−0.12	0.73	576	336

Notes: Sharp RDD estimates for the probability of observing an individual in a low-tax municipality. Estimates account for sampling weights.

Hypothesis 1. This is even more evident in Figure 11, where we exclude a region of 7,000 Swiss francs around the income threshold. The *levels* above and below the threshold are not systematically different as can be seen from the sharp regression discontinuity (RDD) estimates in Table 2.

Figure 10 shows a substantial fall in the probability of observing foreigners in low-tax municipalities just below the threshold and an increase just above it. This pattern corresponds exactly to the *income adjustment* as postulated in Hypothesis 2.

We conclude that there is no evidence of a migration response but evidence of income adjustment. We will next quantify the income response *conditional* on prior location choice by investigating whether individuals bunch at the preferential side of the income threshold, as described in Section 5.1.

We now study the bunching behavior around the income threshold implied by Hypothesis 2. Conditional on locating in a high-tax municipality, we would expect missing density mass above threshold and excess mass below. Figure 12 shows the estimated income distribution in high-tax municipalities for foreign individuals with temporary residence permit (B) during their first five years in Switzerland. We estimate the density function using nonparametric local likelihood density estimation separately below and above the income threshold, as described in Section 5. We see a clear negative jump in the density at the threshold. We also see an upward swing just to the left of the threshold, pointing to excess mass below the threshold, and a downward swing just to the right of the threshold, pointing to missing mass above the threshold. Foreigners in high-tax municipalities therefore systematically seek to keep their income below 120,000 Swiss francs to profit from the lower tax rates in the special tax regime. Figure 13 shows the income distribution in low-tax municipalities. We see a positive jump in the density at the income threshold, excess mass above and missing mass below. Foreigners in low-tax

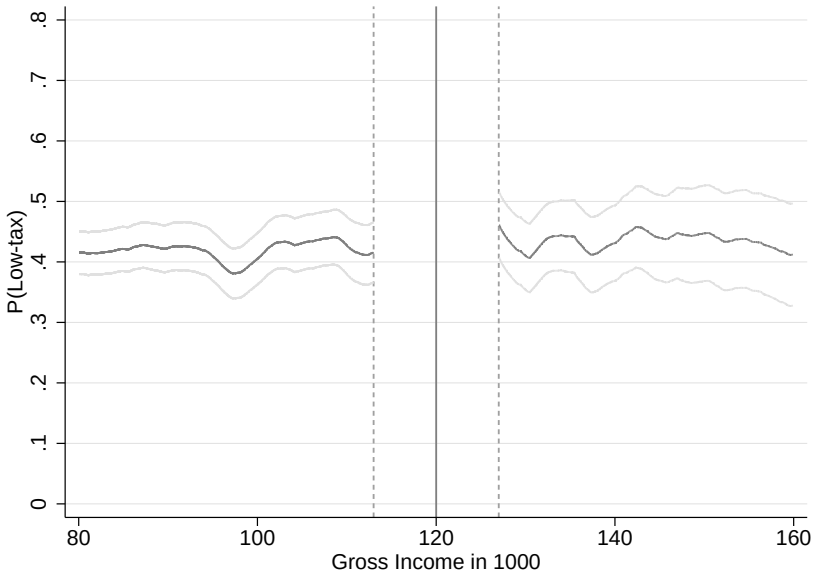


Figure 11: Probability of locating in a low-tax municipality, excluding manipulation range of $+/-7k$. Foreigners not holding the permanent residence permit (C) during their first five years in Switzerland. Local linear regression, bandwidth $5k$ and taking sampling weights into account.

municipalities therefore systematically seek to increase their income above 120,000 Swiss Francs to profit from the lower tax rates in the ordinary tax regime. These visual findings are exactly as predicted by Hypothesis 2. Tables 3 and Table 4 show the statistical test on the null hypothesis of no discontinuity at the threshold. The density just below the threshold is $\widehat{f}_l/\widehat{f}_r = 1.7$ times higher than just above it in high-tax municipalities, while it is 2.5 times higher above the threshold in low-tax municipalities. The empirical likelihood ratio test (see Section 5.1), shows that the discontinuity is statistically significant ($p < 0.01$) for low-tax municipalities but not for high-tax municipalities.

Recall that these findings cannot be attributed to tax evasion, as wages of workers in our sample are third-party reported to the tax authority (see, e.g., Kleven et al., 2011). There is also no room for measurement error in our sample, as wage income is provided by the social security administration.

7.1.1 Income adjustment and information acquisition

As our population of interest consists of foreigners, it is conceivable that those not speaking one of the three national languages are less well informed about the incentives that arise through this special taxation scheme. Those who speak one of the three major national languages in Switzerland (German, French and Italian) are likely better informed, as they have easier access to social networks. This is in line with Chetty et al. (2013), who found that such information disseminates through network peers. Foreigners speaking one of the national languages also have more leverage in wage negotiations, as they can generally better advocate for their interests

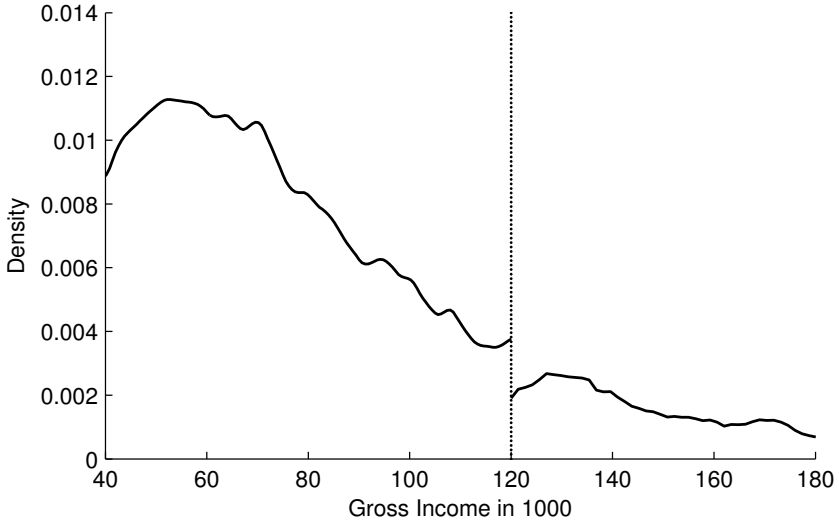


Figure 12: Income distribution in high-tax municipalities for foreigners not holding the permanent residence permit (C) in their first five years of stay in Switzerland. Local likelihood estimates using a bandwidth of $5k$ and taking the sampling weights into account. The corresponding density discontinuity estimates can be found in row I of Table 3.

Table 3: Density discontinuity estimates - High-tax municipalities

	c	h	$\hat{f}_l$	$\hat{f}_r$	$\hat{\theta}$	$\hat{lr}$	$\hat{f}_r/\hat{f}_l$	p-value	N
Whole sample									
I.	120	5	0.003 7	0.002 2	−0.001 5	1.716 5	0.583 1	0.190 1	7314
National language									
II.	120	5	0.004 1	0.001 5	−0.002 6	4.872 7**	0.355 4	0.027 3	5600
National language & not moved									
III.	120	5	0.004 4	0.001 7	−0.002 8	3.996 1**	0.378 3	0.045 6	4984
National language & age<34									
IV.	120	5	0.001 3	0.001 3	−0.000 0	0.001 0	0.979 1	0.974 9	2914
National language & age≥34									
V.	120	5	0.007 7	0.001 8	−0.005 9	6.493 7**	0.239 2	0.010 8	2686
National language & not changed job									
VI.	120	5	0.005 2	0.001 8	−0.003 4	4.860 9**	0.342 7	0.027 5	4404

Notes: Local likelihood ratio results for the discontinuity in the income distribution.
Sample specifications: Estimates I use the basic estimation sample of foreigners not holding the permanent residence permit in their first five years in Switzerland. The sample is reduced to those speaking a national language since estimates II. Estimates III are further reduced to those who have not moved within the last year. Estimates IV are produced using those younger than 34 years, and estimates V using those older than 34 years, respectively. Estimates VI use the sample of individuals who have not changed their job recently. N stands for the number of observations with regard to the observations available to estimate the whole density in the sample.
Significance levels: * $.05 < p < .1$, ** $.01 < p < .05$, *** $p < .01$.

Both channels predict that bunching is stronger for these individuals. Consequently, we repeat our estimation for the sub-sample of those speaking a national language. The income distributions in high- and low-tax municipalities are visualized in Figure A.1. The

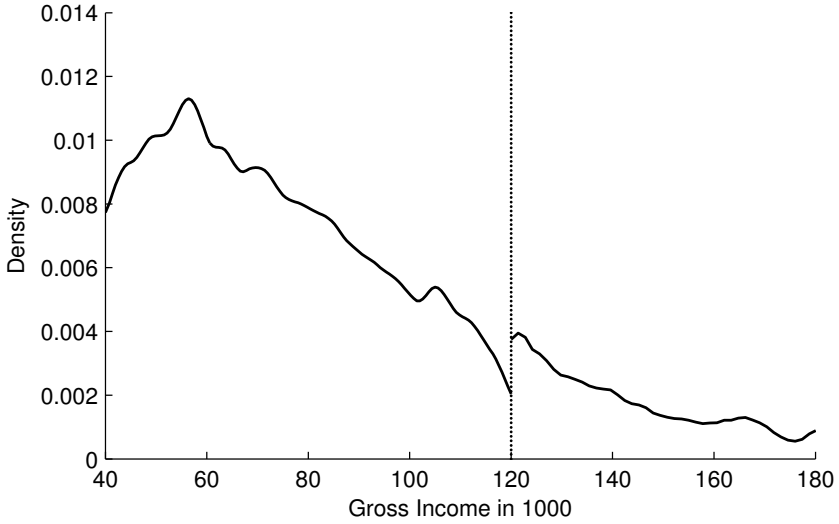


Figure 13: Income distribution in low-tax municipalities for foreigners not holding the permanent residence permit (C) in their first five years of stay in Switzerland. Local likelihood estimates using a bandwidth of $5k$ and taking the sampling weights into account. The corresponding density discontinuity estimates can be found in row I of Table 4.

Table 4: Density discontinuity estimates - Low-tax municipalities

	c	h	$\hat{f}_l$	$\hat{f}_r$	$\hat{\theta}$	$\hat{lr}$	$\hat{f}_r/\hat{f}_l$	p-value	N
Whole sample									
I.	120	5	0.001 5	0.003 8	0.002 3	9.073 6***	2.513 1	0.002 6	6173
National language									
II.	120	5	0.001 3	0.004 0	0.002 7	9.261 1***	3.131 1	0.002 3	4623
National language & not moved									
III.	120	5	0.001 3	0.004 3	0.003 0	8.491 5***	3.238 3	0.003 6	4060
National language & age<34									
IV.	120	5	0.001 0	0.000 9	−0.000 0	0.002 5	0.967 9	0.960 4	2117
National language & age≥34									
V.	120	5	0.001 6	0.007 4	0.005 8	10.557 5***	4.532 7	0.001 2	2506
National language & not changed job									
VI.	120	5	0.001 5	0.004 7	0.003 2	8.450 2***	3.175 1	0.003 7	3649

Notes: Local likelihood ratio results for the discontinuity in the income distribution. Sample specifications: Estimates I use the basic estimation sample of foreigners not holding the permanent residence permit in their first five years of stay in Switzerland. The sample is reduced to those speaking a national language since estimates II. Estimates III are further reduced to those who have not moved within the last year. Estimates IV are produced using those younger than 34 years, and estimates V using those older than 34 years, respectively. Estimates VI use the sample of individuals who have not changed their job recently. N stands for the number of observations with regard to the observations available to estimate the whole density in the sample.
Significance levels: * $.05 < p < .1$, ** $.01 < p < .05$, *** $p < .01$.

estimates of the discontinuities are reported in row II of Tables 3 and 4, respectively. The discontinuity is more pronounced and significantly different from zero in high-tax municipalities ($p < 0.05$), while it remains significant in low-tax municipalities ($p <$

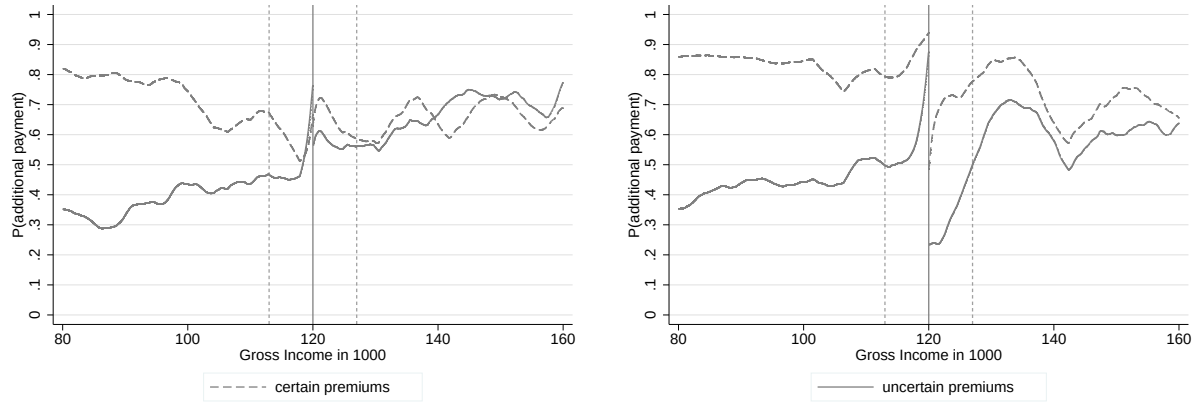
0.01). Note that significance is harder to achieve in these sub-samples due to the smaller sample size. We define this sample to be our baseline and condition the following results on those speaking a national language. It constitutes a rather homogeneous group of foreigners, who in principle could be informed about the incentives if there were no other optimization frictions, which will be addressed hereafter. Somewhat surprisingly, we find that income adjustment is more pronounced in low-tax municipalities, where adjustment implies effort and costs, than in low-tax municipalities, where bunching has low costs. Hence positive tax notches in our setting seem to produce lower bunching incentives than negative notches.

We next investigate whether information acquisition plays a role in income adjustment. It will likely take some time for foreigners to learn about the optimal tax regime for them in their location of residence. Recall that the dominated income region differs from municipality to municipality. Furthermore, income cannot be adjusted from one day to the next. If information acquisition is relevant, we would expect that income adjustment is stronger for foreigners who have lived in a municipality for some time. We therefore restrict our sample of foreigners speaking a national language further to those who have not moved within the last 12 months. The discontinuity estimates for high- and low-tax municipalities are reported in Tables 3 and 4, respectively. The discontinuity is slightly increased from a ratio of 3.1 to 3.2 in low-tax municipalities, while it is falls from 2.8 to 2.6 in high-tax municipalities. Again, the reaction tends to be stronger in low-tax municipalities.

Hypothesis 2 postulated that foreigners in high-tax municipalities with income just above 120,000 should unanimously reduce their income below the threshold. However, there might be a long-run cost to reducing income. Income is typically highly persistent over time (Oreopoulos et al., 2012; Kahn, 2010). Reducing income payment today can substantially reduce lifetime income. This lifetime effect is much more important for younger workers than for more senior workers. We test this by splitting our sample of foreigners speaking a national language into those aged below 34 (the median age in the estimation sample) and those aged 34 and above.

Rows IV and V in Tables 3 and 4 report the results for these two sub-samples. We find no income adjustment for young workers in either high- or low-tax municipalities while we find strong and significant effects for senior workers in high-tax municipalities ($p < 0.05$) and low-tax municipalities ($p < 0.01$).

The above finding could alternatively be explained by weaker bargaining power in the wage negotiations of younger workers. We will therefore next explore whether wage bargaining could play an important role in income adjustment. Individuals employed by the same employer for a longer time should arguably have a better bargaining position. While the initially agreed yearly wage might be determined by standards, they have time to find out how to adjust the flexible wage parts or the timing of money flows and to settle this with their employer. We therefore reduce the sample to those who have not



(a) High-tax municipalities **(b)** Low-tax municipalities

Figure 14: Probability of reporting to receive additional payments. Local linear smooth of the probability that someone receives additional payments for foreigners, not holding the permanent residence permit, speaking a national language, and in their first five years years of stay. The dashed line indicates the probability of receiving basically certain additional payments like a Christmas bonus, while the solid line indicates rather uncertain premiums like performance bonus and profit participation. Yearly gross income is measured in thousands of Swiss francs, and a bandwidth of $5k$ is used and sampling weights are taken into account.

changed their job recently. We find that, compared to the basic estimation of individuals speaking a national language,[23] bunching is amplified in both municipality types, while the ratio between the low and the high density side is still higher for the low-tax type (see row VI of Table 3 and Table 4 respectively). This suggests that it is wage bargaining with the employer through which the manipulation takes place.

7.1.2 Margins of income adjustment

Income adjustment through wage bargaining can occur via fixed wage components and via flexible wage components such as profit participation, Christmas bonuses (13th month's salary) or other bonus payments. Flexible wage components are ex-ante unknown and introduce a risk for individuals who want to target a specific income. A higher bonus than anticipated would risk the special tax regime of an individual in a high-tax municipality, while a lower bonus than anticipated would risk the ordinary tax regime of an individual in a low-tax municipality. We would therefore expect that individuals who target an income prefer the fixed wage over the flexible wage and bargain for a lower share of flexible wage components. Individuals in low-tax municipalities with income just above the threshold as well as individuals in high-tax municipalities with income just below the threshold should have a lower share of flexible wage components. In this respect, upside adjustment (in low-tax municipalities) is easier than downside adjustment (in high-tax municipalities). Individuals in low-tax municipalities seek to push their *certain* income above the threshold

[23]Unfortunately, we cannot estimate the effects in the sub-sample of those who have not recently changed their job because the sample size is too small.

by negotiating for more fixed income in exchange for less flexible income. Any remaining bonus would not impair their strategy. Individuals in high-tax municipalities seek to reduce their *certain* income below the threshold by negotiating for more fixed income in exchange for less flexible income. For them, any remaining bonus would impair their strategy and leave some uncertainty about the tax regime. We therefore expect to see a clearer effect in high-tax municipalities.

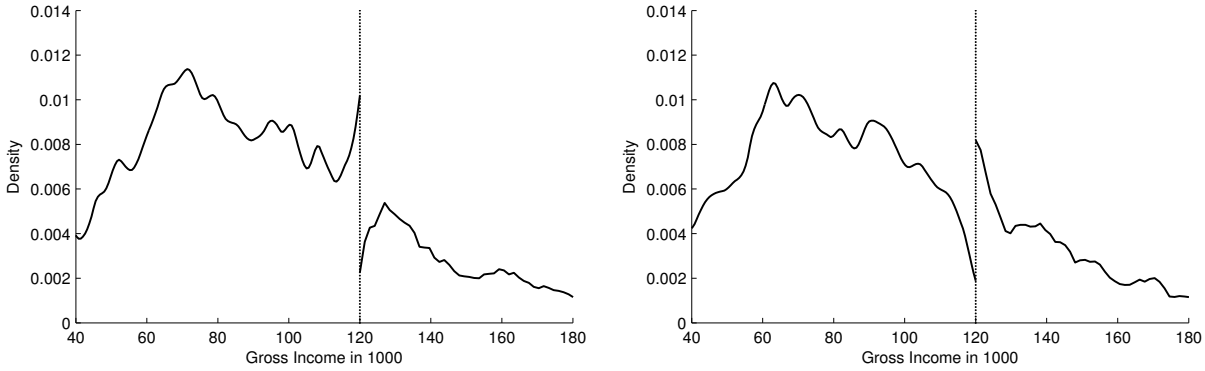
Figure 14 shows local linear smooth of the probability that a worker has received such additional payments beyond the base salary for foreigners with a temporary residence permit (B) during their first five years in Switzerland. The dashed line indicates additional payments that are basically certain, like the Christmas bonus, while the solid line indicates rather uncertain additional payments like performance premiums and profit participation. The resulting patterns are in line with our hypotheses formulated above. We see a distinct drop in the probability of receiving additional payments in the low-tax case, while the drop is less distinct in high-tax municipalities. To access whether this pattern can really be attributed to the tax scheme, we repeat the analysis for a quasi-placebo group, namely foreigners holding the permanent residence permit and residing in Switzerland for longer than seven years. We find no such pattern there, see Figure A.2 in the Appendix. We conclude that flexible wage components and bargaining over their height and division, seems to be an important channel for bunching.

Finally, we gather the most sensitive groups by conditioning on the characteristics that we have found to support bunching[24] and obtain the income distributions presented in Figure 15. The income distributions in both municipality types visually feature clear jumps with clear peaks at the advantageous sides of the threshold. Both discontinuity estimates presented in Table 5 are significantly different from zero at the 5% level and the ratio measure shows that the density at the advantageous side of the threshold is about four times higher than at the disadvantageous side. We interpret this as evidence that these groups are able to and do manipulate their income systematically and precisely.

7.1.3 Effects of immigrant networks

Chetty et al. (2013) present clear evidence that the neighborhood's knowledge about incentives plays an important role for individuals' reactions to these incentives. Devillanova (2008) shows that social networks significantly foster health care utilization of undocumented immigrants in Italy. Inspired by these findings, we expect that foreigners residing in municipalities with a high level of knowledge about the consequences of the two tax regimes should be more probable to respond. We use the municipal share of foreigners that already crossed one of the thresholds as a proxy for knowledge about the differences between the ordinary and the special tax regime. To construct a measure which is meant to capture the share of individuals in a municipality that are potentially informed about

[24]These are those older than 34 years, speaking a national language, residing in the same municipality longer than one year, and holding the same job for longer than one year.



(a) High-tax municipalities **(b)** Low-tax municipalities

Figure 15: Income distribution for foreigners not holding the permanent residence permit - most sensitive groups. Local likelihood estimates using $bandwidth = 5k$ and taking the sampling weights into account for the sample of the most sensitive groups, meaning age$>$ 34, speaking one of the national languages, not moved or changed job within the last year. The corresponding density discontinuity estimates can be found in rows I of Table 5.

Table 5: Density discontinuity estimates for the most sensitive groups

	c	h	$\hat{f}_l$	$\hat{f}_r$	$\hat{\theta}$	$\hat{lr}$	$\hat{f}_r/\hat{f}_l$	p-value	N
High-tax type									
I.	120	5	0.010 0	0.002 4	−0.007 6	5.995 2**	0.235 5	0.014 3	2028
I.	120	6	0.009 9	0.002 6	−0.007 3	5.633 8**	0.261 1	0.017 6	2028
I.	120	7	0.009 8	0.003 0	−0.006 8	4.856 3**	0.303 0	0.027 5	2028
I.	120	8.39	0.009 5	0.003 3	−0.006 1	4.103 5**	0.350 4	0.042 8	2028
Low-tax type									
I.	120	5	0.001 6	0.007 8	0.006 2	8.381 2***	4.915 8	0.003 8	1849
I.	120	6	0.001 9	0.008 2	0.006 3	9.652 7***	4.346 5	0.001 9	1849
I.	120	7	0.002 2	0.008 4	0.006 3	10.084 4***	3.869 7	0.001 5	1849
I.	120	11.87	0.003 0	0.008 9	0.005 9	10.229 3***	2.965 8	0.001 4	1849

Notes: Local likelihood ratio results for the discontinuity in the income distribution for the sample of the most sensitive groups, meaning age$>$ 34, speaking one of the national languages, not moved or changed job within the last year. N stands for the number of observations with regard to the observations available to estimate the whole density in the sample.
Significance levels: * $.05 < p < .1$, ** $.01 < p < .05$, *** $p < .01$.

the two taxation schemes, the evolving incentives, and can share this information, we first calculate $prop_{special}$, the share of foreigners within the municipality who have not crossed one of the thresholds (those not holding the permanent residence permit, residing in Switzerland for fewer than five years and having an income below the income threshold). To circumvent the possibility that our measure is confounded by bunching individuals, we further restrict the sample to those with an income below 113,000 or above 127,000 Swiss francs, and thus exclude the manipulation range. We then calculate the knowledge

Table 6: Density discontinuity estimates for high- and low-knowledge municipalities

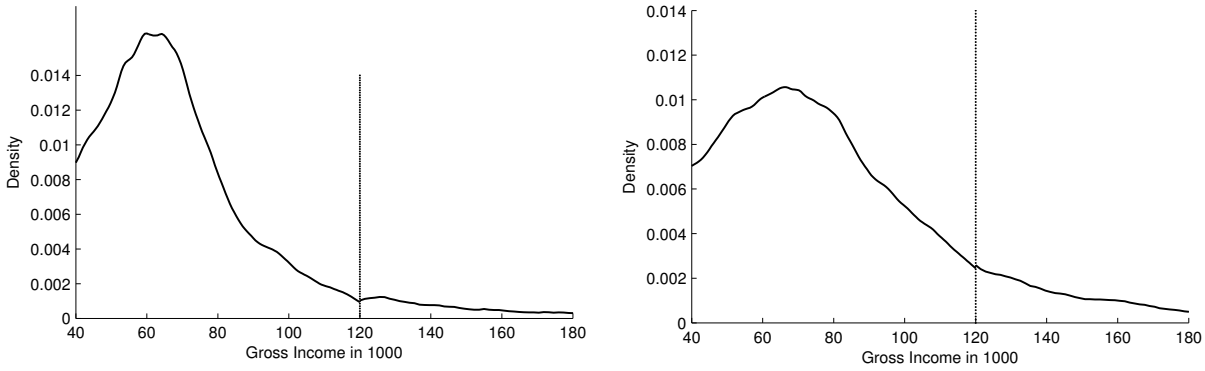
	c	h	$\hat{f}_l$	$\hat{f}_r$	$\hat{\theta}$	$\hat{lr}$	$\hat{f}_r/\hat{f}_l$	p-value	N
High-tax type									
knowledge<median									
I.	120	5	0.001 4	0.002 2	0.000 9	0.924 8	1.629 3	0.336 2	4027
I.	120	6	0.001 4	0.002 2	0.000 8	0.996 7	1.583 9	0.318 1	4027
I.	120	7	0.001 5	0.002 2	0.000 7	0.797 2	1.447 3	0.371 9	4027
I.	120	10.68	0.001 6	0.002 1	0.000 6	0.814 5	1.357 2	0.366 8	4027
knowledge≥median									
II.	120	5	0.006 6	0.002 2	−0.004 4	4.087 6**	0.335 3	0.043 2	3268
II.	120	6	0.006 9	0.002 1	−0.004 7	5.374 8**	0.308 1	0.020 4	3268
II.	120	7	0.007 0	0.002 2	−0.004 8	5.925 2**	0.310 8	0.014 9	3268
II.	120	10.82	0.005 9	0.002 0	−0.003 9	5.843 2**	0.343 3	0.015 6	3268
Low-tax type									
knowledge<median									
I.	120	5	0.001 1	0.001 5	0.000 4	0.392 2	1.362 3	0.531 2	3153
I.	120	6	0.001 2	0.001 7	0.000 5	0.509 6	1.386 7	0.475 3	3153
I.	120	7	0.001 2	0.001 9	0.000 6	0.925 9	1.517 1	0.335 9	3153
I.	120	12.64	0.001 2	0.002 6	0.001 4	5.038 2**	2.224 4	0.024 8	3153
knowledge≥median									
II.	120	5	0.002 0	0.006 4	0.004 4	8.400 7***	3.158 1	0.003 8	3005
II.	120	6	0.002 4	0.006 3	0.003 8	7.181 8***	2.579 7	0.007 4	3005
II.	120	7	0.002 8	0.006 3	0.003 5	6.071 2**	2.217 5	0.013 7	3005
II.	120	9.53	0.003 6	0.006 3	0.002 7	3.944 2**	1.734 1	0.047 0	3005

Notes: Local likelihood ratio results for the discontinuity in the income distribution. The samples for both municipality types are separated at the median of the generated knowledge measure. N stands for the number of observations with regard to the observations available to estimate the whole density in the sample.
Significance levels: * $.05 < p < .1$, ** $.01 < p < .05$, *** $p < .01$.

measure as $1 - prop_{special}$.[25] We would expect bunching to be stronger in communities with a higher knowledge measure. To test this, and to retain a comparable sample size, we separate the sample at the median knowledge measure of the respective sample.[26] Table 6 summarizes the results. In line with the findings of Chetty et al. (2013), we find that there are only persistent discontinuities in municipality with an above median knowledge for both municipality types. This suggests that networks play a role in the information transmission and that the knowledge of neighbors plays a major role for the reactions to the arising incentives for strategic income adjustment.

[25]Please note that we calculate this measure from our sample of observations as there is no official information available. Further, we calculate the share accounting for the survey weights.

[26]The median turns out to be 68% for high- and low-tax municipalities in the income adaption sample and 70% in the duration threshold sample.



(a) High-tax municipalities

(b) Low-tax municipalities

Figure 16: Income distribution of Swiss residents. Local likelihood estimates using $bandwidth = 5k$ and taking the sampling weights into account. The corresponding density discontinuity estimates can be found in rows II of table A.4 in the Appendix.

7.1.4 Robustness

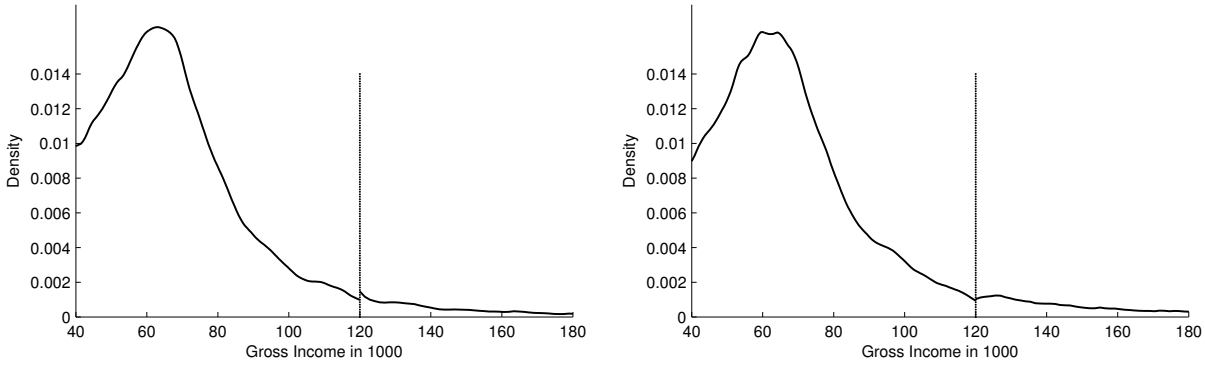
In order to validate that our findings are indeed due to the change in the tax regime we propose several robustness checks. First, we include estimates of discontinuities in the income distribution at two placebo thresholds for the basic sample, see row VII and VIII in Table A.3 in the Appendix. We find no significant discontinuities in either low-tax or high-tax municipalities at 100,000 or 140,000 Swiss francs.

Second, we choose Swiss citizens as a valuable quasi-placebo group, as the regime used is not in force for them. If there is no other regulation using the same income threshold, and we are not aware of another, our findings should be due to the change of the taxation regime for foreigners. The placebo test shown in Figure 16, for the sample of Swiss nationals, features no discontinuity in the income distribution at the income threshold of 120,000 Swiss francs for the population which is not affected by the special tax scheme either for low-tax or for high-tax municipalities.

We repeat the procedure for foreign residents already holding the permanent residence permit (C) and thus already taxed ordinarily who have lived in Switzerland for at least seven years, see Figure 17. These individuals have no incentive to manipulate their income at the income threshold of 120,000 Swiss francs. Again, we do not find a discontinuity in their income distribution (see Table A.4).

To validate that our findings are not sensitive to the bandwidth choice, we repeat the former analysis for different bandwidths and find that the basic pattern persists; results are reported in Tables A.2 and A.3 in the Appendix, where the last row reports results for the optimal bandwidth choice for the sample.

Summing up, we find that individuals react to the incentives generated by the two tax regimes by strategic income adjustment. Our results suggest that information, learning and bargaining ability play an important role in the behavioral reactions to tax incentives. Speaking one of the national languages and living in the same municipality for a longer



(a) High-tax municipalities **(b)** Low-tax municipalities

Figure 17: Income distribution of foreigners holding a C permit and living in Switzerland for at least 7 years. Local likelihood estimates using $bandwidth = 5k$ and taking the sampling weights into account. The corresponding density discontinuity estimates can be found in rows I of table A.4 in the Appendix.

time amplify, bunching which is in line with the findings of, e.g., Chetty and Saez (2013) and Chetty et al. (2013). We also find that knowledge within one's neighborhood plays a major role in the behavioral reactions to the investigated tax scheme. We find bunching behavior both at positive tax notches (high-tax municipalities) and at negative tax notches (low-tax municipalities). We find sharp and precise income adjustment for wage earners, while there is little evidence about strategic income responses of wage earners in the bunching literature (see, e.g., Saez et al., 2012; Chetty et al., 2013; Spencer and Selin, 2014). Most studies do not observe sharp bunching of wage earners, but rather a diffuse excess mass around the threshold, which is mostly attributed to the missing opportunity for misreporting. Note again that our findings cannot be attributed to tax evasion or misreporting as wages of workers in our sample are third-party reported to the tax authority (see, e.g., Kleven et al., 2011).

7.2 Duration Threshold

As described in the theoretical considerations in Section 4, the duration threshold generates incentives for tax induced mobility. Individuals located in high-tax municipalities when receiving the permanent residence permit (C) have an incentive to move, as they would experience a tax increase otherwise.

Thus, we would expect a positive jump in the probability of moving for those who receive the permanent residence permit (C) in a high-tax municipality, as stated in Hypothesis 3. This effect should rise with income (see Section 4). We test this jump for foreigners receiving the permanent residence permit after five years of stay using a fuzzy RD (see Section 5.2).

We first check the RDD assumption of a jump in the probability of treatment, the first stage, by looking at the probability of treatment at the duration threshold. Figure A.3 validates this assumption as it shows a clear discontinuous jump in the probability of

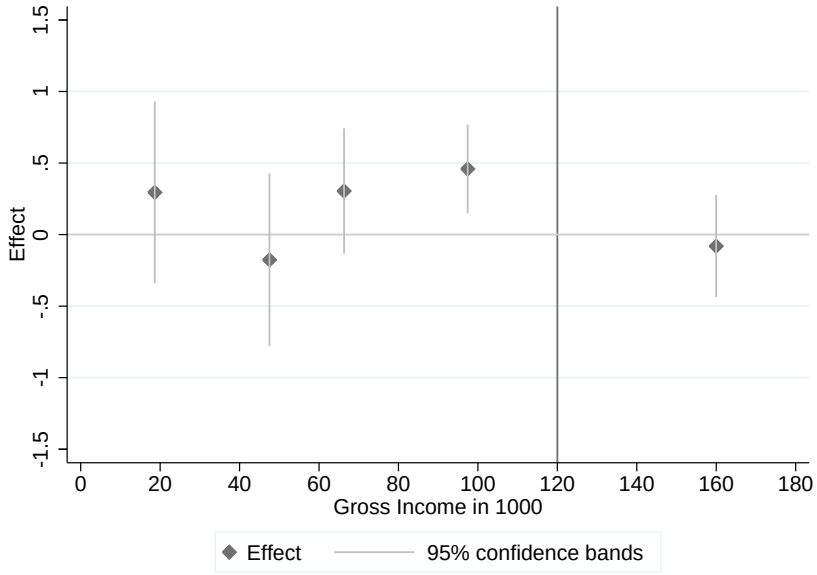


Figure 18: Jump in the probability of moving for residents in a high-tax municipality. Foreigners located in a high-tax municipality when receiving the permanent residence permit after 5 years in Switzerland. Fuzzy RD estimates using a bandwidth of 730 days (Table A.5 in the Appendix).

holding permit C at the threshold for both high- and low-tax municipalities. This also holds for high income individuals (see A.4 for the highest income quartile below the income threshold). This very much speaks for the monotonicity assumption, as incentives to accept the treatment seem to work in one direction.

When looking at the second stage separately, as if it were a sharp design for the highest income quartile below the income threshold of 120,000 Swiss francs, we observe a positive jump in the probability of moving in the high-tax case, while there is none in the low-tax type (see Figure A.5).

We estimate the effects separately for four different income quantiles below 120,000 Swiss francs. We additionally include a placebo-group of foreigners whose income exceeds 120,000 Swiss francs and who thus are already taxed ordinarily. If the observed mobility can be attributed to the shift in the tax regime, we should not observe an effect for those who are already taxed ordinarily. Foreigners in high-tax municipalities experience an increase in their tax liability after receiving a permanent residence permit (C). Figure 18 shows the fuzzy RD estimates for foreigners located in a high-tax municipality when receiving the permanent residence permit after five years in Switzerland.

Consistent with Hypothesis 3, we find that high income foreigners in these municipalities are indeed very likely to move once they receive the permanent residence permit. The probability of moving jumps, e.g., by about 33 to 62 percentage points for the highest income quartile below 120,000 Swiss francs. Further, we do not observe any systematic change either for the income group above the income threshold or for the three lower income quartiles.

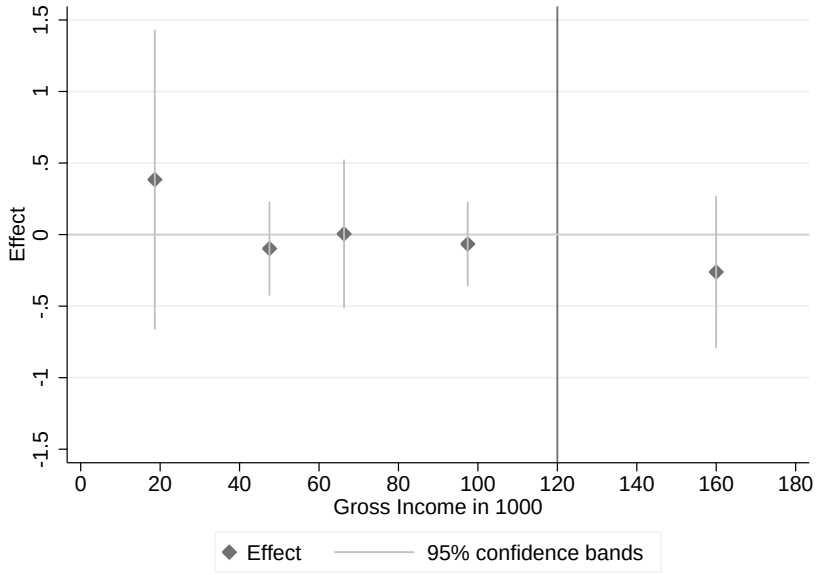


Figure 19: Jump in the probability of moving for residents in a low-tax municipality. Foreigners located in a low-tax municipality when receiving the permanent residence permit after 5 years in Switzerland. The sample is additionally reduced to those speaking one of the national languages. Fuzzy RD estimates using a bandwidth of 730 days (Table A.7 in the Appendix).

It seems reasonable to assume that individuals are not perfectly informed when they enter the country and information is costly, thus the effect should be stronger for those individuals who speak one of the national languages. As argued before, they should have easier access to information and social networks. Additionally, those speaking a national language could have better chances of finding new housing and have better opportunities on the labor market, which would make them more mobile. Another possibility would be that their language skills make them less mobile because they integrate faster and thus build up social ties that prevent them from changing their location because of taxes. The results for a sample reduced to those speaking one of the national languages are reported in Table A.6 in the Appendix. The effects for this subsample are indeed higher in magnitude (between 39 and 79 percent), which is again in line with the findings of Chetty et al. (2013), who find that information works through network peers and with the reasoning that language skills enhance the ability to respond to tax incentives.

Foreigners in low-tax municipalities experience a decrease in their tax liability after receiving a permanent residence permit. These residents therefore have weak incentives to move after they obtain their permanent residence permit. They face a poorer choice set as they start in the lower part of the tax distribution, even though those starting in the upper part of the tax distribution for low-tax municipalities might have some incentives to move when reconsidering their location choice, as there would be municipalities with lower tax rates available. Thus, we would expect no or only a weaker positive jump in the probability of moving for those who receive the permanent residence permit at the

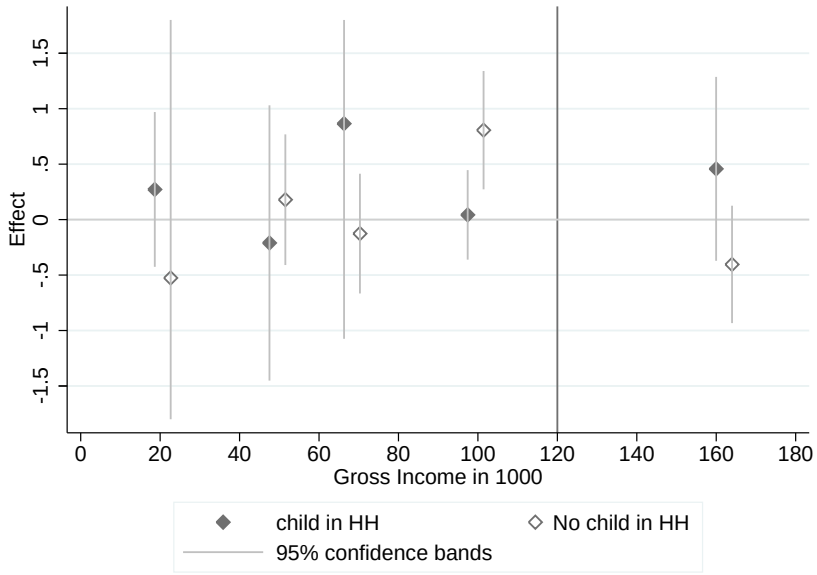


Figure 20: Jump in the probability of moving for residents in a high-tax municipality. Foreigners located in a high-tax municipality when receiving the permanent residence permit after 5 years in Switzerland. The sample is additionally reduced to those speaking one of the national languages and separated into households with and without a child. Fuzzy RD estimates using a bandwidth of 730 days (Table A.8 and Table A.9 in the Appendix).

threshold while residing in a low-tax municipality. We repeat the preceding fuzzy RD estimation for individuals located in low-tax municipalities, while we report the estimates for the sample of foreigners speaking a national language, as we would expect the effect to be most pronounced in this group. Results are presented in Figure 19 and Table A.7. Consistent with Hypothesis 3, we find that there is no systematic jump in the probability of moving for foreigners in these municipalities.

We would further expect that mobility reactions of households with children are less likely. Children can increase social bonds, and parents might hesitate to change their school or kindergarten once they settled in a municipality. Mobility is arguably more costly for households with children. Testing this for those individuals located in high-tax municipalities, for whom we have seen that there is a mobility reaction, we find that the effect seems to be primarily driven by households without children. Results are summarized in Figure 20. Estimation results can be found in Table A.8 and A.9.

The change in tax rates induced by switching from the special to the ordinary tax regime is heterogeneous across municipalities. Hence, the treatment intensity is heterogeneous. We therefore test whether the jump in the moving probabilities is higher if the expected (or experienced) tax rise is stronger. We build two groups within the two municipality types, one for which the tax change is below the median tax change and one where it is above the median. We indeed find strong effects in the very high high-tax municipalities and virtually no effect in the less high-tax municipalities. Results are

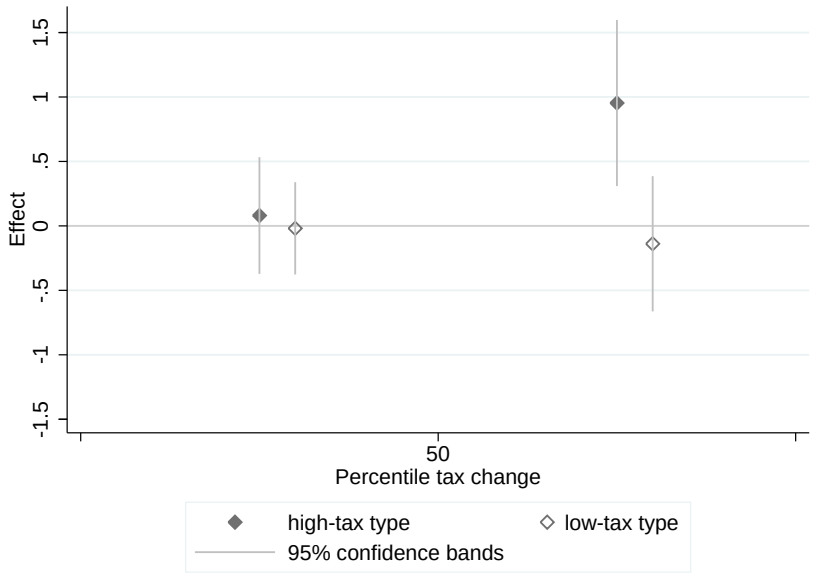


Figure 21: Effects for tax change below and above median. Fuzzy RD estimates of discontinuity in the probability of moving for foreigners in the highest income quartile below the income threshold of 120,000 Swiss francs. The sample is split with regard to the municipality types and by the expected tax change when staying in the municipality after receiving the permanent residence permit. The sample is additionally reduced to those speaking one of the national languages. Fuzzy RD estimates using a bandwidth of 730 days (Table A.10 in the Appendix).

summarized in Figure 21 and Table A.10.[27] Consistent with Hypothesis 3, we find again no systematic effect for foreigners residing in low-tax municipalities when receiving the permanent residence permit, although those with a tax change above the median would have an incentive to move to a municipality with even lower taxes.

Based on the reasoning in the income adjustment analysis we would also expect individuals located in municipalities with a high knowledge measure to be more likely to react to the generated incentives. There are potentially more individuals who can advise them where the rise in taxes comes from and how they could respond. We thus again separate the sample of the highest income quartile below the income threshold at the median knowledge measure. If knowledge plays a role, individuals in high knowledge areas should move away from high-tax municipalities with greater likelihood when receiving permit C. This is exactly what we find. Estimation results are reported in Table A.11.

In order to check for the robustness of our results, we also report parametric FRD estimates of our basic estimates using standard instrumental variable estimation[28] incorporating survey weights. We follow Lee and Lemieux (2010) and Angrist and Pischke (2009) in the specification, and use all relevant observations of individuals with a duration of stay in Switzerland below 10 years, because otherwise the estimation window would be

[27]Tax changes are coded such that a higher value represents a stronger tax change in the positive (high-tax) or the negative (low-tax) direction.

[28]More precisely, we use a two stage least squares, stata's *ivregress 2sls* function.

highly unsymmetric. Results of estimates using first and second order polynomials can be found in Table A.12 and Table A.13.[29] In general we find that parametrically estimated effects are somewhat lower, which could be attributed to the usage of the longer timespan included. However, the pattern persists.[30]

Finally, to validate local randomization and exclude that individuals sort with regard to their duration of stay in Switzerland, we perform the test for a discontinuity of the density of durations as we have estimated it in the analysis of the income threshold. Estimates for the density discontinuity at the five year duration threshold are reported in Table A.17 and are visualized in Figure A.6. We find no striking discontinuities at the duration threshold and thus conclude that there does not seem to be any indication that local randomization is violated. To further validate that our findings are due to the duration threshold we repeat the estimates for the high-tax municipalities for two placebo-thresholds at four and six years of stay in Switzerland. Results are reported in Tables A.15 and A.16. We find no jump in the probability at these placebo-thresholds. This further underpins the causal interpretation of our findings.

Summing up, we find that local tax differences can induce intra-national mobility for high income foreigners in Switzerland. There are no significant effects for incomes below the fourth income quartile (below 120,000 Swiss francs). Individuals tend to move when they would experience a tax rise and do not move if they expect a tax drop, while this holds even if there were a municipality with lower taxes in their choice set. We further find that this effect is driven by individuals without children and by those who would experience a higher tax rise and thus if returns from relocation are higher. Again, the knowledge about the tax scheme change and the incentives produced seem to play a major role in individuals' behavioral reactions.

8 Conclusion

We study behavioral responses to local income taxes exploiting a special tax regime which applies to foreign employees residing in Switzerland. Using administrative income data and exploiting thresholds generated by the institutional setting, we are able to draw causal inference on income adjustment to preferential tax regimes (income bunching) and on tax induced intra-national mobility within a quasi-random setting.

The institutional setting used generates two thresholds: foreigners with income below 120,000 Swiss francs (income threshold) and less than 5 years of stay (duration threshold) are subject to a special tax regime. For taxpayers in the ordinary tax regime, income tax rates differ across individual municipalities within cantons, while taxpayers in the special tax regime pay a single rate within each canton. The single tax rate is the weighted average

[29] A specification controlling for individual characteristics can be found in Table A.14.

[30] For these estimates we follow the specification in Angrist and Pischke (2009)[p.261ff] where we choose the polynomial order to be one or two.

of the ordinary local tax rate. The approximately 2600 municipalities in Switzerland can therefore be grouped into two types: high-tax municipalities where the ordinary tax rate is higher than the special tax rate and low-tax municipalities where the ordinary tax rate is lower than the special tax rate. Low-tax municipalities are, ceteris paribus, attractive to foreigners in the ordinary tax regime, while high-tax municipalities are attractive to foreigners in the special tax regime.

We find that foreign individuals do not systematically choose their initial location in Switzerland based on the tax regime associated with their income. Conditional on their location, however, foreign employees systematically increase or decrease their gross income depending on what leads to a lower tax liability. We propose and apply a novel discontinuity in density design to statistically test for such income adjustments (bunching). We find that upward income adjustments at negative tax notches are more pronounced than downward income adjustments at positive tax notches. We also find evidence that the observed income adjustment is driven by negotiating a lower share of uncertain flexible income components. In contrast to most studies investigating income bunching, we find that there is very precise income bunching for wage earners, although their income is third-party reported.

Applying a fuzzy regression discontinuity design (RDD) to the duration threshold, we find that local tax differences induce intra-national mobility of foreigners in Switzerland. Individuals who experience a tax rise when switching from the special to the ordinary tax scheme tend to relocate with higher probability than those individuals who experience a drop in tax rates. The observed mobility rise is mainly driven by individuals without children in the household and those who face a higher tax rise.

In line with the findings of Chetty et al. (2013), we find that knowledge and information transmission through local networks plays a major role in the behavioral response to tax incentives estimated at both thresholds.

As in all RDD studies, our results are *local* estimates for a particular group. In our case we estimate effects for high-income foreign workers during their first years in Switzerland. This group is actually a very interesting group for the questions at hand. First, we can learn about the importance of information, as foreigners are not fully informed when entering the country. Second, immigrants are a rather mobile group, because of the absence of social bonds and ties. Consequently, their reactions could be more unconfounded by factors that otherwise prevent people from moving for tax reasons. In this case our estimates of tax induced mobility could be seen as an upper bound for the reactions of native residents – in some sense the pure response.

References

Angrist, J. D. and J.-S. Pischke (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton NJ: Princeton University Press.

Basten, C., M. v. Ehrlich, and A. Lassmann (2014). Income Taxes, Sorting, and the Costs of Housing: Evidence from Municipal Boundaries in Switzerland. *KOF Working Papers, KOF Swiss Economic Institute, ETH Zurich, No. 362*.

BFS (2004). Die schweizerische Arbeitskräfteerhegung (SAKE). Konzepte-Methodische Grundlagen-Praktische Ausfürung. *Bundesamt für Statistik (BFS). Neuchâtel*.

BFS (2012). SAKE-Daten Benutzerhandbuch. *Bundesamt für Statistik (BFS). Neuchâtel*.

Bokil, H., Andrews Peter, J. Kulkarni, S. Mehta, and P. P. Mitra (2010). Chronux: A Platform for Analyzing Neural Signals. *Journal of Neuroscience Methods 192*(1), 146–151.

Calonico, S., M. D. Cattaneo, and R. Titiunik (2013). Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. *Econometrica: Journal of the Econometric Society 82*(6).

Chetty, R., J. N. Friedman, T. V. Olsen, and L. Pistaferri (2011). Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records. *The Quarterly Journal of Economics 126*(2), 749–804.

Chetty, R., J. N. Friedman, and E. Saez (2013). Using Differences in Knowledge Across Neighborhoods to Uncover the Impacts of the EITC on Earnings. *American Economic Review 103*(7), 2683–2721.

Chetty, R. and E. Saez (2013). Teaching the Tax Code: Earnings Responses to an Experiment with EITC Recipients. *American Economic Journal: Applied Economics 5*(1), 1–31.

Devillanova, C. (2008). Social Networks, Information and Health Care Utilization: Evidence from Undocumented Immigrants in Milan. *Journal of Health Economics 27*(2), 265–286.

Fan, J. and I. Gijbels (1996). *Local Polynomial Modelling and its Applications*, Volume 66 of *Monographs on Statistics and Applied Probability*. London: Chapman and Hall.

Feld, L. (2000). *Steuerwettbewerb und seine Auswirkungen auf Allokation und Distribution: Ein Überblick und eine empirische Analyse für die Schweiz*, Volume 10. Tübingen: Mohr (Siebeck).

Feld, L. and G. Kirchgässner (2001). Income tax Competition at the State and Local Level in Switzerland. *Regional Science and Urban Economics 31*(2), 181–213.

Feldstein, M. (1995). The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1986 Tax Reform Act. *Journal of Political Economy 103*(3), 551.

Feldstein, M. (1999). Tax Avoidance and the Deadweight Loss of the Income Tax. *The Review of Economics and Statistics 81*(4), 674–680.

Hahn, J., P. Todd, and W. van der Klaauw (2001). Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design. *Econometrica 69*(1), 201–209.

Harms, T. and P. Duchesne (2010). On Kernel Nonparametric Regression Designed for Complex Survey Data. *Metrika 72*(1), 111–138.

Imbens, G. and J. Angrist (1994). Identification and Estimation of Local Average Treatment Effects. *Econometrica: Journal of the Econometric Society*, 467–475.

Imbens, G. and T. Lemieux (2008). Regression Discontinuity Designs: A Guide to Practice. *Journal of Econometrics 142*(2), 615–635.

Kahn, L. B. (2010). The Long-Term Labor Market Consequences of Graduating from College in a Bad Economy. *Labour Economics 17*(2), 303–316.

Kitamura, Y. (2008). Empirical Likelihood Methods in Econometrics: Theory and Practice. *Cowles Foundation Discussion Paper No. 1254*.

Kleven, H., M. B. Knudsen, C. T. Kreiner, S. Pedersen, and E. Saez (2011). Unwilling or Unable to Cheat? Evidence from a Tax Audit Experiment in Denmark. *Econometrica 79*(3), 651–692.

Kleven, H., C. Landais, E. Saez, and E. Schultz (2014). Migration and Wage Effects of Taxing Top Earners: Evidence from the Foreigners' Tax Scheme in Denmark. *The Quarterly Journal of Economics 129*(1), 333–378.

Kleven, H. and M. Waseem (2013). Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan. *The Quarterly Journal of Economics 128*(2), 669–723.

Kleven, H. J., C. Landais, and E. Saez (2013). Taxation and International Migration of Superstars: Evidence from the European Football Market. *The American Economic Review 103*(5), 1892–1924.

Lee, D. and T. Lemieux (2010). Regression Discontinuity Designs in Economics. *Journal of Economic Literature 48(2)*, 281–355.

Liebig, T. and A. Sousa-Poza (2006). The Influence of Taxes on Migration: Evidence from Switzerland. *Cambridge Journal of Economics 30*(2), 235–252.

Loader, C. R. (1996). Local Likelihood Density Estimation. *The Annals of Statistics 24*(4), 1602–1618.

McCrary, J. (2008). Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test. *Journal of Econometrics 142*(2), 698–714.

Moore, J., L. Stinson, and E. Welniak (2000). Income Measurement Error in Surveys: A Review. *Journal of Official Statistics 16*(4), 331–362.

Morger, M. (2013). What Do Immigrants Value Most About Switzerland? Evidence of the Relative Importance of Income Taxes. *CESifo Working Paper No. 4134*.

Oreopoulos, P., v. T. Wachter, and A. Heisz (2012). The Short-and Long-Term Career Effects of Graduating in a Recession. *American Economic Journal: Applied Economics 1*(1), 1–29.

Otsu, T., K.-l. Xu, and Y. Matsushita (2013). Estimation and Inference of Discontinuity in Density. *Journal of Business & Economic Statistics 31*(4), 507–524.

Owen, A. B. (2001). *Empirical Likelihood*, Volume 92 of *Monographs on Statistics and Applied Probability*. Boca Raton and Fla: Chapman and Hall.

Parchet, R. (2014). Are Local Tax Rates Strategic Complements or Strategic Substitutes? *IdEP Economic Papers, University of Lugano* (2014/07).

Porter, J. (2003). Estimation in the Regression Discontinuity Model. *Unpublished Manuscript, Department of Economics, University of Wisconsin at Madison*.

Saez, E. (2010). Do Taxpayers Bunch at Kink Points? *American Economic Journal: Economic Policy 2*(3), 180–212.

Saez, E., J. Slemrod, and S. Giertz (2012). The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review. *Journal of Economic Literature 50*(1), 3–50.

Schaltegger, C., F. Somogyi, and J. Sturm (2011). Tax Competition and Income Sorting: Evidence from the Zurich Metropolitan Area. *European Journal of Political Economy 27*, 455–470.

Schmidheiny, K. (2006a). Income Segregation and Local Progressive Taxation: Empirical Evidence from Switzerland. *Journal of Public Economics 90*(3), 429–458.

Schmidheiny, K. (2006b). Income Segregation from Local Income Taxation When Households Differ in Both Preferences and Incomes. *Regional Science and Urban Economics 36*(2), 270–299.

Slemrod, J. (2010). Buenas Notches: Lines and Notches in Tax System Design. *Unpublished Manuscript, Department of Economics, University of Wisconsin at Madison*.

Spencer, B. and H. Selin (2014). Bunching and Non-Bunching at Kink Points of the Swedish Tax Schedule. *Journal of Public Economics 109*, 36–49.

Stadelmann, D. and S. Billon (2012). Capitalization of Fiscal Variables and Land Scarcity. *Urban Studies 49*(7), 1571–1594.

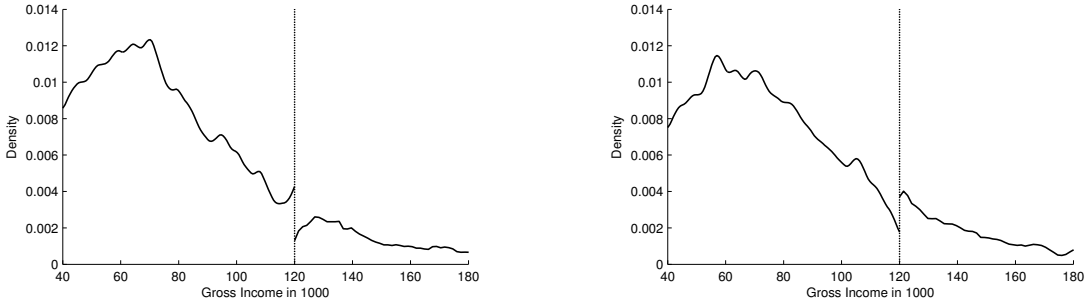
Young, C. and C. Varner (2011). Millionaire Migration and State Taxation of Top Incomes: Evidence from a Natural Experiment. *National Tax Journal 64*(2), 255–284.

A Appendix

Table A.1: Descriptive Statistics

Variable	Mean	Se	Median	Min	Max	N
SESAM Dataset sample						
Gross Income in 1000	87.8	293.9	67.6	0	42317.35	24157
Duration in 100	17.3	8.84	15.9	3.65	36.5	24157
Permit B	0.28	0.45	0	0	1	24157
Move	0.1	0.29	0	0	1	24157
Nat. Language	0.77	0.42	1	0	1	24157
Age	35.42	8.69	34	16	78	24157
Tax Data in sample						
Low-Tax	0.45	0.5	0	0	1	24157
$\Delta_{ord-special}$ (low-tax)	-0.05	0.05	-0.04	-0.47	-0.00	10914
$\Delta_{ord-special}$ (high-tax)	0.04	0.03	0.04	0.00	0.32	13243
Tax Data (municipalities 2001-2012)						
Low-Tax	0.37	0.48	0	0	1	29257
$\Delta_{ord-special}$ (low-tax)	-0.05	0.05	-0.04	-0.49	-0.00	10764
$\Delta_{ord-special}$ (high-tax)	0.05	0.04	0.04	0.00	0.32	18493

Notes: Summary statistics for the main variables. The summary statistics for the sample are based on the basic sample and the most relevant range of duration of stay (up to 10 years of stay in Switzerland). $\Delta_{ord-special}$ stands for the difference between the ordinary and special tax scheme measured in percent of the special tax rate at an income of 120,000 Swiss francs.



(a) High-tax municipalities **(b)** Low-tax municipalities

Figure A.1: Income distribution for foreigners not holding the permanent residence permit - speaking national language. Local likelihood estimates using $bandwidth = 5k$ and taking the sampling weights into account. Foreigners not holding the permanent residence permit in their first five years of stay in Switzerland, speaking one of the national languages. The corresponding density discontinuity estimates can be found in row II of Table 3 and Table 4, respectively.

Table A.2: Density discontinuity estimates - High-tax municipalities

	c	h	$\hat{f}_l$	$\hat{f}_r$	$\hat{\theta}$	$\hat{lr}$	$\hat{f}_r/\hat{f}_l$	p-value	N
Whole sample									
I.	120	5	0.003 7	0.002 2	−0.001 5	1.716 5	0.583 1	0.190 1	7314
I.	120	6	0.003 7	0.002 1	−0.001 6	2.198 8	0.567 1	0.138 1	7314
I.	120	7	0.003 8	0.002 1	−0.001 6	2.556 9	0.566 9	0.109 8	7314
I.	120	12.21	0.003 4	0.002 0	−0.001 3	2.927 3*	0.607 3	0.087 1	7314
National language									
II.	120	5	0.004 1	0.001 5	−0.002 6	4.872 7**	0.355 4	0.027 3	5600
II.	120	6	0.004 1	0.001 5	−0.002 6	5.250 4**	0.364 4	0.021 9	5600
II.	120	7	0.004 1	0.001 6	−0.002 5	5.095 3**	0.392 9	0.024 0	5600
II.	120	9.17	0.003 8	0.001 8	−0.002 1	4.104 6**	0.460 7	0.042 8	5600
National language & not moved									
III.	120	5	0.004 4	0.001 7	−0.002 8	3.996 1**	0.378 3	0.045 6	4984
III.	120	6	0.004 4	0.001 7	−0.002 7	4.243 9**	0.389 4	0.039 4	4984
III.	120	7	0.004 4	0.001 9	−0.002 6	3.992 6**	0.420 8	0.045 7	4984
III.	120	11.27	0.003 6	0.002 0	−0.001 7	2.342 4*	0.547 3	0.125 9	4984
National language & age<34									
IV.	120	5	0.001 3	0.001 3	−0.000 0	0.001 0	0.979 1	0.974 9	2914
IV.	120	6	0.001 4	0.001 2	−0.000 2	0.051 8	0.867 4	0.820 0	2914
IV.	120	7	0.001 5	0.001 1	−0.000 3	0.168 7	0.787 5	0.681 3	2914
IV.	120	11.24	0.001 0	0.000 9	−0.000 1	0.042 2	0.901 5	0.837 3	2914
National language & age≥34									
V.	120	5	0.007 7	0.001 8	−0.005 9	6.493 7**	0.239 2	0.010 8	2686
V.	120	6	0.007 6	0.002 0	−0.005 6	6.157 1**	0.264 8	0.013 1	2686
V.	120	7	0.007 6	0.002 3	−0.005 3	5.385 9**	0.306 8	0.020 3	2686
V.	120	8.87	0.007 3	0.002 7	−0.004 6	4.470 6**	0.365 8	0.034 5	2686
National language & not changed job									
VI.	120	5	0.005 2	0.001 8	−0.003 4	4.860 9**	0.342 7	0.027 5	4404
VI.	120	6	0.005 1	0.001 9	−0.003 3	4.933 8**	0.361 8	0.026 3	4404
VI.	120	7	0.005 2	0.002 1	−0.003 1	4.587 3**	0.397 2	0.032 2	4404
VI.	120	10.25	0.004 5	0.002 3	−0.002 2	2.802 5*	0.515 2	0.094 1	4404
National language, placebo threshold 100k									
VII.	100	5	0.005 3	0.006 3	0.001 0	0.185 9	1.187 0	0.666 4	5600
VII.	100	6	0.005 7	0.006 0	0.000 3	0.018 6	1.047 1	0.891 4	5600
VII.	100	7	0.006 0	0.005 6	−0.000 4	0.049 3	0.937 0	0.824 3	5600
VII.	100	10.17	0.006 7	0.005 7	−0.001 0	0.519 7	0.852 4	0.471 0	5600
National language, placebo threshold 140k									
VIII.	140	5	0.001 6	0.002 7	0.001 1	1.404 2	1.681 0	0.236 0	5600
VIII.	140	6	0.001 7	0.002 7	0.000 9	1.223 6	1.552 7	0.268 6	5600
VIII.	140	7	0.001 9	0.002 6	0.000 7	0.839 7	1.401 4	0.359 5	5600
VIII.	140	9.09	0.001 9	0.002 5	0.000 6	0.592 7	1.298 6	0.441 4	5600

Notes: Local likelihood ratio results for the discontinuity in the income distribution. Sample specifications: Estimates I use the basic estimation sample of foreigners not holding the permanent residence permit in their first five years of stay in Switzerland. The sample is reduced to those speaking a national language since estimates II. Estimates III are reduced to those who have not moved within the last year. Estimates IV are produced using those younger than 34 years, and estimates V using those older than 34 years, respectively. Estimates VI use the sample of individuals who have not changed their job recently. Estimates VII present results of placebo threshold at 100,000 Swiss francs and use the sample from estimate II, while estimates VIII present results for the placebo threshold of 140,000 Swiss francs. N stands for the number of observations with regard to the observations available to estimate the whole density. Significance levels: * $.05 < p < .1$, ** $.01 < p < .05$, *** $p < .01$.

Table A.3: Density discontinuity estimates - Low-tax municipalities

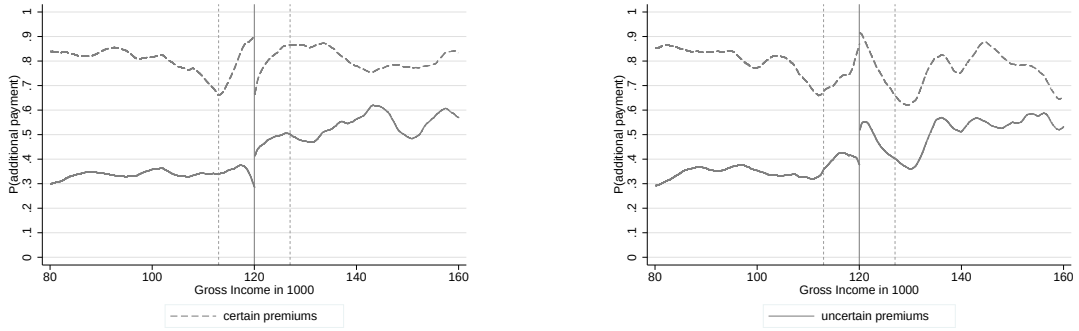
	c	h	$\hat{f}_l$	$\hat{f}_r$	$\hat{\theta}$	$\hat{lr}$	$\hat{f}_r/\hat{f}_l$	p-value	N
Whole sample									
I.	120	5	0.001 5	0.003 8	0.002 3	9.073 6***	2.513 1	0.002 6	6173
I.	120	6	0.001 8	0.003 9	0.002 1	7.739 4***	2.149 5	0.005 4	6173
I.	120	7	0.002 0	0.004 0	0.002 0	7.253 7***	1.980 6	0.007 1	6173
I.	120	12.99	0.002 6	0.004 4	0.001 8	7.074 5***	1.703 8	0.007 8	6173
National language									
II.	120	5	0.001 3	0.004 0	0.002 7	9.261 1***	3.131 1	0.002 3	4623
II.	120	6	0.001 6	0.003 9	0.002 3	7.635 4***	2.512 7	0.005 7	4623
II.	120	7	0.001 8	0.004 0	0.002 3	7.387 0***	2.292 8	0.006 6	4623
II.	120	11.66	0.002 3	0.004 4	0.002 2	7.287 1***	1.946 7	0.006 9	4623
National language & not moved									
III.	120	5	0.001 3	0.004 3	0.003 0	8.491 5***	3.238 3	0.003 6	4060
III.	120	6	0.001 6	0.004 1	0.002 6	7.093 3***	2.615 0	0.007 7	4060
III.	120	7	0.001 8	0.004 3	0.002 5	6.892 8***	2.383 6	0.008 7	4060
III.	120	11.33	0.002 3	0.004 5	0.002 2	6.044 2**	1.952 0	0.014 0	4060
National language & age<34									
IV.	120	5	0.001 0	0.000 9	−0.000 0	0.002 5	0.967 9	0.960 4	2117
IV.	120	6	0.001 2	0.000 8	−0.000 4	0.371 1	0.690 2	0.542 4	2117
IV.	120	7	0.001 4	0.001 0	−0.000 4	0.410 7	0.705 0	0.521 6	2117
IV.	120	12.79	0.001 8	0.001 3	−0.000 5	0.590 5	0.725 2	0.442 2	2117
National language & age≥34									
V.	120	5	0.001 6	0.007 4	0.005 8	10.557 5***	4.532 7	0.001 2	2506
V.	120	6	0.001 9	0.007 4	0.005 4	10.850 3***	3.821 6	0.001 0	2506
V.	120	7	0.002 2	0.007 5	0.005 3	10.975 7***	3.423 6	0.000 9	2506
V.	120	11.36	0.002 7	0.008 0	0.005 3	12.589 2***	2.924 3	0.000 4	2506
National language & not changed job									
VI.	120	5	0.001 5	0.004 7	0.003 2	8.450 2***	3.175 1	0.003 7	3649
VI.	120	6	0.001 8	0.004 6	0.002 9	7.490 8***	2.615 8	0.006 2	3649
VI.	120	7	0.002 0	0.004 8	0.002 8	7.304 5***	2.387 6	0.006 9	3649
VI.	120	11.13	0.002 7	0.005 2	0.002 5	6.243 7**	1.926 1	0.012 5	3649
National language, placebo threshold 100k									
VII.	100	5	0.004 8	0.006 2	0.001 4	0.917 1	1.303 3	0.338 2	4623
VII.	100	6	0.005 0	0.006 2	0.001 2	0.760 3	1.245 5	0.383 2	4623
VII.	100	7	0.005 1	0.006 0	0.000 9	0.492 2	1.177 7	0.483 0	4623
VII.	100	13.21	0.005 2	0.006 0	0.000 8	0.638 9	1.145 6	0.424 1	4623
National language, placebo threshold 140k									
VIII.	140	5	0.001 8	0.001 7	−0.000 1	0.027 7	0.932 1	0.867 7	4623
VIII.	140	6	0.001 8	0.002 0	0.000 2	0.088 8	1.127 0	0.765 7	4623
VIII.	140	7	0.001 7	0.002 3	0.000 5	0.465 3	1.313 4	0.495 2	4623
VIII.	140	10.72	0.001 8	0.002 4	0.000 6	0.591 2	1.338 9	0.442 0	4623

Notes: Local likelihood ratio results for the discontinuity in the income distribution. Sample specifications: Estimates I use the basic estimation sample of foreigners not holding the permanent residence permit in their first five years of stay in Switzerland. The sample is reduced to those speaking a national language since estimates II. Estimates III are reduced to those who have not moved within the last year. Estimates IV are produced using those younger than 34 years, and estimates V using those older than 34 years, respectively. Estimates VI use the sample of individuals who have not changed their job recently. Estimates VII present results of placebo threshold at 100,000 Swiss francs and use the sample from estimate II, while estimates VIII present results for the placebo threshold of 140,000 Swiss francs. N stands for the number of observations with regard to the observations available to estimate the whole density. Significance levels: * $.05 < p < .1$, ** $.01 < p < .05$, *** $p < .01$.

Table A.4: Density discontinuity estimates for quasi-placebo groups

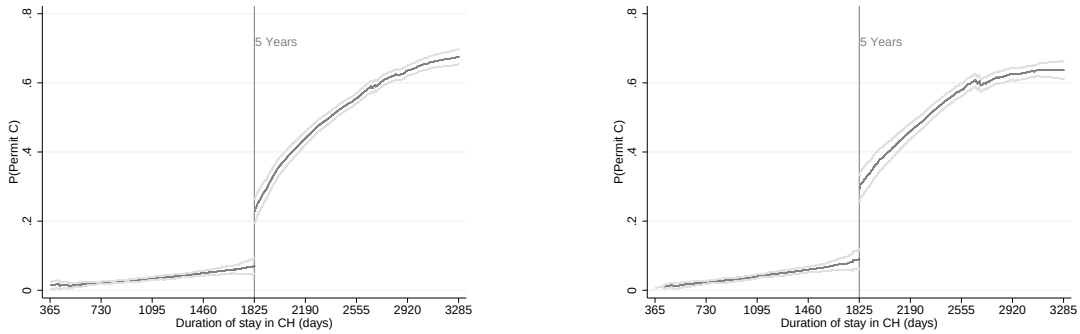
	c	h	$\hat{f}_l$	$\hat{f}_r$	$\hat{\theta}$	$\hat{lr}$	$\hat{f}_r/\hat{f}_l$	p-value	N
High-tax type									
Foreign residents, holding permit C									
I.	120	50	0.001 3	0.001 7	0.000 4	0.174 0	1.272 2	0.676 6	19411
I.	120	6	0.001 3	0.001 4	0.000 1	0.035 2	1.108 2	0.851 1	19411
I.	120	7	0.001 1	0.001 2	0.000 1	0.087 3	1.071 9	0.767 7	19411
I.	120	10.05	0.001 1	0.001 2	0.000 1	0.056 7	1.103 6	0.811 8	19411
Swiss residents									
II.	120	5	0.003 1	0.003 6	0.000 5	1.319 7	1.175 6	0.250 6	42311
II.	120	6	0.003 1	0.003 6	0.000 5	1.327 6	1.159 9	0.249 2	42311
II.	120	7	0.003 1	0.003 5	0.000 4	1.053 0	1.130 4	0.304 8	42311
II.	120	13.90	0.002 9	0.003 3	0.000 4	2.299 3	1.138 4	0.129 4	42311
Low-tax type									
Foreign residents, holding permit C									
I.	120	5	0.001 5	0.002 2	0.000 7	0.660 7	1.502 0	0.416 3	15802
I.	120	6	0.001 4	0.002 1	0.000 6	0.686 8	1.458 0	0.407 3	15802
I.	120	7	0.001 5	0.002 0	0.000 5	0.490 8	1.336 1	0.483 6	15802
I.	120	9.10	0.001 5	0.001 9	0.000 4	0.369 8	1.241 3	0.543 1	15802
Swiss residents									
II.	120	5	0.002 7	0.003 0	0.000 3	0.368 9	1.097 9	0.543 6	39702
II.	120	6	0.002 7	0.002 9	0.000 2	0.214 9	1.067 3	0.643 0	39702
II.	120	7	0.002 7	0.002 8	0.000 1	0.059 4	1.032 2	0.807 4	39702
II.	120	11.55	0.002 7	0.002 6	−0.000 1	0.089 0	0.970 7	0.765 5	39702

Notes: Local likelihood ratio results for the discontinuity in the income distribution for the quasi-placebo groups of foreigners holding the permanent resident permit and Swiss residents.
Sample specifications: Estimates I use the a sample of foreign residents holding the permanent residence permit and residing in Switzerland for at least seven years. Estimates II use the sample of Swiss residents. N stands for the number of observations with regard to the observations available to estimate the whole density in the sample.
Significance levels: * $.05 < p < .1$, ** $.01 < p < .05$, *** $p < .01$.



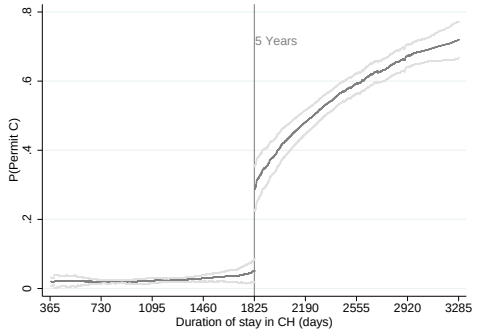
(a) High tax municipalities **(b)** Low tax municipalities

Figure A.2: Probability of reporting to receive additional payments. Local linear smooth of the probability that an individual reports having received additional payments for the sample of foreigners holding the permanent residence permit, speaking a national language, and residing in Switzerland for at least seven years. The dashed line indicates the probability of receiving rather certain additional payments like Christmas bonus, while the solid line indicates rather uncertain premiums like performance bonus and profit participation. Yearly gross income is measured in thousands of Swiss francs and a bandwidth of $5k$ is used.

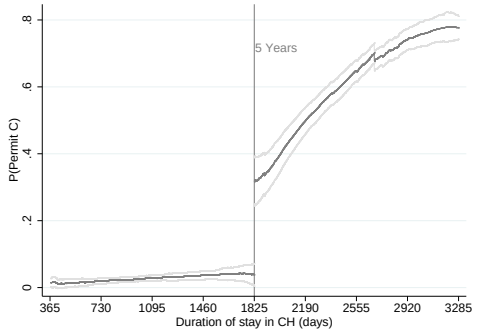


(a) High-tax municipalities **(b)** Low-tax municipalities

Figure A.3: Probability of holding permit C - first stage of fuzzy design. These graphs show local linear smooths of the probability that an individual holds the permanent residence permit around the duration threshold for all income quartiles below the income threshold using a bandwidth of 730 days.

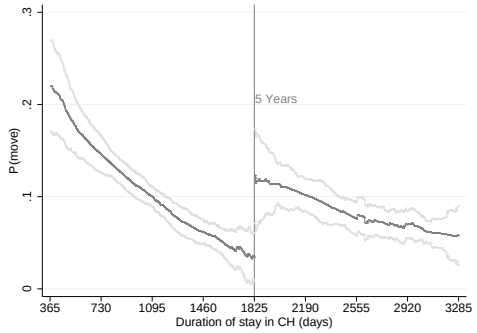


(a) High-tax municipalities

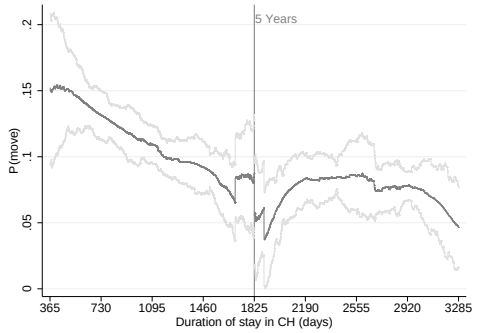


(b) Low-tax municipalities

Figure A.4: Probability of holding permit C - first stage of fuzzy design. These graphs show local linear smooths of the probability that an individual holds the permanent residence permit around the duration threshold for the highest income quartile below the income threshold using a bandwidth of 730 days.



(a) High-tax municipalities



(b) Low-tax municipalities

Figure A.5: Probability of moving - second stage of fuzzy design. These graphs show local linear smooths of the probability that an individual will move around the duration threshold for the highest income quartile below the income threshold using a bandwidth of 730 days.

Table A.5: FRD estimate for the probability of moving and different income groups

High-tax type								
p.	h	τ	se.	Confidence Interval 95% upper	lower	p-value	N_l	N_r
$y \leq q25$								
x	730	0.30	0.38	0.93	−0.34	0.44	588	556
x	365	0.80	1.90	3.95	−2.35	0.68	280	282
x	1460	0.10	0.10	0.27	−0.07	0.33	1244	1067
x	629	0.43	0.63	1.48	−0.61	0.49	505	493
$q25 > y \leq q50$								
x	730	−0.18	0.37	0.43	−0.78	0.63	691	536
x	365	−0.48	3.58	5.45	−6.41	0.89	293	285
x	1460	−0.05	0.13	0.15	−0.26	0.67	1408	986
x	565	−0.21	0.52	0.65	−1.07	0.69	524	421
$q50 > y \leq q75$								
x	730	0.30	0.27	0.74	−0.13	0.25	674	495
x	365	0.53	0.54	1.42	−0.37	0.33	328	263
x	1460	0.23	0.18	0.53	−0.07	0.21	1294	848
x	729	0.30	0.27	0.74	−0.13	0.25	674	495
$y > q75$								
x	730	0.46	0.19	0.77	0.15	0.02	998	674
x	365	0.62	0.35	1.20	0.04	0.08	463	372
x	1460	0.33	0.10	0.48	0.17	0.00	2026	1094
x	498	0.54	0.26	0.97	0.11	0.04	635	493
$y > 120,000$								
x	730	−0.08	0.22	0.28	−0.44	0.71	632	423
x	365	−0.11	0.25	0.30	−0.52	0.66	295	249
x	1460	−0.04	0.12	0.17	−0.25	0.75	1251	663
x	501	−0.05	0.23	0.33	−0.43	0.84	426	312

Notes: Foreigners located in a high-tax municipality when receiving the permanent residence permit after 5 years in Switzerland. FRD estimates for different income groups, where q represents the quartile below the income threshold. Estimates account for sampling weights.

Table A.6: FRD estimate for the probability of moving and different income groups - speaking national language

High-tax type								
p.	h	τ	se.	Confidence Interval 95% upper	lower	p-value	N_l	N_r
$y \leq q25$								
x	730	0.67	1.54	3.23	−1.89	0.66	390	414
x	365	0.98	4.72	8.83	−6.87	0.84	186	212
x	1460	0.13	0.18	0.42	−0.16	0.46	811	789
x	553	−14.54	586.52	959.90	−988.98	0.98	302	314
$q25 > y \leq q50$								
x	730	0.01	0.33	0.55	−0.53	0.98	497	413
x	365	0.15	0.75	1.40	−1.10	0.84	210	215
x	1460	0.00	0.14	0.22	−0.22	0.00	1022	751
x	597	0.06	0.42	0.76	−0.64	0.88	397	336
$q50 > y \leq q75$								
x	730	0.19	0.30	0.68	−0.31	0.53	581	411
x	365	0.57	0.77	1.84	−0.71	0.46	282	223
x	1460	0.15	0.19	0.46	−0.16	0.42	1126	686
x	723	0.19	0.30	0.69	−0.31	0.53	575	408
$y > q75$								
x	730	0.57	0.23	0.94	0.19	0.01	868	596
x	365	0.79	0.45	1.54	0.04	0.08	410	332
x	1460	0.39	0.11	0.57	0.21	0.00	1733	969
x	540	0.63	0.30	1.13	0.14	0.04	625	470
$y > 120,000$								
x	730	−0.14	0.22	0.23	−0.50	0.53	366	251
x	365	−0.08	0.26	0.36	−0.51	0.78	170	147
x	1460	−0.02	0.12	0.18	−0.23	0.85	649	376
x	753	−0.14	0.22	0.22	−0.50	0.53	374	257

Notes: Foreigners located in a high-tax municipality when receiving the permanent residence permit after 5 years in Switzerland. The sample is additionally reduced to those speaking one of the national languages. FRD estimates for different income groups, where q represents the quartile below the income threshold. Estimates account for sampling weights.

Table A.7: FRD estimate for the probability of moving and different income groups - speaking national language

Low-tax type								
p.	h	$\boldsymbol{\tau}$	se.	Confidence Interval 95% upper	lower	p-value	N_l	N_r
$y \leq q25$								
x	730	0.38	0.63	1.43	−0.66	0.54	375	304
x	365	−0.04	0.12	0.17	−0.24	0.75	192	157
x	1460	−0.31	0.28	0.15	−0.78	0.27	746	571
x	639	0.23	0.35	0.81	−0.34	0.50	331	272
$q25 > y \leq q50$								
x	730	−0.10	0.20	0.23	−0.43	0.62	373	289
x	365	0.02	0.25	0.43	−0.40	0.95	173	154
x	1460	0.08	0.13	0.29	−0.14	0.56	751	509
x	585	−0.09	0.23	0.29	−0.47	0.69	298	232
$q50 > y \leq q75$								
x	730	0.01	0.32	0.52	−0.51	0.99	419	288
x	365	0.40	2.48	4.49	−3.69	0.87	179	163
x	1460	−0.11	0.18	0.18	−0.40	0.53	834	462
x	707	0.03	0.33	0.57	−0.52	0.94	402	285
$y > q75$								
x	730	−0.07	0.18	0.23	−0.36	0.71	691	447
x	365	−0.18	0.30	0.32	−0.68	0.56	321	259
x	1460	−0.04	0.12	0.16	−0.23	0.76	1361	736
x	572	−0.07	0.22	0.30	−0.44	0.75	520	373
$y > 120,000$								
x	730	−0.26	0.32	0.27	−0.79	0.42	378	340
x	365	−0.93	1.26	1.15	−3.00	0.46	185	183
x	1460	−0.05	0.13	0.16	−0.26	0.72	833	568
x	605	−0.44	0.46	0.32	−1.20	0.34	313	291

Notes: Foreigners located in a low-tax municipality when receiving the permanent residence permit after 5 years in Switzerland. The sample is additionally restricted to individuals speaking the national language. FRD estimates for different income groups, where q represents the quartile below the income threshold. Estimates account for sampling weights.

Table A.8: FRD estimate for the probability of moving and different income groups- with a child in the household

High-tax type								
p.	h	τ	se.	Confidence Interval 95% upper	lower	p-value	N_l	N_r
$y \leq q25$								
x	730	0.27	0.42	0.97	−0.43	0.52	197	255
x	365	−0.39	1.46	2.07	−2.85	0.79	97	124
x	1460	0.11	0.13	0.32	−0.10	0.39	353	534
x	550	2.90	46.66	81.38	−75.59	0.95	156	188
$q25 > y \leq q50$								
x	730	−0.21	0.75	1.03	−1.45	0.78	166	190
x	365	0.52	4.42	7.86	−6.82	0.91	77	94
x	1460	−0.11	0.23	0.28	−0.49	0.65	283	378
x	597	0.06	1.12	1.92	−1.80	0.96	134	156
$q50 > y \leq q75$								
x	730	0.87	1.17	2.81	−1.07	0.46	128	155
x	365	10.96	96.86	172.74	−150.83	0.91	61	77
x	1460	0.14	0.30	0.64	−0.36	0.64	229	290
x	546	1.11	1.53	3.66	−1.44	0.47	91	118
$y > q75$								
x	730	0.04	0.24	0.45	−0.36	0.86	174	163
x	365	−0.11	0.67	1.02	−1.23	0.88	80	86
x	1460	0.11	0.13	0.33	−0.10	0.39	308	270
x	682	0.03	0.26	0.46	−0.40	0.91	160	156
$y > 120,000$								
x	730	0.46	0.50	1.29	−0.37	0.36	150	83
x	365	1.44	2.44	5.49	−2.62	0.56	72	46
x	1460	0.08	0.20	0.41	−0.25	0.69	272	149
x	715	0.48	0.53	1.36	−0.39	0.36	146	83

Notes: Foreigners located in a high-tax municipality when receiving the permanent residence permit after 5 years in Switzerland. The sample is additionally reduced to those speaking one of the national languages and with children in the household. FRD estimates for different income groups, where q represents the quartile below the income threshold. Estimates account for sampling weights..

Table A.9: FRD estimate for the probability of moving and different income groups- no children in the household

High-tax type								
p.	h	τ	se.	Confidence Interval 95% upper	lower	p-value	N_l	N_r
$y \leq q25$								
x	730	−0.53	1.79	2.45	−3.50	0.77	193	159
x	365	0.70	2.03	4.09	−2.69	0.73	89	88
x	1460	0.27	1.24	2.33	−1.78	0.83	458	255
x	855	−0.36	1.55	2.22	−2.94	0.82	247	176
$q25 > y \leq q50$								
x	730	0.18	0.35	0.77	−0.41	0.61	331	223
x	365	0.25	0.73	1.49	−0.98	0.73	133	121
x	1460	0.12	0.18	0.42	−0.17	0.49	739	373
x	775	0.18	0.34	0.74	−0.39	0.61	360	235
$q50 > y \leq q75$								
x	730	−0.13	0.33	0.41	−0.67	0.70	453	256
x	365	−0.29	0.54	0.61	−1.19	0.59	221	146
x	1460	0.13	0.25	0.54	−0.29	0.62	897	396
x	743	−0.13	0.32	0.41	−0.66	0.70	458	260
$y > q75$								
x	730	0.81	0.32	1.34	0.27	0.01	694	433
x	365	1.03	0.53	1.90	0.16	0.05	330	246
x	1460	0.56	0.16	0.82	0.30	0.00	1425	699
x	713	0.81	0.33	1.35	0.27	0.01	673	423
$y > 120,000$								
x	730	−0.40	0.32	0.13	−0.93	0.21	306	228
x	365	−0.36	0.31	0.15	−0.87	0.25	141	133
x	1460	−0.14	0.15	0.10	−0.39	0.34	564	345
x	640	−0.42	0.34	0.14	−0.98	0.21	256	203

Notes: Foreigners located in a high-tax municipality when receiving the permanent residence permit after 5 years in Switzerland. The sample is additionally reduced to those speaking one of the national languages and no children in the household. FRD estimates for different income groups, where q represents the quartile below the income threshold. Estimates account for sampling weights.

Table A.10: FRD estimate for the probability of moving-
heterogeneity in tax change

p.	h	τ	se.	Confidence Interval 95% upper	lower	p-value	N_l	N_r
High-tax type								
tax change<median								
x	730	0.08	0.27	0.53	−0.37	0.77	404	294
x	365	0.06	0.45	0.80	−0.68	0.89	193	161
x	1460	0.15	0.14	0.38	−0.08	0.29	829	464
x	802	0.09	0.25	0.51	−0.33	0.73	443	313
tax change≥median								
x	730	0.95	0.39	1.60	0.31	0.02	464	302
x	365	1.28	0.79	2.58	−0.03	0.11	217	171
x	1460	0.62	0.18	0.92	0.33	0.00	904	505
x	612	1.04	0.48	1.83	0.26	0.03	384	267
Low-tax type								
tax change<median								
x	730	−0.02	0.22	0.34	−0.38	0.93	324	217
x	365	−0.05	0.27	0.40	−0.49	0.87	152	122
x	1460	−0.02	0.15	0.22	−0.26	0.89	646	358
x	708	−0.02	0.22	0.35	−0.38	0.94	311	214
tax change≥median								
x	730	−0.14	0.32	0.39	−0.66	0.66	367	230
x	365	−0.73	1.45	1.67	−3.13	0.62	169	137
x	1460	−0.06	0.20	0.27	−0.38	0.77	715	378
x	744	−0.14	0.31	0.37	−0.66	0.65	377	238

Notes: Foreigners when receiving their permanent residence permit after 5 years in Switzerland. The sample is additionally reduced to those speaking one of the national languages. FRD estimates for the highest income quartiles below the income threshold of 120,000 Swiss francs, for the two municipality types and separated at the median of the tax change at the threshold. Estimates account for sampling weights.

Table A.11: FRD estimate for the probability of moving-
heterogeneity in knowledge

p.	h	τ	se.	Confidence Interval 95% upper	lower	p-value	N_l	N_r
High-tax type								
knowledge<median								
x	730	−0.13	0.18	0.16	−0.42	0.47	531	209
x	365	−0.12	0.31	0.40	−0.63	0.71	247	120
x	1460	0.01	0.11	0.19	−0.17	0.92	1090	334
x	627	−0.13	0.20	0.19	−0.46	0.51	440	184
knowledge≥median								
x	730	0.81	0.37	1.42	0.20	0.03	467	465
x	365	0.89	0.58	1.86	−0.08	0.13	216	252
x	1460	0.51	0.16	0.77	0.26	0.00	936	760
x	7039	0.81	0.38	1.44	0.19	0.03	449	451
Low-tax type								
knowledge<median								
x	730	−0.21	0.34	0.35	−0.78	0.53	423	111
x	365	−0.42	0.92	1.09	−1.94	0.64	186	65
x	1460	−0.03	0.15	0.22	−0.27	0.86	843	192
x	769	−0.21	0.31	0.30	−0.71	0.50	450	125
knowledge≥median								
x	730	0.13	0.18	0.44	−0.17	0.47	418	397
x	365	0.08	0.27	0.53	−0.37	0.77	198	232
x	1460	0.05	0.14	0.29	−0.18	0.72	822	639
x	653	0.15	0.20	0.48	−0.19	0.47	368	366

Notes: Foreigners when receiving their permanent residence permit after 5 years in Switzerland. FRD estimates for the highest income quartiles below the income threshold of 120,000 Swiss francs, for the two municipality types and separated at the median of the constructed knowledge measure. Estimates account for sampling weights.

Table A.12: Parametric FRD estimate for the probability of moving and different income groups-first order polynomial

High-tax type					
	I.	II.	III.	IV.	V.
τ	0.184*	0.038	0.057	0.279***	0.057
	(0.097)	(0.056)	(0.063)	(0.066)	(0.073)
No. of obs.	2695	2845	3089	3245	1716
Low-tax type					
	I.	II.	III.	IV.	V.
τ	–0.039	0.193	0.007	0.031	–0.007
	(0.065)	(0.118)	(0.098)	(0.072)	(0.063)
No. of obs.	2101	2146	2329	2589	1828

Notes: Standard parametric FRD (IV) estimates using a linear model for foreigners whose duration of stay in Switzerland is under 10 years. The sample is further restricted to those speaking one of the national languages. Column I corresponds to the sample $y \leq q25$, column II to $q25 > y \leq q50$, column III $q50 > y \leq q75$, column IV $y > q75$, and column V to the quasi-placebo group $y > 120,000$, where q represents the quartile below the income threshold. Estimates account for sampling weights. Standard errors in parentheses.

Significance levels: * .05<p<.1, ** .01<p<.05, *** p<.01.

Table A.13: Parametric FRD estimate for the probability of moving and different income groups- second order polynomial

High-tax type					
	I.	II.	III.	IV.	V.
τ	0.320	0.050	0.061	0.252**	–0.055
	(0.297)	(0.103)	(0.102)	(0.114)	(0.241)
No. of obs.	2695	2845	3089	3245	1716
Low-tax type					
	I.	II.	III.	IV.	V.
τ	–0.038	0.020	–0.239	0.049	–0.111
	(0.112)	(0.279)	(0.171)	(0.152)	(0.126)
No. of obs.	2101	2146	2329	2589	1828

Notes: Standard parametric FRD (IV) estimates using a specification with a polynomial of second order for foreigners whose duration of stay in Switzerland is under 10 years and the sample is further restricted to those speaking one of the national languages. Column I corresponds to the sample $y \leq q25$, column II to $q25 > y \leq q50$, column III $q50 > y \leq q75$, column IV $y > q75$, and column V to the quasi-placebo group $y > 120,000$, where q represents the quartile below the income threshold. Estimates account for sampling weights. Standard errors in parentheses.
Significance levels: * .05<p<.1, ** .01<p<.05, *** p<.01.

Table A.14: Parametric FRD estimate for the probability of moving and different income groups-first order polynomial & covariates

High-tax type					
	I.	II.	III.	IV.	V.
τ	0.209	0.003	0.051	0.232***	0.019
	(0.140)	(0.061)	(0.076)	(0.046)	(0.045)
No. of obs.	2651	2809	3062	3224	1709
$N_{cluster}$	24	24	24	24	21
Low-tax type					
	I.	II.	III.	IV.	V.
τ	–0.021	0.185	0.032	0.020	–0.029
	(0.055)	(0.126)	(0.087)	(0.056)	(0.050)
No. of obs.	2082	2120	2307	2572	1816
$N_{cluster}$	24	24	24	24	23

Notes: Standard parametric FRD (IV) estimates for foreigners whose duration of stay in Switzerland is under 10 years and the sample is further restricted to those speaking one of the national languages. These estimates additionally control for age, age squared, children, household size, marital status, profession classes , education classes and canton fixed effects. Column I corresponds to the sample $y \leq q25$, column II to $q25 > y \leq q50$, column III $q50 > y \leq q75$, column IV $y > q75$, and column V to the quasi-placebo group $y > 120,000$, where q represents the quartile below the income threshold. Estimates account for sampling weights. Standard errors in parentheses and clustered at the cantonal level.

Significance levels: * .05<p<.1, ** .01<p<.05, *** p<.01.

Table A.15: FRD estimate for the probability of moving and different income groups-placebo-threshold 4 years of stay

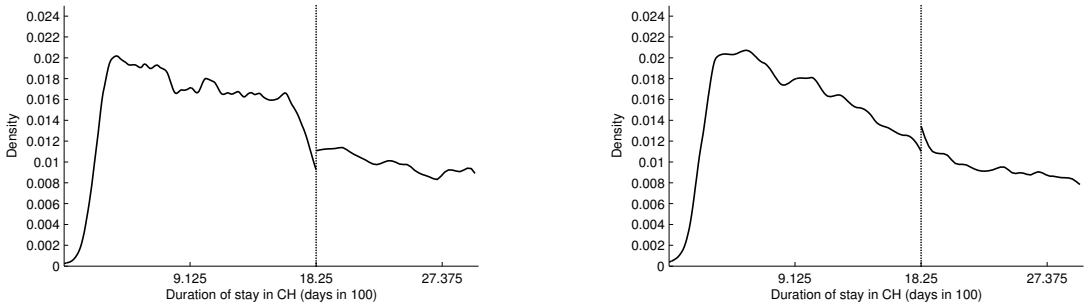
High-tax type								
p.	h	τ	se.	Confidence Interval 95% upper	lower	p-value	N_l	N_r
$y \leq q25$								
x	730	0.60	0.53	1.48	−0.29	0.26	442	400
x	365	1.47	1.85	4.57	−1.62	0.43	202	188
x	1460	0.18	0.28	0.63	−0.28	0.52	623	815
x	462	0.97	0.87	2.42	−0.49	0.27	275	245
$q25 > y \leq q50$								
x	730	−0.08	0.71	1.09	−1.25	0.91	578	425
x	365	2.08	6.49	12.84	−8.68	0.75	287	210
x	1460	0.06	0.50	0.88	−0.77	0.91	812	808
x	500	−21.74	981.72	1 602.82	−1 646.29	0.98	393	301
$q50 > y \leq q75$								
x	730	0.07	0.45	0.80	−0.67	0.88	614	505
x	365	−0.93	5.15	7.57	−9.44	0.86	299	282
x	1460	−0.05	0.35	0.52	−0.62	0.88	845	847
x	646	0.18	0.58	1.14	−0.78	0.76	537	461
$y > q75$								
x	730	0.26	0.80	1.59	−1.06	0.74	938	744
x	365	1.09	1.90	4.23	−2.05	0.57	456	412
x	1460	−0.32	0.61	0.69	−1.32	0.60	1321	1233
x	550	0.37	4.54	7.86	−7.12	0.94	694	581
$y > 120,000$								
x	730	0.14	0.66	1.23	−0.96	0.84	462	392
x	365	1.17	1.48	3.62	−1.27	0.43	243	213
x	1460	−0.80	0.97	0.79	−2.39	0.41	625	630
x	532	0.51	0.85	1.93	−0.90	0.55	330	309

Notes: FRD estimates for different income groups residing in high-tax municipalities and a placebo threshold at 4 years of stay in Switzerland. The sample is additionally reduced to those speaking one of the national languages. Estimates for different income groups, where q represents the quartile below the income threshold. Estimates account for sampling weights.

Table A.16: FRD estimate for the probability of moving and different income groups-placebo-threshold 6 years of stay

High-tax type								
p.	h	$\boldsymbol{\tau}$	se.	Confidence Interval 95% upper	lower	p-value	N_l	N_r
$y \leq q25$								
x	730	−0.07	0.38	0.56	−0.70	0.85	398	415
x	365	−22.86	2 650.71	4 377.95	−4 423.66	0.99	212	202
x	1460	0.02	0.11	0.20	−0.16	0.86	842	744
x	834	−0.05	0.33	0.49	−0.59	0.89	465	475
$q25 > y \leq q50$								
x	730	0.18	1.46	2.59	−2.23	0.90	425	383
x	365	−0.74	2.12	2.78	−4.25	0.73	214	198
x	1460	0.09	0.23	0.47	−0.29	0.71	1003	707
x	670	0.09	2.31	3.91	−3.73	0.97	390	353
$q50 > y \leq q75$								
x	730	−0.46	1.61	2.19	−3.10	0.78	505	342
x	365	0.43	1.43	2.79	−1.94	0.77	222	188
x	1460	0.13	0.27	0.58	−0.32	0.62	1119	569
x	787	−0.64	2.38	3.29	−4.57	0.79	550	369
$y > q75$								
x	730	0.12	1.71	2.93	−2.69	0.94	742	489
x	365	0.49	1.84	3.52	−2.55	0.79	330	264
x	1460	0.10	0.16	0.37	−0.17	0.55	1682	761
x	612	1.66	25.07	42.99	−39.67	0.95	605	413
$y > 120,000$								
x	730	−0.34	0.32	0.19	−0.86	0.30	392	238
x	365	−0.36	0.31	0.16	−0.87	0.25	178	132
x	1460	−0.28	0.15	−0.04	−0.52	0.05	854	388
x	619	−0.32	0.31	0.19	−0.82	0.30	312	208

Notes: FRD estimates for different income groups residing in high-tax municipalities and a placebo threshold at 6 years of stay in Switzerland. The sample is additionally reduced to those speaking one of the national languages. Estimates for different income groups, where q represents the quartile below the income threshold. Estimates account for sampling weights.



(a) High-tax municipalities **(b)** Low-tax municipalities

Figure A.6: Distributions of duration of stay in Switzerland for the two municipality types. Local likelihood estimates using $bandwidth = 1.20$ separately from both sides of the duration threshold 1825 days, or 5 years. The corresponding density discontinuity estimates can be found in rows II of Table A.17.

Table A.17: Density discontinuity estimates for duration of stay

	c	h	$\hat{f}_l$	$\hat{f}_r$	$\hat{\theta}$	$\hat{l}$	$\hat{f}_r/\hat{f}_l$	p-value	N
High-tax type									
I.	18.25	0.90	0.010 3	0.011 1	0.000 7	0.165 8	1.072 7	0.683 9	32237
II.	18.25	1.20	0.010 0	0.011 2	0.001 2	0.543 2	1.120 2	0.461 1	32237
III.	18.25	1.50	0.009 5	0.011 3	0.001 8	1.545 0	1.190 3	0.213 9	32237
IV.	18.25	2.01	0.009 5	0.011 1	0.001 6	1.530 6	1.169 9	0.216 0	32237
Low-tax type									
I.	18.25	0.90	0.010 4	0.013 2	0.002 8	1.057 6	1.264 7	0.303 8	26357
II.	18.25	1.20	0.010 6	0.012 3	0.001 7	0.432 5	1.157 5	0.510 7	26357
III.	18.25	1.50	0.010 9	0.012 6	0.001 6	0.451 0	1.149 1	0.501 9	26357
IV.	18.25	1.86	0.011 3	0.012 7	0.001 4	0.398 1	1.127 7	0.528 1	26357

Notes: Local likelihood ratio results for the discontinuity in the distribution of the duration of stay. N stands for the number of observations with regard to the observations available to estimate the whole density in the sample.
Significance levels: * $.05 < p < .1$, ** $.01 < p < .05$, *** $p < .01$.