

Roller, Marcus; Schmidheiny, Kurt

Conference Paper

Mobility and Progressive Taxation

54th Congress of the European Regional Science Association: "Regional development & globalisation: Best practices", 26-29 August 2014, St. Petersburg, Russia

Provided in Cooperation with:

European Regional Science Association (ERSA)

Suggested Citation: Roller, Marcus; Schmidheiny, Kurt (2014) : Mobility and Progressive Taxation, 54th Congress of the European Regional Science Association: "Regional development & globalisation: Best practices", 26-29 August 2014, St. Petersburg, Russia, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/124484>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Mobility and Progressive Taxation*

- very preliminary: do not cite or quote -

Marcus Roller[§] Kurt Schmidheiny[§]

February 28, 2014

Abstract

In fiscally centralised countries with a unique country-wide income tax schedule, it is straightforward to quantify the degree of progressivity. In fiscally decentralised countries with varying local tax schedules, however, this is not the case. In these countries, the effective tax progressivity depends on the distribution of taxpayers across local jurisdiction as well as on local income distributions. The latter might differ systematically because high income households can partly or fully avoid high tax rates by sorting into low tax locations.

This paper develops an empirical approach in order to quantify the effective tax progressivity in a fiscally highly decentralised country - Switzerland - taking the relative size of jurisdictions and the actually observed income sorting into account. Exploiting data on the universe of Swiss taxpayers, we find that rich households face significantly lower tax rates and lower progressivity than in the benchmark case that does not take the income sorting into account. The results are stronger for singles than for families indicating that singles are more sensitive to tax differentials than families. Furthermore, we find suggestive evidence that due to this income sorting the Swiss income tax system is not only less progressive but even regressive for single households with very high incomes.

Keywords: Progressive Taxation, Fiscal Decentralisation, Income Segregation

JEL classification: H71, H73, R23

*We thank the Swiss Federal Tax Administration for granting us access to individual-level federal tax data and the Swiss National Science Foundation (Sinergia grants 130648 and 147668) for financial support.

[§]Wirtschaftswissenschaftliche Fakultät, Universität Basel, Peter-Merian-Weg 6 4002 Basel, Switzerland. E-mail: marcus.roller@unibas.ch

[§]Wirtschaftswissenschaftliche Fakultät, Universität Basel, Peter-Merian-Weg 6 4002 Basel, Switzerland. E-mail: kurt.schmidheiny@unibas.ch. Also affiliated with the Centre for Economic Policy Research (CEPR) and CESifo.

1 Introduction

Progressive income tax rates are the most common mechanism to redistribute income from high income households to low income households (Boadway and Keen, 2000). In fiscally centralised countries with a unique country-wide income tax schedule, it is straightforward to quantify the degree of progressivity. In fiscally decentralised countries with varying local tax schedules, however, this is not the case. In these countries, the effective tax progressivity depends on the distribution of taxpayers across local jurisdiction as well as on local income distributions. The latter might differ systematically because high income households can partly or fully avoid high tax rates by sorting into low tax locations. This paper seeks to quantify the effective tax progressivity in a highly fiscally decentralised country.

Switzerland is especially suitable for this purpose because of its highly decentralised income tax system: In addition to a federal income tax, 26 cantons set their own income tax schedules and about 2'700 municipalities charge a local income tax that is a multiple of the cantonal tax rate. Hodler and Schmidheiny (2006) show theoretically that local income tax differentials lead to income sorting and result in an effective canton-wide tax progressivity which is lower than in each individual jurisdiction. They argue that high income households sort into low-tax municipalities because they are more sensitive to tax rates than low income households. Schmidheiny (2006a) provides empirical evidence of this sorting for the Basel area. Based on publicly available survey data, Hodler and Schmidheiny (2006) provide first estimates of the effective progressivity of the income tax system for the canton of Zurich. They show that the effective canton-wide tax progressivity is indeed lower than the tax progressivity in each individual municipality of the canton. These estimates, however, rest upon strong parametric assumptions because survey data contain only very few observations in the upper tail of the income distribution.

This paper empirically estimates the effective progressivity for the entire Swiss income tax system based on the universe of individual Swiss taxpayers. Unlike most survey data, our data contains uncensored high income tax payers allowing us to estimate the progressivity of the income tax system even for top incomes with high precision. We show that high income people pay indeed comparably low taxes and that the effective progressivity of the Swiss income tax system is far lower for high incomes than the tax codes of the single jurisdictions suggest. We find suggestive evidence that the effective tax schedule is even *regressive* for very high incomes for singles. This result is strongest in the agglomeration of Zurich.

In this paper we discuss the properties of three progressivity measures that were suggested by Musgrave and Thin (1948) - in particular the change in average rate progression. This progressivity measure and the tax rates are location-specific, i.e. *conditional* on the jurisdiction. In order to quantify the country-wide progressivity, we seek a country-wide, i.e. an *unconditional* progressivity measure. Such a measure must account for the fact that high income households may systematically sort into low-tax locations. A

measure that satisfies these conditions is the *expected average tax rate* (see Hodler and Schmidheiny (2006)). Additionally, we define the *population weighted mean average tax rate* which does not take the income sorting into account and equals the former in the absence of income sorting. This property makes it a perfect benchmark for our estimations. The respective progressivity measures are the average rate progression of the expected average tax rate and the population weighted mean average tax rate. Building up on the two municipality model of Schmidheiny (2006a), we theoretically show that the expected average tax rate for high income households is lower than the population weighted mean average tax rates. For low income household the expected average tax rate is higher than this benchmark. We also show that the income sorting results in a lower progressivity for high incomes.

In order to estimate the expected average tax rate and its progressivity, we take individual data on the universe of Swiss taxpayers and calculate for each taxpayer the average tax rates at his residential municipality and perform a local polynomial regression of these average tax rates on the natural logarithm of the gross income. This local polynomial regression produces a consistent¹ estimator of the expected average tax rate as well as a consistent estimator for the country-wide average rate progression: Note that this estimated expected average tax rate is *not* conditional on an individual municipality. Our local polynomial regression therefore integrates the location out - accounting for the actually observed income sorting. The resulting estimator is basically a weighted mean of the conditional average tax rates weighted by the number of taxpayers with a specific income in each municipality. Different from Hodler and Schmidheiny (2006), we do not have to make any parametric assumptions on the income distributions. Due to our full population income data we have enough observations to apply this estimation strategy up to very high income levels.

As a benchmark we also estimate the weighted mean of the the local total average tax rates weighted by the local population². This weighted mean would be equal to the above's expected average tax rate in the absence of income sorting, i.e. if the income distribution in each municipality would exactly equal the country-wide distribution.

For our estimations, use the universe of uncensored individual incomes of Swiss taxpayers. This data has been made accessible by the Sinergia Project 130648 founded by the Swiss National Science Foundation in cooperation with the Swiss Federal Tax Administration (ESTV). We focus on two typical household types: single households and married couples with two children. In 2009, we observe data on 1'649'167 single households and 303'283 families with two children.

For singles including all Swiss municipalities in 2009 we find that starting from 100'000 CHF on, the expected average tax rate is significantly lower than the population weighted mean of the average tax rates. The gap is constantly increasing up to about ten percentage

¹The local polynomial regression estimator is generally biased. But including higher order polynomials reduces this bias especially at the boundaries (see e.g. Fan and Gijbels (1996))

²Amount of households of a type in municipality m . See data section for discussion.

points for an income of 10 Mio. CHF. This difference reflects the effect of the income sorting. Thus, tax payers with incomes above 100'000 CHF live in municipalities with comparably low tax rates. Although the expected average tax rate starts flattening and decreasing after 1 Mio. CHF, the progressivity is never significantly negative. If we neglect the very progressive federal tax in our calculation, we find a lower but neither negative progressivity. Because the expected average tax rates seem to peak at 500'000 CHF, we conduct a second test for regressivity. We test the difference of the estimates of the expected average tax rate at 500'000 CHF income with those of higher incomes. We observe significantly lower tax rates for incomes above 1.5 Mio. CHF. These results suggest that the insignificance in our first estimation stems from taking the limit of the income differences rather than from insignificant differences in the tax rates. Hence, although the progressivity measure is not significantly negative, we find suggestive evidence for regressivity in the Swiss income tax system for singles.

When we reduce the sample to households living in a specific agglomeration, we get further interesting results: While we find the strongest effect for the agglomeration of Zurich with even a range of negative progressivity between 1 and 2 Mio. CHF, we do not find strong effects for Basel. There are three facts that help explaining these findings: First, the possibilities to avoid taxes in the Basel agglomeration are lower due to a small variation in tax rates among these municipalities. Second, the tax rates in Basel are comparably high. Third, we observe fewer people with very high incomes living in this agglomeration. All these findings together suggest that people with very high incomes are very mobile and optimise their tax payments at least on national level by moving to very low tax municipalities.

We repeat the estimation also for families with two children and can show that the effect of the income sorting on the expected average tax rate is lower than for singles. In none of the samples we find any evidence for regressivity.

Overall, we show that high income people are much more likely to live in low tax municipalities which lowers the progressivity of the Swiss income tax systems significantly. Especially, when we consider the singles in the agglomeration of Zurich, we even find the income tax system to be regressive for high incomes despite taking the very progressive federal income tax into account.

The paper is organised as follows: Section 2 summarises the basic facts of the Swiss income tax system, discusses the issue of measuring progressivity, and finally presents theoretical implications from a two municipality model. Section 3 derives the empirical strategy while section 4 describes the data at hand and its processing. Section 5 presents the results of the estimations before section 6 concludes.

2 Theoretical Background

2.1 The Swiss Income Tax System

As mentioned above the Switzerland taxes the income on three different levels: First, the federal state charges the gross income y with tax $T^f(y)$. This income tax schedule is highly progressive and independent of the location choice within Switzerland.

Second, the cantons set their own income tax schedules $T^c(y, m)$. c denotes the residential canton and m the residential municipality. Because each municipality is located in only one canton we can write the cantonal tax as a function of the residential municipality m . Municipalities also charge an income tax. Their tax is a multiple of the cantonal tax rate: $M_m T^c(y, m)$. In addition to the political jurisdictions the churches charge a tax that is also a multiple of the cantonal tax rates. Thus, let $K_m T^c(y, m)$ denote the church tax in municipality³ m . The total income tax liability of an individual with gross income y living in municipality m its therefore:

$$T(y, m) = (1 + M_m + K_m)T^c(y, m) + T^f(y) \quad (1)$$

The average tax rate in municipality m at gross income y is the ratio of the income tax liability $T(y, m)$ and the gross income:

$$ATR(y, m) = \frac{T(y, m)}{y} \quad (2)$$

The federal tax cannot be avoided by moving within Switzerland, i.e. it is dependent on the location choice. In order to disentangle the progressivity of the local tax system and that part that stems from the federal tax rate, we define the local income tax liability as the sum of all income taxes that vary across locations:

$$T^l(y, m) = (1 + M_m + K_m)T^c(y, m) \quad (3)$$

The local average tax rate is then defined as:

$$ATR^l(y, m) = \frac{T^l(y, m)}{y} \quad (4)$$

This income tax system can lead to different tax rates in each of the about 2700 municipalities in Switzerland. Thus, already neighbouring municipalities can have different tax rates. This highly decentralised income tax system with its potential and actual large variation in the local tax rates makes Switzerland especially suitable for the purpose of quantifying the effective progressivity of an decentralised income tax system. Furthermore, we can also investigate how the federal tax rate affects the progressivity of the income tax system.

³The geographic borders of the municipalities and the church communities are not necessarily congruent. See data section for discussion.

Table 1: Local Progressivity Measures

Name	Formula	progressive	regressive
Average Rate Progression	$\alpha(y) = \frac{\partial ATRR(y)}{\partial y}$	> 0	< 0
Liability Progression	$\lambda(y) = \frac{\partial T(y)}{\partial y} \frac{y}{T(y)}$	> 1	< 1
Residual Income Progression	$\rho(y) = \frac{\partial [y-T(y)]}{\partial y} \frac{y}{[y-T(y)]}$	< 1	> 1

2.2 Measures of Progressivity

In order to quantify the effective progressivity, we first have to define our understanding of progressivity. This section discusses the selection of the progressivity measure. The decision whether a tax schedule is progressive or not depends on the measure of progressivity (Pigou, 1951). Generally, we distinguish two families of progressivity measures: The first family focuses on income inequalities and redistribution. Therefore, they build on a comparison of prae- and after-tax income distributions (see e.g. Atkinson (1970), Jakobsson (1976), Rothschild and Stiglitz (1973)). One of the most common concepts is the Lorenz criterion: A tax schedule is unambiguously more redistributive than another one if and only if the Lorenz curve of the resulting after tax income lies entirely inside the one of the resulting after tax income of the other tax schedule (Jakobsson, 1976).

The second family of progressivity measures focuses on properties of the tax schedule itself. These measures are independent of the actual income distribution. Furthermore, they are usually locally defined, i.e. they are a function of the underlying tax base. Although they are independent of the income distribution, some of them allow for a comparison of the redistributive effect of different tax schedules (Jakobsson, 1976). In this paper, we focus on the latter progressivity measures because we are interested in the properties of the tax schedules themselves.

Pigou (1951) defines two kinds of progressivity. Either a tax formula is called *progressive* if the average rate increases in the tax base (in our case income) or if the marginal tax increases in the tax base. Both concepts differ and there is no reason to consider one being superior. Because we want to investigate how the tax rates change with income, we will apply the first concept. Based on this concept, Musgrave and Thin (1948) discuss three local progressivity measures: The average rate progression, the liability progression, and the residual income progression. Table 1 gives an overview and summarises their main properties.

Average Rate Progression

The first measure, the average rate progression, is the first derivative of the average tax rate with respect to gross income. Note that the average rate progression as well as the

following progressivity measures are defined over the gross income instead of the taxable income. This accounts for the possible deductions which are part of the income tax system and can cause huge differences in the resulting average tax rates in particular if we consider families with children. A tax schedule is called progressive (regressive) if the average tax rate is increasing (decreasing) in income, i.e. the first derivative is positive (negative).

Liability Progression

The liability progression is a similar concept that measures the mechanical elasticity of the tax payments $T(y)$ to the gross income. Slitor (1948) calls this measure built-in flexibility because it depicts the sensitivity of individuals' tax yields with respect to changes in their personal income and therefore is measure of a property of the tax schedule alone and not of the income distribution. A tax schedule is progressive (regressive) if the the liability progression is greater (smaller) than one. It is easy to show that for strictly positive average tax rates and incomes, these conditions translate into the same as for the average rate progression:

$$\begin{aligned}
1 &< \frac{\partial T(y)}{\partial y} \frac{y}{T(y)} \\
1 &< \frac{\partial [ATR(y)y]}{\partial y} \frac{1}{ATR(y)} \\
1 &< \left[\frac{\partial ATR(y)}{\partial y} y + ATR(y) \right] \frac{1}{ATR(y)} \\
0 &< \frac{\partial ATR(y)}{\partial y} \frac{y}{ATR(y)} \\
0 &< \frac{\partial ATR(y)}{\partial y}
\end{aligned}$$

While the sorting into progressive or regressive tax schedules is consistent with the average rate progression, this is not necessarily the case for the degree of progression. Consider two tax schedules with the average tax rate functions $ATR(y)$ and $ATR'(y)$ with positive average rate progressions $\alpha(y) > \alpha'(y) > 0$. Then $\lambda(y) > \lambda'(y)$ only if $\frac{\alpha(y)}{\alpha'(y)} > \frac{ATR(y)}{ATR'(y)}$:

$$\begin{aligned}
&\lambda(y) > \lambda'(y) \\
&\frac{\partial ATR(y)}{\partial y} \frac{1}{ATR(y)} > \frac{\partial ATR'(y)}{\partial y} \frac{1}{ATR'(y)} \\
&\frac{\alpha(y)}{\alpha'(y)} > \frac{ATR(y)}{ATR'(y)} \\
&\frac{\alpha(y)}{\alpha'(y)} \frac{ATR'(y)}{ATR(y)} > 1
\end{aligned} \tag{5}$$

If $0 > \alpha(y) > \alpha'(y)$, i.e. both tax schedules are regressive at y , condition (5) reads as:

$$\frac{\alpha(y)}{\alpha'(y)} < \frac{ATR(y)}{ATR'(y)} \tag{6}$$

Unlike the average rate progression the liability progression allows for comparisons of the redistributive effects of tax schedules. If $\lambda^1(y) > \lambda^2(y)$ for all y , the tax schedule 1 is more redistributive than tax schedule 2 (Jakobsson, 1976). This means that the Lorenz Curve of the resulting after tax income distribution of schedule 1 lies completely inside the one of tax schedule 2.

Residual Income Progression

The residual income progression measures the elasticity of the after tax income (residual income) with respect to the gross income. A tax schedule is progressive (regressive) if the the liability progression is smaller (greater) than one. It is easy to show that for strictly positive gross incomes and average tax rates smaller than 100%, these conditions translate into the same as for the average rate progression:

$$\begin{aligned}\rho(y) &= \frac{\partial [y - T(y)]}{\partial y} \frac{y}{[y - T(y)]} < 1 \\ &1 - \frac{\frac{\partial ATR(y)}{\partial y} y}{1 - ATR(y)} < 1 \\ &\frac{\partial ATR(y)}{\partial y} < 0\end{aligned}$$

Thus, the sorting of tax schedules into progressive and regressive is consistent with the average rate progression and consequently also with the liability progression. Again this is not necessarily the case for the degree of progression. Consider two tax schedules with the average tax rate functions $ATR(y)$ and $ATR'(y)$ with positive average rate progressions $\alpha(y) > \alpha'(y) > 0$. Then $\rho(y) < \rho'(y) < 1$ only if $\frac{\alpha(y)}{\alpha'(y)} > \frac{1-ATR(y)}{1-ATR'(y)}$:

$$\begin{aligned}\rho(y) &< \rho'(y) \\ \frac{\partial [y - T(y)]}{\partial y} \frac{y}{[y - T(y)]} &< \frac{\partial [y - T'(y)]}{\partial y} \frac{y}{[y - T'(y)]} \\ \left[1 - \frac{\partial T(y)}{\partial y}\right] \frac{1}{1 - ATR(y)} &< \left[1 - \frac{\partial T'(y)}{\partial y}\right] \frac{1}{1 - ATR'(y)} \\ \left[1 - \frac{\partial ATR(y)}{\partial y} y - ATR(y)\right] \frac{1}{1 - ATR(y)} &< \left[1 - \frac{\partial ATR'(y)}{\partial y} y - ATR'(y)\right] \frac{1}{1 - ATR'(y)} \\ 1 - \frac{\alpha(y)y}{1 - ATR(y)} &< 1 - \frac{\alpha'(y)y}{1 - ATR'(y)} \\ \frac{\alpha(y)}{\alpha'(y)} &> \frac{1 - ATR(y)}{1 - ATR'(y)}\end{aligned}\tag{7}$$

For regressive average tax rates, the inequality in condition (7) changes the direction. Given that $\lambda(y) > \lambda'(y)$ and $1 - \lambda'(y)ATR'(y) > 0$ the ordering of the degree of progression

of the liability and the residual income progression is consistent if:

$$\begin{aligned}
& \rho(y) < \rho'(y) \\
& \frac{\partial [y - T(y)]}{\partial y} \frac{y}{[y - T(y)]} < \frac{\partial [y - T'(y)]}{\partial y} \frac{y}{[y - T'(y)]} \\
& \left[1 - \frac{\partial T(y)}{\partial y}\right] \frac{1}{1 - ATR(y)} < \left[1 - \frac{\partial T'(y)}{\partial y}\right] \frac{1}{1 - ATR'(y)} \\
& [1 - \lambda(y)ATR(y)] \frac{1}{1 - ATR(y)} < [1 - \lambda'(y)ATR'(y)] \frac{1}{1 - ATR'(y)} \\
& \frac{1 - \lambda(y)ATR(y)}{1 - \lambda'(y)ATR'(y)} < \frac{1 - ATR(y)}{1 - ATR'(y)} \tag{8}
\end{aligned}$$

If $1 - \lambda'(y)ATR'(y) < 0$ the inequality in condition (8) changes.

Thus, we have shown that, although all three progressivity measures build on the same idea of progressivity, they are not necessarily consistent in the way they order the degree of progressivity of different tax schedules. Thus, the choice of the progressivity measure is crucial if one does comparisons of tax schedules beyond the simple ordering in regressive and progressive. For the remainder of the paper we will focus on the average rate progression. As we have shown this is sufficient to classify the tax schedules as progressive or regressive. Furthermore, the average rate progression is the most basic definition of this kind of progressivity measure and easy to interpret because it simply answers the question whether the average tax rate is increasing in income, i.e. it is the limit of the difference in the average tax rates over a change in income:

$$\frac{\partial ATR(y)}{\partial y} = \lim_{\Delta y \rightarrow 0} \frac{\Delta ATR(y)}{\Delta y} \tag{9}$$

This enables us to analyse the progressivity not only at income y but also for a certain range of income. Additionally, we can derive very general theoretical results that apply for the average rate progression in the following section.

2.3 Model

Many studies have been dealing with income segregation across different jurisdictions (see e.g. Epple and Sieg (1999), Hardman and Ioannides (2004), Ioannides (2004), Schmidheiny (2006a), Schmidheiny (2006b)). Schmidheiny (2006a) identifies local progressive taxation as a new source of income segregation. Within a two municipality model he shows that the municipality with the lower tax rate must have higher housing prices. Because the marginal income loss of living in the high tax municipality is higher for high income people, they are more likely to live in the low tax municipality. We build up on this two municipality model of Schmidheiny (2006a) except that we relax some assumptions made on the tax schedule. Instead of assuming that the tax functions in both municipalities are a multiple of each other, we only assume that the tax functions in both municipalities do not cross, i.e. the sorting of municipalities into high and low tax

municipalities is perfect and consistent over gross income and that higher tax rates imply higher progressivity.

There are two municipalities $m = 1, 2$ with different progressive tax functions $ATR(y, m)$: $\frac{\partial ATR(y)}{\partial y} > 0 \forall y$. Without loss of generality we assume that $ATR(y, 1) > ATR(y, 2)$ and $\alpha(y, 1) > \alpha(y, 2)$ for all y . The economy consist of three goods: housing, private consumption, and a local public good. There is a continuum of households that differ in income $y \in [\underline{y}, \bar{y}]$ and their taste for housing. Hodler and Schmidheiny (2006) show that in the equilibrium of this model the probability of a household with gross income y living in municipality 1, $P(1|y)$, is weakly decreasing for all y and strictly for some y . Thus, high income people are more likely to live in low tax municipalities. In order to quantify the tax level conditional on the gross income y , they define the *expected average tax rate* $ATR(y)$ which is the expected tax rate a household with gross income y faces in equilibrium unconditionally on his location and housing preferences:

$$\begin{aligned} ATR(y) &= \sum_{m=1}^2 P(m|y) ATR(y, m) \\ &= P(1|y) ATR(y, 1) + [1 - P(1|y)] ATR(y, 2) \end{aligned} \quad (10)$$

This is exactly the measure that integrates out any location and thereby is accounting for income sorting. Consequently, we use this measure throughout the remaining paper as the main measure of the average tax level of the income tax systems. If there was no income sorting this expected average tax rate would equal the *population weighted mean average tax rate*, which we define as:

$$wATR(y) = \sum_{m=1}^2 P(m) ATR(y, m) \quad (11)$$

Due to this equivalence, the population weighted mean average tax rate will serve as a benchmark for the taxation level for the remaining paper because the expected average tax rate exactly equals this benchmark in the absence of income sorting, i.e. if the income distribution of each municipality equals the country-wide distribution: If $f(y) = f(y|m) \forall m = 1, \dots, M$, then

$$\begin{aligned} ATR(y) &= \sum_{m=1}^2 P(m|y) ATR(y, m) \\ &= \sum_{m=1}^2 \frac{f(y|m) P(m)}{\sum_{m=1}^2 f(y|m) P(m)} ATR(y, m) \end{aligned} \quad (12)$$

$$= \sum_{m=1}^2 P(m) ATR(y, m) = wATR(y) \quad (13)$$

This equality make the population weighted mean average tax rate an ideal benchmark for our estimations. Let the difference of both level measures denote as:

$$\Delta(y) = ATR(y) - wATR(y) \quad (14)$$

$$= [P(1|y) - P(1)] [ATR(y, 1) - ATR(y, 2)] \quad (15)$$

Then we can derive the following proposition:

Proposition 1. *If both municipalities are populated and if the housing prices p_m differ and $ATR(y, 1) > ATR(y, 2) \forall y$, then there exists a y^* where $\Delta(y^*) = 0$, $\Delta(y) \geq 0$ for all $y < y^*$, and $\Delta(y) \leq 0$ for all $y > y^*$ and $\Delta(y) > 0$ for some $y < y^*$, and $\Delta(y) < 0$ for some $y > y^*$.*

Proof: $[ATR(y, 1) - ATR(y, 2)] > 0$ by assumption. Further, we know that $P(1|y)$ is decreasing in y , therefore $P(1|\underline{y}) > P(1|\bar{y})$. Because the unconditional probability is just a weighted integral of the conditional probabilities $P(1) = \int_{\underline{y}}^{\bar{y}} f(y)P(1|y)dy$ while $\int_{\underline{y}}^{\bar{y}} f(y)dy = 1$, we know that $P(1|\underline{y}) > P(1) > P(1|\bar{y})$. Therefore, $\Delta(y) > 0$ and $\Delta(\bar{y}) < 0$. From the continuity of $\Delta(y)$ it follows that there must be at least one y^* , where $\Delta(y) = 0$. From the fact that $\frac{\partial P(1|y)}{\partial P(1)} \leq 0$ we know that there exists no y_3 that lies between two y_1^* and y_2^* , both satisfying $\Delta(y_1^*) = \Delta(y_2^*) = 0$ that satisfies $\Delta(y_3) \neq 0$.

Thus, our model predicts that the expected average tax rate of low income people is higher than the population weighted mean average tax rate, while it is lower for people with high incomes.

Beyond the level predictions our model has also some interesting implications for the progressivity of the tax system: Let denote $\alpha(y)$ the average rate progression of the expected average tax rate:

$$\alpha(y) = \frac{\partial ATR(y)}{\partial y} = \sum_{m=1}^2 \frac{\partial P(m|y)}{\partial y} ATR(y, m) + \sum_{m=1}^2 P(m|y) \alpha(y, m) \quad (16)$$

$$= \underbrace{\frac{\partial P(1|y)}{\partial y} [ATR(y, 1) - ATR(y, 2)]}_{\leq 0} + \underbrace{P(1|y) \alpha(y, 1) + [1 - P(1|y)] \alpha(y, 2)}_{E\{\alpha(y, m)\}} \quad (17)$$

Because $\frac{\partial P(1|y)}{\partial y} < 0$ and $ATR(y, 1) - ATR(y, 2) > 0$, the first addend is negative. The second addend is the expected value of the local average rate progression taking the income sorting into account. Therefore, without making any further assumption about the relation between the tax level and the progressivity, we can conclude, that the average rate progression of the income tax system is lower than the expected value of the local average progression.

If a higher tax rate implies a higher average rate progression like in the model⁴ of Hodler and Schmidheiny (2006), we can derive the following relationship in analogy to lemma 1:

$$w\alpha(y) = \frac{\partial w ATR(y)}{\partial y} = \sum_{m=1}^2 P(m) \alpha(y, m) = P(1) \alpha(y, 1) + [1 - P(1)] \alpha(y, 2) \quad (18)$$

⁴This is the case in Switzerland for the local taxes within a canton.

Then the difference of this progression measure and the average rate progression of the expected average tax rate is:

$$\begin{aligned}\delta(y) &= \alpha(y) - w\alpha(y) \\ &= \underbrace{\frac{\partial P(1|y)}{\partial y} [ATR(y, 1) - ATR(y, 2)]}_{\leq 0} + [P(1|y) - P(1)][\alpha(y, 1) - \alpha(y, 2)]\end{aligned}\quad (19)$$

Proposition 2. *If both municipalities are populated and if the housing prices p_m differ and $ATR(y, 1) > ATR(y, 2)$ and $\alpha(y, 1) > \alpha(y, 2) \forall y$, then there exists a y^{**} such that for all $y > y^{**}$ $\delta(y) \leq 0$ and $\delta(y) < 0$ some.*

Proof: The proof directly follows as a corollary from proposition 1 and the fact that $ATR(y, 1) > ATR(y, 2)$ and $\alpha(y, 1) > \alpha(y, 2)$ and the strict negativity of the first addend. Note that we cannot make a similar statement for any $y < y^{**}$.

Therefore, our model not only implies that level of the expected average tax rate is below the population weighted mean average tax rate for high incomes but also that its progressivity is lower than the one of the benchmark.

3 Method

This section describes the estimation of the country wide average tax rate function and its progressivity. Furthermore, based on the discussion above, we derive benchmark measures that help interpreting these estimates.

We have data on the average tax rates on the federal, cantonal, and municipal level including church taxes at several income points⁵. We are mainly interested in the overall income tax burden, therefore we sum all income tax burdens:

$$\begin{aligned}ATR(y, m) &= \frac{T(y, m)}{y} = \frac{(1 + M_m + K_m)T^c(y, m) + T^f(y)}{y} \\ &= (1 + M_m + K_m)ATR^c(y, m) + ATR^f(y) \\ \frac{\partial ATR(y, m)}{\partial y} &= (1 + M_m + K_m)\frac{\partial ATR^c(y, m)}{\partial y} + \frac{\partial ATR^f(y)}{\partial y}\end{aligned}$$

Both, the $ATR(y, m)$ and the progressivity measure are location-specific, i.e. *conditional* on the jurisdiction. In order to quantify the country-wide progressivity, we seek a country-wide, i.e. an *unconditional* progressivity measure. Such a measure must account for the fact that high income households may systematically sort into low-tax locations. As discussed above, the following measure for the average tax rate satisfies these conditions:

⁵see data section for discussion

$$ATR(y) = \sum_{m=1}^M P(m|y) ATR(y, m) = \sum_{m=1}^M \frac{P(m)f(y|m)}{f(y)} ATR(y, m) \quad (20)$$

While we observe the relative amount of tax payers $P(m)$ living in municipality m and the average tax rate $ATR(y, m)$, we do not observe the continuous income distributions $f(y)$ and $f(y|m)$. Hodler and Schmidheiny (2006) estimate these functions assuming log-normal income distributions, thereby imposing parametric restrictions. Especially for municipalities with very few observations, it is questionable if we can get reliable density estimates. The last argument also holds for non-parametric kernel density estimations. Kernel density estimations for 3000 municipalities are only feasible if we use a rule for the bandwidth choice (Silverman, 1986). But in order to be confident with the results one would have to check all 3000 density estimates. Therefore, we suggest another way, that is based on the idea of taking sample averages of the average tax rates that people with income y pay. We can show that this leads to an estimator of $ATR(y)$:

$$\widehat{ATR}(y) = \frac{1}{N_y} \sum_{i|y_i=y}^{N_y} ATR_i \quad (21)$$

$$\begin{aligned} &= \frac{1}{N_y} \sum_{m=1}^M \sum_{i|y_i=y,m}^{N_{y,m}} ATR_i \\ &= \frac{1}{N_y} \sum_{m=1}^M N_{y,m} ATR(y, m) = \sum_{m=1}^M \frac{N_{y,m}}{N_y} ATR(y, m) \\ &= \sum_{m=1}^M P(\hat{m}|y) ATR(y, m) \end{aligned} \quad (22)$$

Unfortunately, this estimator in its simple form is not consistent because the likelihood of observing an individual with income y is by construction zero:

$$\widehat{P(m|y)} = \frac{N_{y,m}}{N_y} \rightarrow P(m|y) \quad (23)$$

$$\Rightarrow \widehat{ATR}(y) \rightarrow ATR(y) \quad (24)$$

The solution to the problem is local fitting. Because we are not only interested in an estimator of $ATR(y)$ but also in an estimator for the average rate progression $\frac{\partial ATR(y_0)}{\partial y}$, we take data on individual taxpayers and calculate for each taxpayer the average tax rates at his residential municipality m and perform a local polynomial regression of these average tax rates on the natural logarithm⁶ of the gross income. Suppose, the individual average tax rate ATR_i is a function of the country-wide average tax rate $ATR(\ln(y_i))$ and an error term $\sigma(\ln(y_i))\epsilon_i$:

$$ATR_i = ATR(\ln(y_i)) + \sigma(\ln(y_i))\epsilon_i \quad (25)$$

⁶The income distribution is extremely positively skewed. The transformation leads to a far less skewed distribution which is easier to handle using a fixed bandwidth and symmetric kernels.

In the neighbourhood of $\ln(y_0)$ we can approximate the country-wide average tax function with a Taylor polynomial of order p :

$$\begin{aligned}
ATR(\ln(y)) \approx & ATR(\ln(y_0)) + \frac{\partial ATR(\ln(y_0))}{\partial \ln(y)} (\ln(y) - \ln(y_0)) \\
& + \frac{\frac{\partial^2 ATR(\ln(y_0))}{\partial \ln(y)^2}}{2!} (\ln(y) - \ln(y_0))^2 \\
& + \dots \\
& + \frac{\frac{\partial^p ATR(\ln(y_0))}{\partial \ln(y)^p}}{p!} (\ln(y) - \ln(y_0))^p
\end{aligned} \tag{26}$$

This polynomial can be fitted by a local weighted least squares estimation:

$$\min_{\beta} \sum_i^N \left\{ ATR_i \sum_{j=0}^p \beta_j (\ln(y_i) - \ln(y_0))^j \right\} K_h(\ln(y_i) - \ln(y_0)) \tag{27}$$

where $K_h(\cdot)$ are Epanechnikov kernel weights with bandwidth h .

Applying a polynomial of degree $p > 0$ allows us also to estimate the average rate progression simultaneously. At income y_0 , this method results in consistent⁷ estimates of $ATR(y_0)$ and $\frac{\partial ATR(y_0)}{\partial y}$:

$$\widehat{ATR}(y_0) = \hat{\beta}_0 \tag{28}$$

$$\frac{\partial \widehat{ATR}(y_0)}{\partial y} = \frac{\hat{\beta}_1}{y_0} \tag{29}$$

Note that this estimated $ATR(y)$ is *not* conditional on an individual municipality m . Our local polynomial regression therefore integrates the location m out - accounting for the actually observed income sorting. The resulting $ATR(y)$ is basically a weighted mean of the conditional $ATR(y, m)$ weighted by the number of taxpayers with income y in municipality m . Because the estimates are local for y_0 , we neither make any assumption on the functional form of $ATR(y)$. Due to our full population income data we have enough observations to apply this estimation strategy up to very high income levels y_0 ⁸.

As a benchmark we also estimate the weighted mean of the the local total average tax rates $ATR(y, m)$ weighted by the local population N_m ⁹:

$$wATR(y) = \sum_{m=1}^M P(m) ATR(y, m) = \sum_{m=1}^M \frac{N_m}{\sum_{m=1}^M N_m} ATR(y, m) \tag{30}$$

This weighted mean would be equal to the above $ATR(y)$ in the absence of income sorting, i.e. if the income distribution in each municipality would exactly equal the country-wide

⁷The local polynomial regression estimator is generally biased. But including higher order polynomials reduces this bias especially at the boundaries (see e.g. Fan and Gijbels (1996)).

⁸In the standard setting, we perform this estimation for the following income levels (in CHF): 20'000; 25'000; 30'000; 35'000; 40'000; 50'000; 60'000; 70'000; 80'000; 90'000; 100'000; 150'000; 200'000; 300'000; 500'000; 1'000'000; 1'500'000; 2'000'000; 2'500'000; 3'000'000; 4'000'000; 5'000'000; 10'000'000

⁹Amount of households of a type in municipality m . See data section for discussion.

distribution. Thus, the difference between the weighted mean of the the local total average tax rates at income y_0 and the estimate $\widehat{ATR}(y_0)$ is the effect that the income sorting has on the average tax rate at income y_0 .

We conduct this estimation for two typical household types: single households and families with two children. For each type, we perform three estimations: first with the full Swiss sample (CH), second with a sample containing people living in the agglomeration of Zurich (ZH), and third with a sample with people living in the agglomeration of Basel (BS).

Additionally, we repeat all estimations without taking the very progressive federal income tax into account. This results in 12 different specifications. Table 2 gives an overview. We focus on this reduced tax burden in order to investigate the impact of the federal tax on the progressivity of the income tax system. By construction the estimates of the average tax rate will just differ by the level of the federal average tax rate $ATR^f(y)$. The difference between the expected average tax rate and the population weighted mean average tax rate remains the same because the federal average tax rate cancels:

$$ATR(y) = ATR^f(y, m) + \sum_{m=1}^M P(m|y)ATR^l(y, m) = ATR^f(y, m) + ATR^l(y) \quad (31)$$

$$wATR(y) = ATR^f(y, m) + \sum_{m=1}^M P(m)ATR^l(y, m) = ATR^f(y, m) + wATR^l(y) \quad (32)$$

Also the difference of the average rate progression of both measures is independent of the federal tax rate:

$$\alpha(y) = \frac{\partial ATR^f(y, m)}{\partial y} + \alpha^l(y) \quad (33)$$

$$w\alpha(y) = \frac{\partial ATR^f(y, m)}{\partial y} + w\alpha^l(y) \quad (34)$$

However, from equation 33 we clearly see that the progressivity of the federal income tax matters for the progressivity of the entire income tax system. Because the progressivity of the local and federal taxes simply add up, the federal income tax can serve as a policy instrument to overcome too low progressivity that might be generated by the local part of the income tax system.

Table 2: Sample Specifications

Sample	Civil Status	Federal Tax	Area
[1]	Singles	included	Switzerland
[2]	Singles	included	Zurich
[3]	Singles	included	Basel
[4]	Singles	excluded	Switzerland
[5]	Singles	excluded	Zurich
[6]	Singles	excluded	Basel
[7]	Families	included	Switzerland
[8]	Families	included	Zurich
[9]	Families	included	Basel
[10]	Families	excluded	Switzerland
[11]	Families	excluded	Zurich
[12]	Families	excluded	Basel

4 Data

We use the universe of uncensored individual incomes of Swiss taxpayers in 2009. This data has been made accessible by the Sinergia Project 130648 founded by the Swiss National Science Foundation in cooperation with the Swiss Federal Tax Administration (ESTV). We focus on two typical household types: single households and married couples with two children. In 2009, we observe data on 1'649'167 single households and 303'283 families with two children.

Average Tax Rates

The Swiss Federal Tax Administration annually publishes the tax burden for a sample of Swiss municipalities¹⁰. This tax burden includes cantonal, municipal, and church taxes. The same publication also includes federal tax burdens. All tax burdens are calculated for several representative households at certain gross income levels¹¹. Because the statutory tax rates are defined over the taxable income rather than the gross income this data contain also implicit deductions at cantonal and federal level.

For the year 2009, this source delivers data on average tax rates for 813 municipalities. Within the Sinergia Project, Marius Brühlhart, Raphael Parchet, and Kurt Schmidheiny calculate the average tax rate at these income levels for the missing municipalities. They collect data on all municipal M_m and church K_m multipliers and exploit the special functional relationship of the local average tax rates $ATR^l(y, m)$ within a canton:

¹⁰Swiss Federal Tax Administration, Steuerbelastung in der Schweiz, Neuchâtel: Swiss Federal Statistical Office.

¹¹20'000; 25'000; 30'000; 35'000; 40'000; 45'000; 50'000; 60'000; 70'000; 80'000; 90'000; 100'000; 150'000; 200'000; 300'000; 400'000; 500'000; 1'000'000;

$$ATR^l(y, m) = \frac{(1 + M_m + K_m)T_m^c(y)}{y} \quad (35)$$

Because the local average tax rates of the municipalities within a canton differ only by the tax multipliers, knowing the tax multipliers is sufficient to construct the average tax rates for the missing municipalities. Thus, we have a data set a hand that contains the average tax rates for 2624 Swiss municipalities in 2009.

Income Data

Our data set contains the universe of uncensored individual taxable incomes of all federal tax payers in Switzerland¹² in the year 2009. Because we do not have data on the individuals' gross income, we need to approximate it. As mentioned above, the annual publication of the Swiss Federal Tax Administration also contains average tax rates on the federal level. From this, we calculate the tax burdens at all published income levels:

$$T^f(y) = ATR^f(y) * y \quad (36)$$

For both household types, singles and families with two kids we can calculate the implicit deductions. The statutory tax¹³ burden $S(y^t)$ is defined over the taxable income $y^t = y - d$. Because the deductions are the only unknown in the following equation, we can calculate the deductions d that the Swiss Federal Tax Administration assumed for the respective household type with gross income y .

$$T^f(y) = S(y^t) = S(y - d) \quad (37)$$

$$d(y) = y - S^{-1}(T^f(y)) \quad (38)$$

It turns out that the relationship between the deductions and the gross income that the ESTV assumed can be very well approximated by the following linear relationship:

$$d = a + by \quad (39)$$

After calculating all deduction for the income levels for both household types separately, we regress the the implicit deduction on the gross income levels. Table 3 represents the results. While singles are assumed to deduct a fix amount of 4679 CHF and 10% of their gross income, families with two kids are assumed to deduct a fix amount of 22'912 CHF and also 10% of their gross income.

Note that calculating the deductions this way does not impose any further assumptions on the relationship between gross income and deductions because this relationship is already implied in the data on tax burdens by the ESTV that we use to calculate the average tax

¹²Also tax payers with a taxable income of zero are included.

¹³Swiss Federal Tax Administration: Table for calculation of the federal income tax of natural persons (Art. 214 DBG)

Table 3: Deductions

	Singles	Families
\hat{a}	4679	22912
\hat{b}	0.1032	0.1036

rate data. We also avoid problems that arise because the federal and the cantonal taxable income differ.

Finally, we calculate the individual gross incomes for each individual from the individual taxable income y_i^t as follows:

$$y_i = \frac{y_i^t + \alpha}{1 - \beta} \quad (40)$$

Individual Average Tax Rates

In the final step, we calculate the individual tax data. For individuals with a gross income below 1'000'000 CHF we interpolate the average tax rates with the average tax rates of the next gross income steps in the residential municipality we have data on. For individuals with a gross income above 1'000'000 CHF, we calculate the average tax rate as follows:

$$ATR_i = \frac{\theta(m_i)(y_i - 1'000'000) + T(1'000'000, m_i)}{y_i} \quad (41)$$

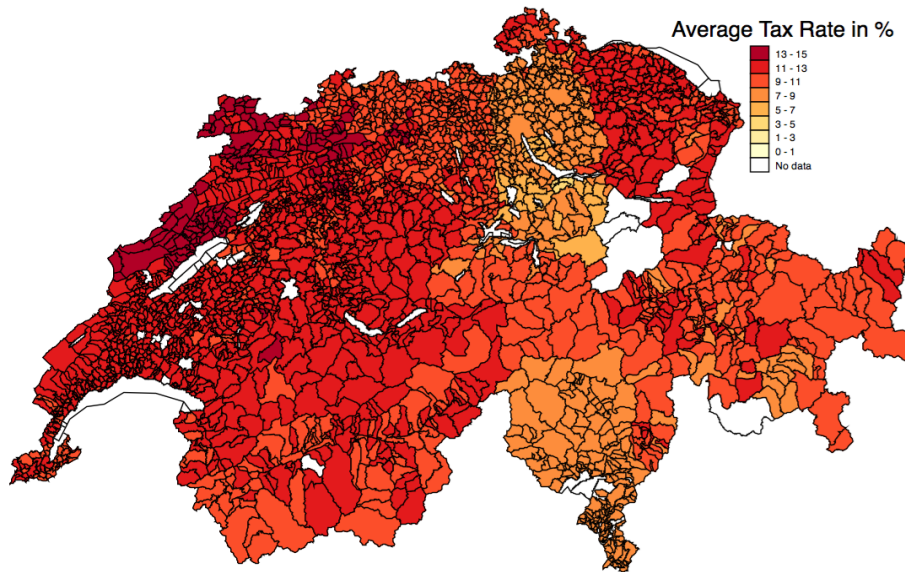
$$\theta(m_i) = \frac{T(1'000'000, m_i) - T(500'000, m_i)}{500'000} \quad (42)$$

This method assumes that the maximum marginal tax rate $\theta(m)$ is already reached at a gross income of 500'000 CHF. Furthermore, the deductions must grow linearly in the gross income which is already assumed in the average tax rate data. We repeat these calculations neglecting the federal average tax rates.

Finally, for singles and families with two kids, we get a data set that contains the gross income, the local average tax rates, and the total average tax rates of all taxpayers. Figure 1 and 2 show the average tax rates that a household with median taxable income¹⁴ faces in the different municipalities. We see that the average tax rates across Switzerland indeed differ a lot between municipalities for both household types and that the level of taxation for the families is lower than for singles.

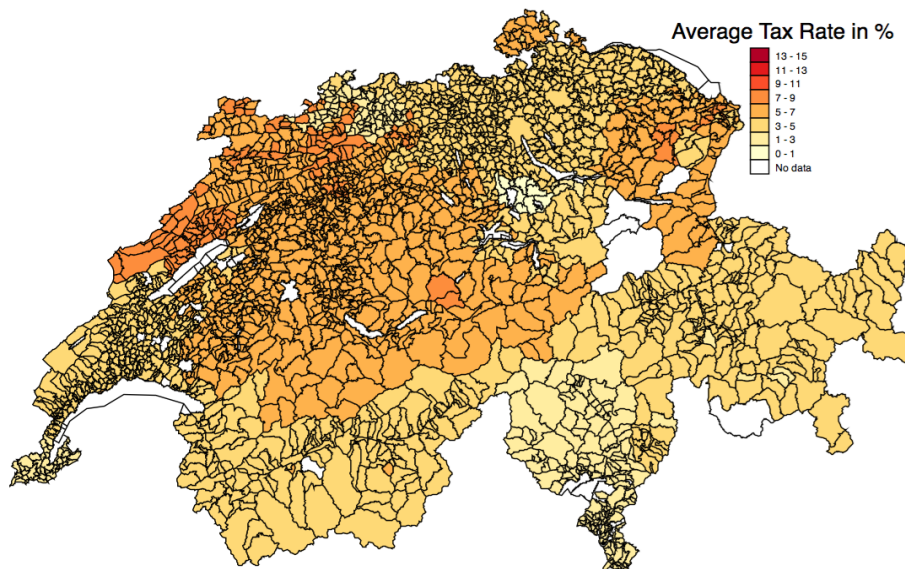
¹⁴Here: the median taxable income of all Swiss tax payers including zero taxable incomes. From this we construct the respective gross incomes.

Figure 1: Total Average Tax Rates for an Income of Singles at Median Income



Notes: The average tax rates are calculated for a median taxable income of 44'500CHF. Using the described transformation this results in a gross income of 54'643 CHF.

Figure 2: Total Average Tax Rates for an Income of Families at Median Income



Notes: The average tax rates are calculated for a median taxable income of 44'500CHF. Using the described transformation this results in a gross income of 74'902 CHF.

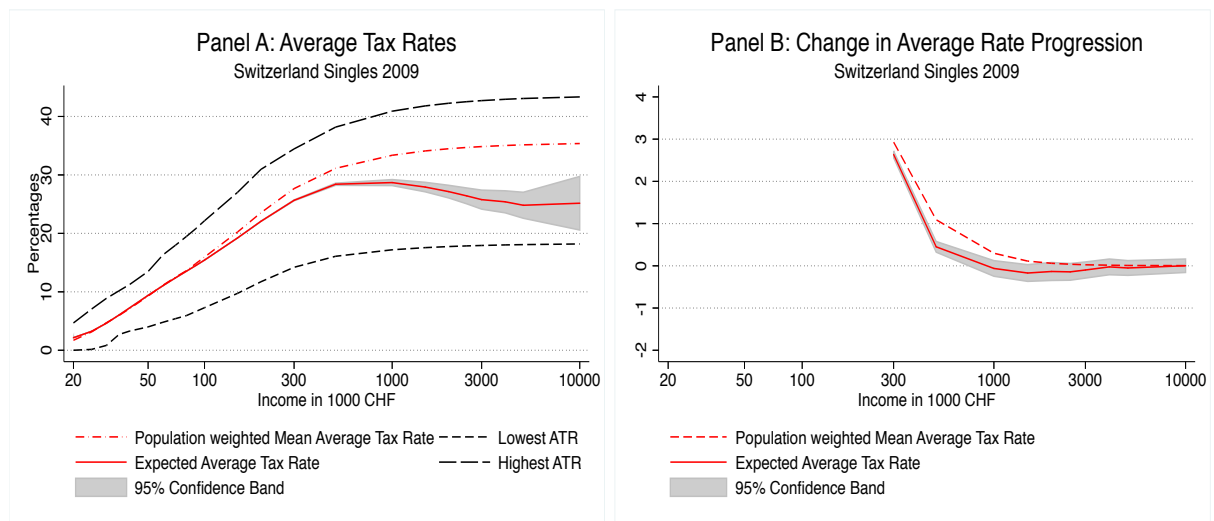
5 Results

5.1 Singles

The left panel of figure 3 shows the results from the estimation of the total expected average tax rate function for singles including all Swiss municipalities in 2009. Starting from 100'000 CHF on, the expected average tax rate is significantly lower than the population weighted mean of $ATR(y, m)$. The gap is constantly increasing up to about ten percentage points for an income of 10 Mio. CHF. This difference reflects the effect of the income sorting. Thus, tax payers with incomes above 100'000 CHF live in municipalities with comparably low tax rates.

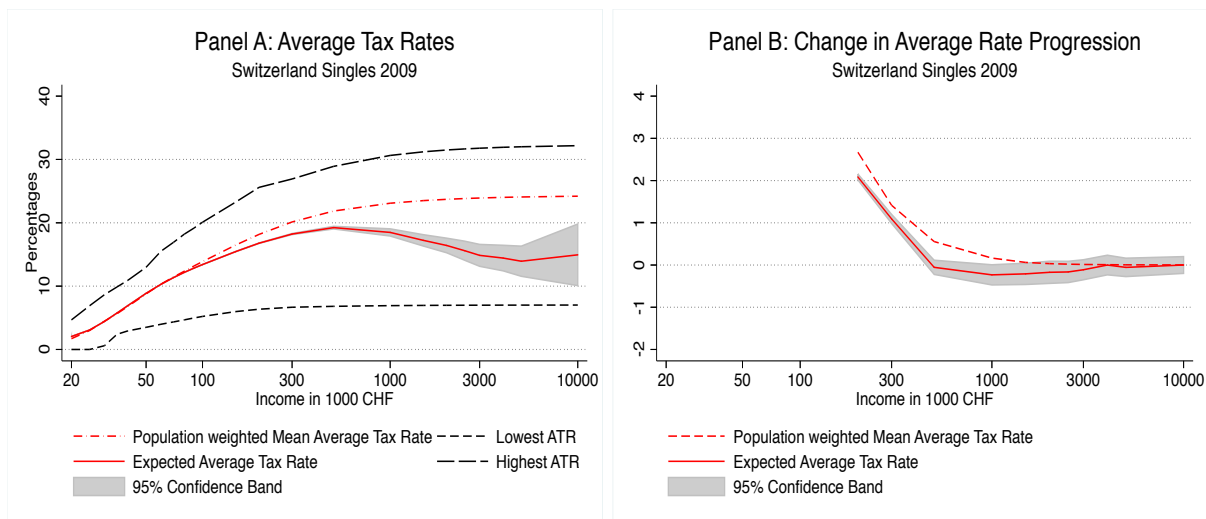
Although the expected average tax rate starts flattening and decreasing after 1 Mio. CHF, the average rate progression is never significantly negative as the right panel of figure 3 indicates. From 100'000 CHF to 2 Mio. CHF, the average rate progression of the expected average tax rate is significantly lower than the average rate progression of the population weighted mean average tax rate. Hence, the income sorting lowers the progressivity for singles within that income range. Above 2 Mio CHF, we cannot observe any significant difference anymore, because the the average rate progression of the expected average tax rate is rising again, while the benchmark average rate progression keeps falling. At least partly, the disappearance of the difference can be accounted to the fact that the limit of the benchmark average tax rate must be zero as all marginal tax rates are constant.

Figure 3: Average Tax Rate and Progressivity - Switzerland: Singles



Notes: Number of observations: 1'649'167; Polynomial degree: 3; Kernel-weights: Epanechnikov; Half-log-bandwidth: 0.343; reported average rate progressions are multiplied by 100'000.

Figure 4: Local Average Tax Rate and Progressivity - Switzerland: Singles



Notes: Number of observations: 1'649'167; Polynomial degree: 3; Kernel-weights: Epanechnikov; Half-log-bandwidth: 0.299; reported average rate progressions are multiplied by 100'000.

If we neglect the very progressive federal tax in our calculation, we find a similar picture. The expected average tax rate is significantly lower than the benchmark for incomes above 60'000 CHF and starts decreasing as of an annual income of 500'000 CHF. The decrease is stronger than before because the high progressivity of the federal income tax is not considered. Despite neglecting the high progressivity of the federal income tax, figure 4 suggests that the expected average tax rate is not significantly regressive.

Although in both specifications we find significantly less progressive expected average tax rates compared to the benchmark, we do not find significant regressivity which seems surprising because the expected average tax rates in the left panel is strongly decreasing. Therefore, we also test the differences in the expected average tax rates if the income increases from 500'000 CHF to the higher income levels. We can do so by applying a simple t-test because the bandwidth is small enough, i.e. the estimates at all higher income points do not include observations used for the estimation of the expected average tax rate at 500'000 CHF. Table 4 reports the results. If we include the federal income tax, we can conclude that for all income levels from 2 Mio. CHF to 5 Mio. CHF the expected average tax rate is significantly smaller than for an income of 500'000 CHF. We find the highest difference at 5 Mio. CHF. These people face an average tax rate that is 3.6 percentage points smaller than the one people with an income of 500'000 face. If we neglect the federal tax rate that cannot be avoided by moving to another municipality, we find even stronger effects. At all income levels above 500'000 CHF, people bare significantly lower average tax rates. Also in this specification, we find the highest difference for an income of 5 Mio. CHF. These people have 5.3 percentage points lower local average tax rates than people with an income of 500'000 CHF. Therefore, these results clearly indicate that the Swiss income tax system is regressive for incomes between 1 Mio. and 5 Mio

Table 4: Differences to the expected average tax rate of 500'000 CHF

Income	with federal tax		without federal tax	
	CH	ZH	CH	ZH
1'000'000	0.279 (0.277)		-0.757*** (0.307)	
1'500'000	-0.487 (0.425)	-1.648*** (0.575)	-2.006*** (0.468)	-3.059*** (0.600)
2'000'000	-1.267** (0.559)	-2.915*** (0.718)	-2.817*** (0.597)	-4.495*** (0.749)
2'500'000	-2.003*** (0.699)	-3.326*** (0.861)	-3.652*** (0.777)	-5.069*** (0.889)
3'000'000	-2.639*** (0.836)	-3.121*** (0.996)	-4.379*** (0.888)	-4.929*** (1.025)
4'000'000	-3.01*** (0.966)	-2.803** (1.249)	-4.806*** (1.036)	-4.496*** (1.273)
5'000'000	-3.601*** (1.138)	-2.485* (1.376)	-5.304*** (1.220)	-4.411*** (1.393)
10'000'000	-3.254 (2.340)	-2.599 (3.360)	-4.295* (2.482)	-4.419 (3.457)
N	1'649'167	286'831	1'649'167	286'831

Notes: Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

CHF.

These results are not contradicting because the test if a function is decreasing in a point, i.e. the derivative is smaller than zero, is not the same as the test if a function is decreasing between two points. Yet, the derivative of the average tax rate is just the limit of the difference in between two income levels when the difference in income gets zero:

$$\frac{\partial ATR(y)}{\partial y} = \lim_{\Delta y \rightarrow 0} \frac{\Delta ATR(y)}{\Delta y} \quad (43)$$

Thus, the insignificance of the regressivity in the average rate progression is caused by taking exactly these limits and not by insignificant differences in the tax rates. Consequently, we can conclude that we find suggestive evidence that the Swiss income tax system for singles is regressive for incomes above 500'000 CHF.

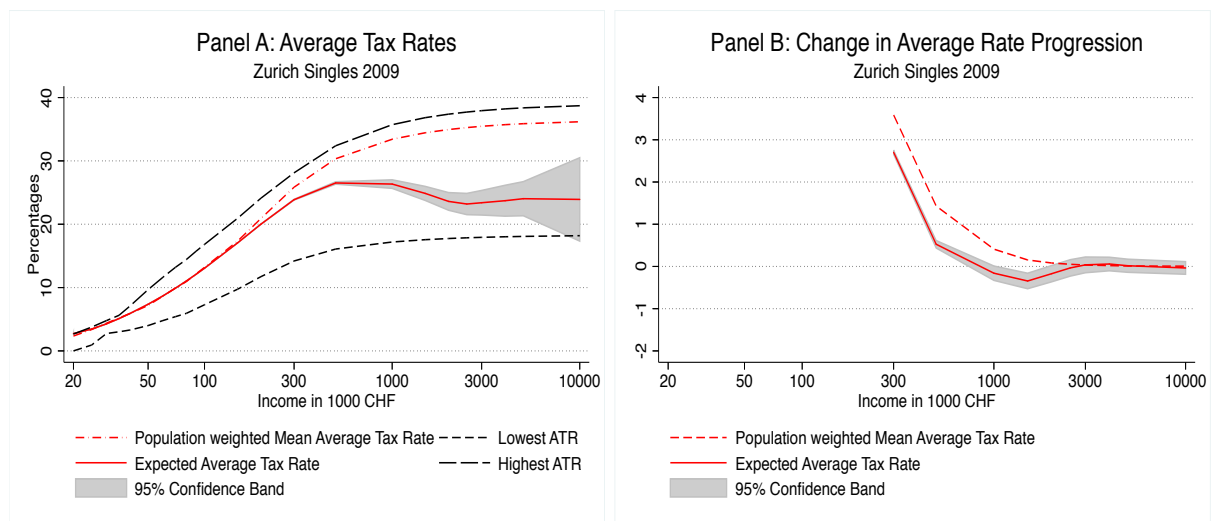
Agglomeration of Zurich

One might be concerned that people rather choose their residential municipality locally than out of all municipalities in Switzerland. Therefore, we repeat the analysis considering only municipalities in the agglomeration of Zurich, the most important economic

agglomeration in Switzerland. Figure 5 presents the results of the estimations including the federal tax rate. We observe an increasing gap between the average tax rates between the lowest and highest possible average tax rates up to 20 percentage points for an income of 10 Mio. CHF which generates large possible gains of moving to the lower tax municipalities. The population weighted mean average tax rate is comparably high which is caused by the city of Zurich having a very high population share and comparably high average tax rates. The expected average tax rate starts being significantly lower than this benchmark as of an income of 200'000 CHF. This gap is increasing up to about 15 percentage points for incomes above 3 Mio. CHF. As of an income of 1 Mio. CHF the expected average tax rate starts declining up to an income of 3 Mio. CHF. Unlike in the case for Switzerland as a whole, we find the average rate progression being significantly negative for incomes between 1 Mio. CHF and 2 Mio. CHF as panel B of figure 5 indicates. If we neglect the federal income tax, these results are even stronger (see figure 6).

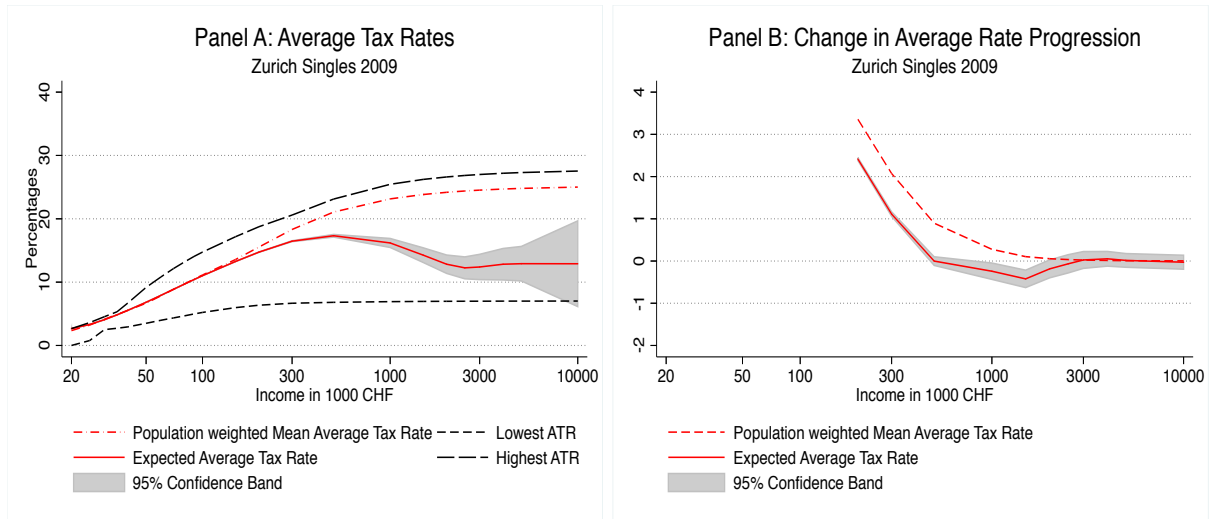
If we consider again the differences in expected average tax rates to the expected average tax rate of an income of 500'000 CHF, we find confirmative results (see table 4). The total expected average tax rates for incomes of 1.5 Mio. up to 5 Mio. CHF are all significantly lower than the expected average tax rate for an income of 500'000 CHF. We observe the highest difference between the benchmark income and an income of 2.5 Mio. CHF neglecting the federal tax: These singles pay 5 percentage points less local income taxes than those that have an annual income of 500'000 CHF. Therefore, we can conclude that the income tax system for singles considering only the agglomeration of Zurich is unambiguously regressive at least for incomes between 1 Mio. CHF and 2 Mio. CHF.

Figure 5: Average Tax Rate and Progressivity - Zurich: Singles



Notes: Number of observations: 286'831; Polynomial degree: 3; Kernel-weights: Epanechnikov; Half-log-bandwidth: 0.459; reported average rate progressions are multiplied by 100'000.

Figure 6: Local Average Tax Rate and Progressivity - Zurich: Singles



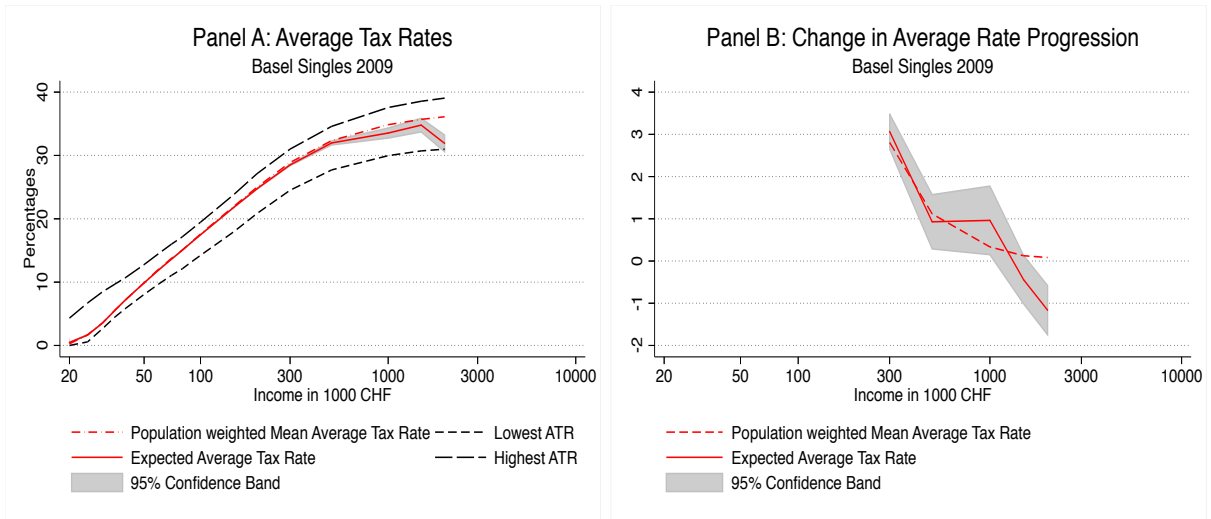
Notes: Number of observations: 286'831; Polynomial degree: 3; Kernel-weights: Epanechnikov; Half-log-bandwidth: 0.437; reported average rate progressions are multiplied by 100'000.

Agglomeration of Basel

We repeat the same analysis for the agglomeration of Basel. Figure 7 presents the results for singles including the federal tax. One can immediately see that the gap between the lowest average tax rates and the highest is smaller than in the agglomeration of Zurich. Up to an income of 500'000 CHF the expected average tax rate and the population weighted mean average tax rate are basically the same. Thereafter, they start deviating slowly. However, only for incomes above 2.5 Mio. CHF, we observe a statistically and economically significant difference of -4.22 percentage points. This sharp decline is also reflected in the average rate progression where we get a negative estimate for an income of 2.5 Mio. CHF. The estimates without the federal tax confirm these results (see figure 8). Despite this evidence for regressivity, the estimates have to be handled with care because we do not observe enough high income individuals in the area of Basel in order to estimate the expected average tax rate and its progressivity for income levels higher than 2.5 Mio.

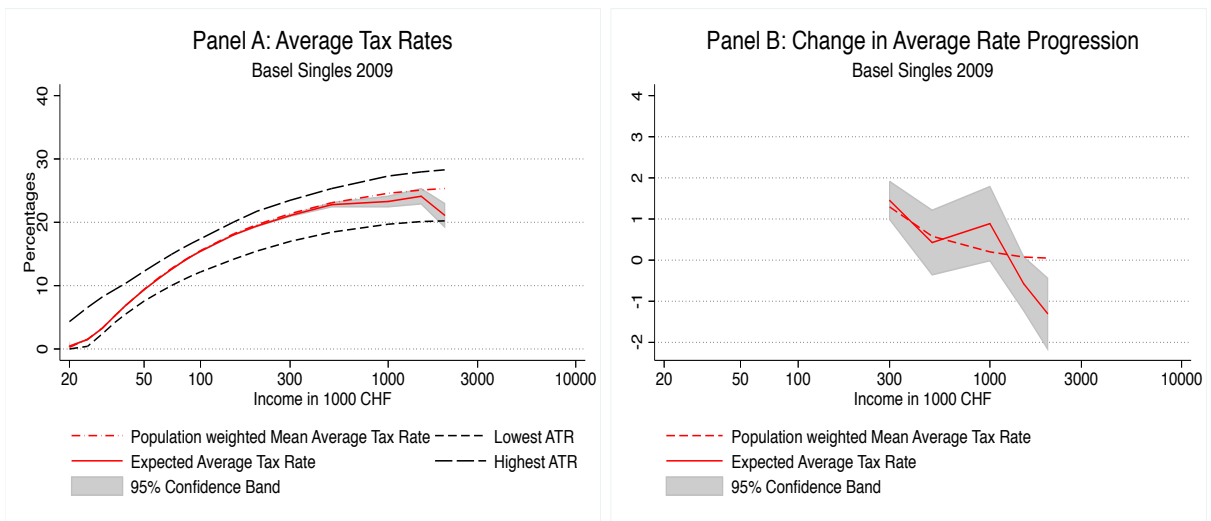
Nevertheless, the example of Basel also provides two interesting insights: First, the actual tax avoidance by moving to low tax municipalities seems to be far more important in agglomerations where the spread between high and low tax municipalities is high, thus people have the possibility to save large amounts of taxes. Second, the fact that we observe relatively few high income people and the taxes within the agglomeration in Basel are comparably high might be an indicator that these people optimise their tax payments at least on national level.

Figure 7: Average Tax Rate and Progressivity - Basel: Singles



Notes: Number of observations: 104'685; Polynomial degree: 1; Kernel-weights: Epanechnikov; Half-log-bandwidth: 0.078; reported average rate progressions are multiplied by 100'000.

Figure 8: Local Average Tax Rate and Progressivity - Basel: Singles



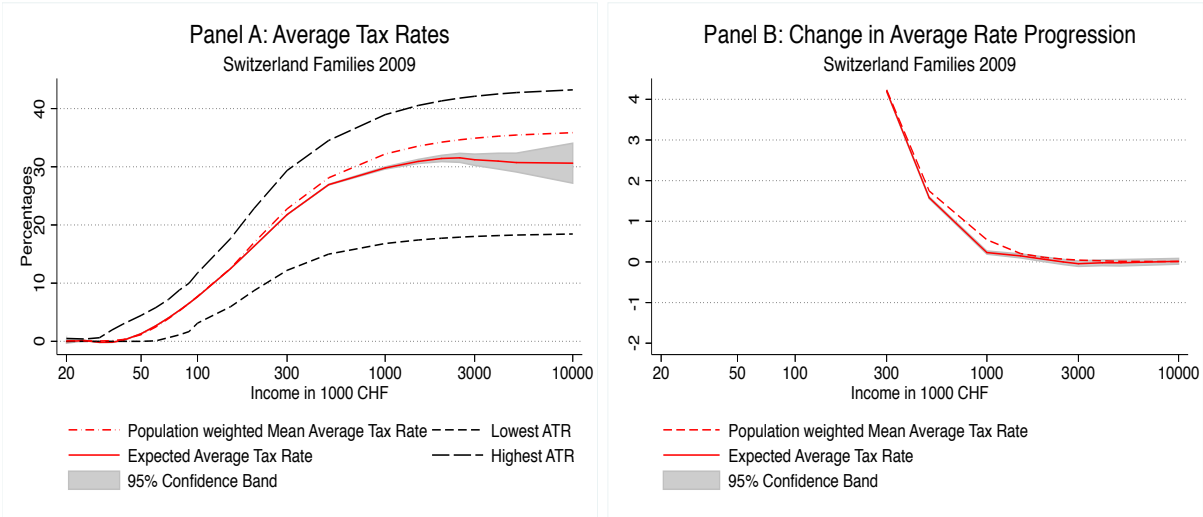
Notes: Number of observations: 104'685; Polynomial degree: 1; Kernel-weights: Epanechnikov; Half-log-bandwidth: 0.072; reported average rate progressions are multiplied by 100'000.

5.2 Families with two Children

This section focuses on the analysis of the expected average tax rates and its progressivity for a typical Swiss family (married couple with two children). Again we performed the analyses for Switzerland and the agglomerations of Zurich and Basel. Figure 9 presents the results for Switzerland including the federal tax rate. The possibilities of saving taxes

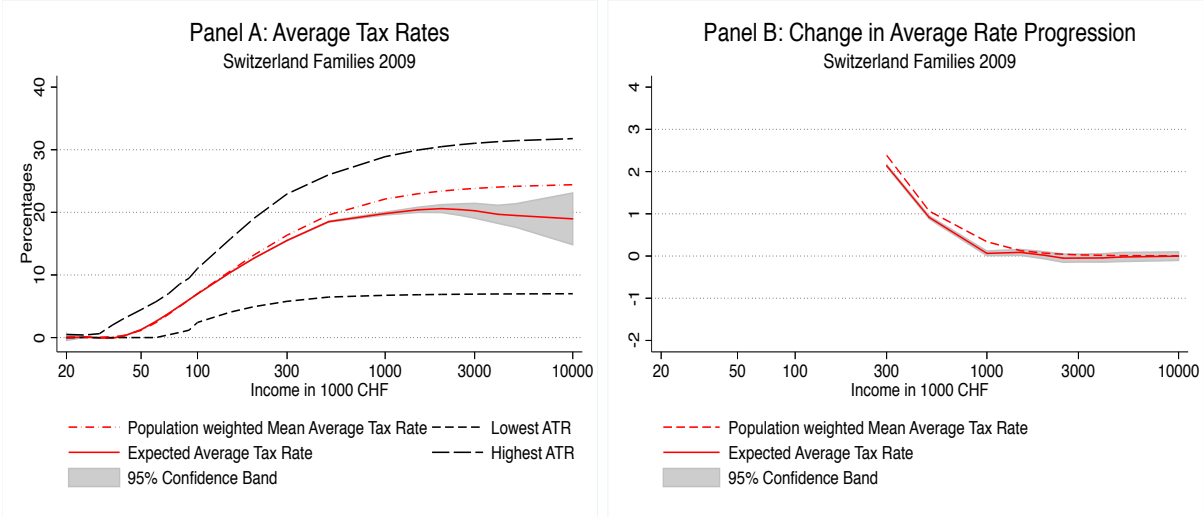
by moving to low tax municipalities is even larger than for singles. For an income of 10 Mio. the difference in the average tax rate between the highest tax municipality and the lowest is 24.8 percentage points. Despite this high saving potential, the actual deviation of the expected average tax rate from the population weighted mean average tax rate is lower than for singles. Starting from 100'000 CHF the expected average tax rate is lower than its benchmark but the difference reaches only -5.24 percentage points for an income of 10 Mio. CHF. Neither, the slope of the expected average tax rate is significantly negative as panel B indicates. Therefore, we also find that families with two children and an income of more than 100'000 CHF live in municipalities with comparably low taxes, however they are far away from pure tax optimisation. The income sorting effect here is too small in order to turn progressivity of the expected average tax rate negative. Thus, we do not find evidence that the Swiss income tax system gets regressive at any income level for families with two children.

Figure 9: Average Tax Rate and Progressivity - Switzerland: Families with 2 kids



Notes: Number of observations: 303'283; Polynomial degree: 3; Kernel-weights: Epanechnikov; Half-log-bandwidth: 0.516; reported average rate progressions are multiplied by 100'000.

Figure 10: Local Average Tax Rate and Progressivity - Switzerland: Families with 2 kids



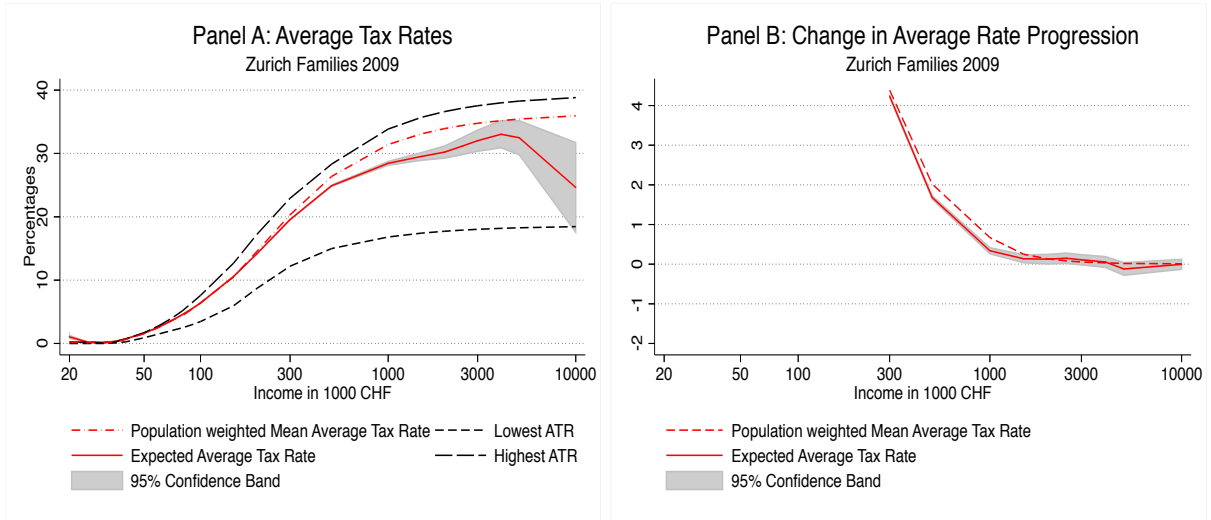
Notes: Number of observations: 303'283; Polynomial degree: 3; Kernel-weights: Epanechnikov; Half-log-bandwidth: 0.440; reported average rate progressions are multiplied by 100'000.

Agglomeration of Zurich

In the agglomeration of Zurich, we observe a similar picture (see figure 11). The expected average tax rate is significantly lower than the population weighted mean average tax rates for incomes higher than 200'000 CHF. This difference stays small up to an income of 5 Mio. CHF. Beyond that, it is increasing up to -11 percentage points for an income of 10 Mio. CHF. But also the confidence interval increases which makes any difference test insignificant. The average rate progression estimates confirm these results.

The estimates without the federal tax rates neither deliver evidence of any regressivity (see figure 12). Despite the huge dip in the average tax rates from 5 to 10 Mio. CHF, we find no significant evidence that the income tax system for families with two children in the agglomeration in Zurich is significantly regressive. But still we observe a significantly negative effect of the income sorting on the average tax rates of high income households. Yet, this effect is much smaller compared to singles.

Figure 11: Average Tax Rate and Progressivity - Zurich: Families with 2 kids



Notes: Number of observations: 41'857; Polynomial degree: 3; Kernel-weights: Epanechnikov; Half-log-bandwidth: 0.461; reported average rate progressions are multiplied by 100'000.

Figure 12: Local Average Tax Rate and Progressivity - Zurich: Families with 2 kids



Notes: Number of observations: 41'857; Polynomial degree: 3; Kernel-weights: Epanechnikov; Half-log-bandwidth: 0.764; reported average rate progressions are multiplied by 100'000.

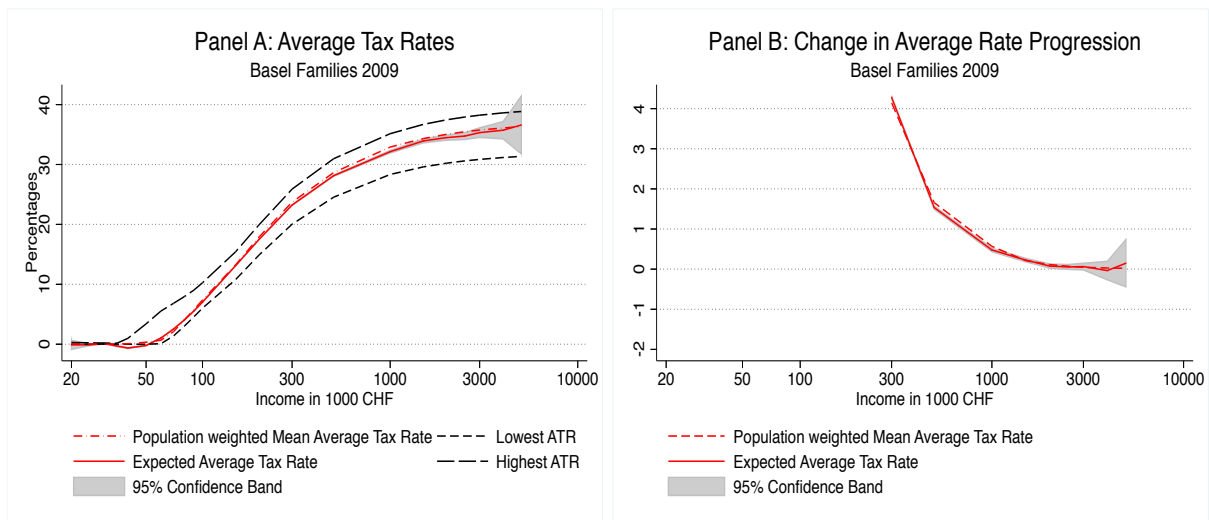
Agglomeration of Basel

The results for the agglomeration of Basel are presented in figure 13. Although the expected average tax rate is significantly below the the population weighted mean average tax rate for certain income levels, it is obvious that the economic magnitude is negligible. Instead of a decreasing progressivity we even observe a rising progressivity of the expected average tax rate for incomes above 4 Mio. CHF. This observation is stronger if we exclude

the federal tax rate (see figure 14). This result confirms our findings from above that the possibilities to avoid taxes in the agglomeration of Basel are fairly small which leads to a much smaller effect of income sorting. We are not able to produce estimates for incomes of more than 5 Mio. CHF, which again indicates that these very rich households rather optimise their tax payments nationwide.

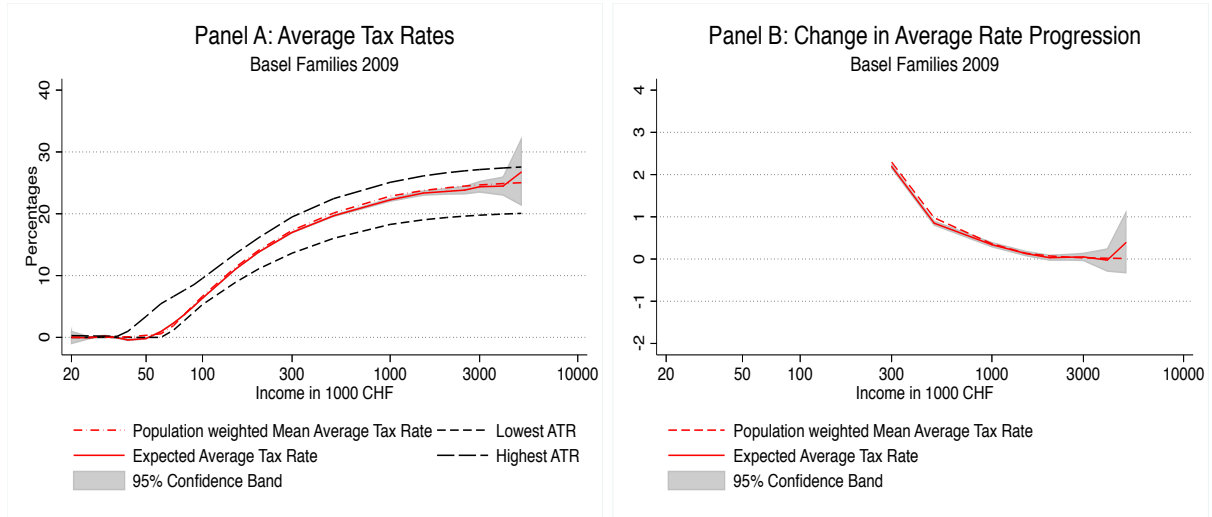
Overall, families seem to be far less sensitive to taxes than the singles. We still find a significantly lower expected average tax rates compared to the benchmark for Switzerland and the agglomeration of Zurich, the magnitude of these differences however is much smaller than for singles. For the latter, we find evidence for the presence of regressivity in the expected average tax rates for all specification while the income tax system for families with two children stays progressive throughout the entire income range. These findings suggest that single tax payers are more mobile than families, i.e. they care more about tax payment optimisation.

Figure 13: Average Tax Rate and Progressivity - Basel: Families with 2 kids



Notes: Number of observations: 17'583; Polynomial degree: 3; Kernel-weights: Epanechnikov; Half-log-bandwidth: 0.524; reported average rate progressions are multiplied by 100'000.

Figure 14: Local Average Tax Rate and Progressivity - Basel: Families with 2 kids



Notes: Number of observations: 17'583; Polynomial degree: 3; Kernel-weights: Epanechnikov; Half-log-bandwidth: 0.479; reported average rate progressions are multiplied by 100'000.

6 Conclusion

We study the progressivity of a fiscally highly decentralised country. In a two community model we are able to show that rich households pay lower average tax rates than they are expected to pay without sorting into low tax municipalities while low income households pay higher average tax rates than expected. Additionally, we show that under certain conditions rich households also encounter a lower progressivity than the local tax schedules suggest.

This paper is the first to empirically estimate the effective progressivity of the income tax system of an entire country with a decentralised income tax system. We analyse two typical household types: singles and families with two children in Switzerland. Within both groups, we find rich households being more likely to live in low tax municipalities than low income households. We compare the expected average tax rate of households with a benchmark that equals the expected average tax rate in absence of systematical income sorting. Our estimations show that the expected average tax rate is lower than the benchmark for high income households and higher for low income households. Additionally, we provide suggestive evidence that the Swiss income tax system is even getting regressive for singles with very high incomes. This is particularly the case in the agglomeration of Zurich. Thus, although all local and federal income tax schedules are progressive, the resulting country-wide income tax system is partly regressive. This strengthens the well known argument that redistribution is more effective on the federal than on the local

level. We point out that a progressive federal income tax cannot be avoided by the tax payers. Increasing its progressivity can directly increase the progressivity of the income system without losing the gains of the federal structure.

We also show that the degree of income sorting depends on the actual potential to avoid taxes. In areas with only small income tax differentials people do not seem to react strongly to these tax differentials. Furthermore, our results suggest that very high income people seem to optimise their tax payments on a national level thus they consider at least all Swiss municipalities in their location choice.

All observed effects are smaller for families with two children than for single households which leads to the conclusion that they are less sensitive to tax rate differentials than single households.

References

- Atkinson, A. B. (1970). On the measurement of inequality. *Journal of economic theory*, 2(3):244–263.
- Boadway, R. and Keen, M. (2000). Redistribution. *Handbook of income Distribution*, 1:677–789.
- Epple, D. and Sieg, H. (1999). Estimating equilibrium models of local jurisdictions. *Journal of Political Economy*, 107(4):645–681.
- Fan, J. and Gijbels, I. (1996). Local polynomial modelling and its applications. *Chapman, Hall, London*.
- Hardman, A. and Ioannides, Y. M. (2004). Neighbors’ income distribution: Economic segregation and mixing in us urban neighborhoods. *Journal of Housing Economics*, 13(4):368–382.
- Hodler, R. and Schmidheiny, K. (2006). How fiscal decentralization flattens progressive taxes. *FinanzArchiv: Public Finance Analysis*, 62(2):281–304.
- Ioannides, Y. M. (2004). Neighborhood income distributions. *Journal of Urban Economics*, 56(3):435–457.
- Jakobsson, U. (1976). On the measurement of the degree of progression. *Journal of public economics*, 5(1):161–168.
- Musgrave, R. A. and Thin, T. (1948). Income tax progression, 1929-48. *The Journal of Political Economy*, 56(6):498–514.
- Pigou, A. C. (1951). *A study in public finance*. Macmillan, London.
- Rothschild, M. and Stiglitz, J. E. (1973). Some further results on the measurement of inequality. *Journal of Economic Theory*, 6(2):188–204.
- 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.
- Silverman, B. W. (1986). *Density estimation for statistics and data analysis*, volume 26. CRC press.
- Slitor, R. E. (1948). The measurement of progressivity and built-in flexibility. *The Quarterly Journal of Economics*, 62(2):309–313.