Effective Tax Rates and Effective Progressivity in a Fiscally Decentralized Country*

Marcus Roller[§] Kurt Schmidheiny[$]

December 22, 2015

Abstract

This paper proposes measures that quantify the effective level and the effective progressivity of tax schedules in a fiscally decentralized country taking income sorting into account. Using data on the universe of Swiss taxpayers, we find that rich households effectively face significantly lower average and marginal tax rates and lower progressivity than in the benchmark case that does not consider income sorting. This is because high-income households systematically avoid high taxation by locating in low-tax jurisdictions. The results are stronger for singles than for families indicating that singles are more sensitive to spatial tax differentials than families. We find evidence that the Swiss income tax system is effectively regressive for households without children and very high incomes. Our results also hold when we account for the disutility from housing prices into which tax rates capitalize.

Keywords: Progressive Taxation, Fiscal Decentralization, Income Segregation
JEL classification: H71, H73, R23

*This paper circulated previously as "Mobility and Progressive Taxation". 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. We are also grateful to the participants of the CESifo Area Conference on Public Sector Economics, the IIPF Annual Congress, the 9th Meeting of the Urban Economics Association, the 4th Meeting of the European Urban Economics Association, the Sinergia Workshop, the Economic Geography and International Trade Workshop, the Young Swiss Economist Meeting, the SSES Annual Congress, the KOF Research Seminar, the Seminar at Centraal Planbureau, and the Conference on Regional Economics and Public Policy.

[§]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

Tax schedules can be described by various measures (see Pigou, 1951): effective average tax rates allow to compare the total tax burden, effective marginal tax rates are used in the design of optimal tax schedules (Mirrlees, 1971; Saez, 2001; Conesa and Krueger, 2006) and different progressivity measures inform about the extend to which tax schedules are redistributive (Boadway and Keen, 2000). These measures are designed for a fiscally centralized country with a single country-wide tax schedule. This paper proposes a country-wide measure of effective tax rates and effective progressivity for a fiscally decentralized country.

In a fiscally decentralized country, mobile taxpayers face a multitude of different local tax schedules. Different types of taxpayers are typically not uniformly distributed across space. High-income households, for example, may more likely live in low-tax jurisdictions than low-income households (see e.g. Schmidheiny, 2006a). Different types of households are therefore differently affected by the local tax schedules. High-income households, for example, may be less affected by high tax rates in high-tax jurisdictions simply because they are less likely to live there. We therefore propose to use a country-wide measure of effective average and marginal tax rates that takes the observed income sorting into account. We propose to use a weighted average of the tax rates across jurisdictions weighted by the actually observed spatial location pattern which depends on the taxpayer type. High tax rates in high-tax jurisdictions will therefore have less weight for the effective country-wide tax rate of high-income households if these households are less likely located in high-tax jurisdictions. We call this the *effective* rate because it is the tax rate that a specific type of taxpayers effectively faces on average in the country. The effective country-wide average tax rate can be consistently estimated by a local polynomial regression using data on the income and location of individual taxpayers.

Following Musgrave and Thin (1948), we focus on three progressivity measures: The change in average rate progression, the liability progression, and the residual income progression. These progressivity measures are also designed for a fiscally centralized country. We apply the same definitions as to our effective country-wide average and marginal tax rates. The resulting country-wide progressivity measures therefore also account for the observed spatial location pattern. We show that important properties of the centralized progressivity measures, such as the consistent classification into progressive and regressive tax systems, also hold for our measures for fiscally decentralized countries.

We apply our proposed country-wide measures of effective tax rates and effective progressivity to local income tax schedules in Switzerland, one of the most fiscally decentralized countries (Brülhart et al., 2015). We use administrative data of the universe of Swiss taxpayers (households) between 1974 and 2009. We compare effective country-wide average tax rate to the mean average tax rate that does not account for income sorting as a natural benchmark. We find that the effective country-wide average tax rates are substantially lower than the benchmark rates in particular for households with very high incomes. This reflects the fact that very high-income households are systematically clustered in low-tax municipalities in Switzerland. The effective marginal tax rates and the effective progressivity are also lower than in the benchmark. We find heterogenous results for different household types: For singles including all Swiss municipalities in 2009 we find that starting from 100,000 CHF on, the effective average tax rate is significantly lower than the benchmark. The effective country-wide average tax rate starts flattening at around 500'000 CHF and decreasing after 1M CHF. However, the progressivity measure is never significantly negative for single households. For married couples without children, our results show that the income system even becomes significantly regressive for very high incomes. For families with children the difference between the benchmark and the effective average tax rate is smaller. We also show that these results have become starker between 1974 and 2009. This finding could be explained by a higher mobility and hence more pronounced income sorting of the taxpayers.

We briefly survey the theoretical literature on income sorting in fiscally decentralized metropolitan areas in section 3. This theoretical literature shows the existence of equilibria in which high-income households tend to live in low-tax jurisdictions with high housing prices and low-income households tend to live in high-tax jurisdictions with low housing prices. The provision of public goods may be higher or lower in low-tax jurisdictions depending on the details of the models. In these models, income sorting is the consequence of systematic location choices of the households based on the local equilibrium tax rates. At the same time, the local tax rates are the result of local votes subject to local budget balance given the local income distribution in equilibrium. Hence, spatial sorting of taxpayers is theoretically both the cause and the consequence of tax differentials. In this paper, we do not need to take a stance on the causal link between local tax rates and local tax base. We simply take the observed location pattern of different households types to calculate our country-wide measures. However, the theoretical framework clarifies that the observed pattern can indeed be an equilibrium outcome of a micro-founded general equilibrium model.

The theoretical models of income sorting point to the capitalization of tax rates in housing prices. In equilibrium, the financial advantages of low local tax rates are offset by higher local housing prices and a different amount of public services. House prices capitalization is empirically well documented (see e.g.Oates 1969, Schmidheiny 2006a, Basten et al. 2014 or Morger 2013). In section 8, we develop a simple structural framework that allows to consider the effect of the disutility from house price capitalization on the effective country-wide progressivity. Applying this approach to the largest Swiss metropolitan area, we show that capitalization leaves our country-wide measures of effective progressivity practically unaltered.

The remainder of this paper is organised as follows: Section 2 discusses the standard measures of tax schedules for a fiscally centralized country and develops their respective counterparts for a fiscally decentralized country. Section 4 presents our estimation strategy. Section 5 describes the institutional background in Switzerland while section 6 describes the data. Section 7 presents the results and section 9 concludes.

2 Tax Rates and Progressivity

2.1 Level of Taxation in a Centralized Country

tax schedules are in most countries defined in tax codes and consist of list of statements relating quantities of payments required from a taxpayer to selected objective conditions (Pigou, 1951). These conditions usually include the gross income and household characteristics. The tax burden is often defined through statutory tax rates applied to taxable income, i.e. gross income minus tax deductions. A meaningful measure of the tax burden has to account for both statutory tax rates as well as tax deductions. The average tax rate, $ATR(y)$, is defined as the ratio of the tax liability $T(y)$ and gross income y. The tax liability $T(y)$ here is itself a function of gross income y:

$$ATR(y) = \frac{T(y)}{y} \tag{1}$$

While the level of taxation influences households' location choice and extensive labour supply decisions, standard literature identifies the marginal tax rate as the relevant determinant of the intensive labour supply decision (see Gruber and Saez (2002), Saez et al. (2012)). The marginal tax rate $MTR(y)$ is defined as the marginal change in the tax liability with gross income:

$$MTR(y) = \frac{\partial T(y)}{\partial y} \tag{2}$$

2.2 Progressivity in a Centralized Country

Progressivity measures are based either on the average tax rate or the marginal tax rate (Pigou, 1951): 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. In this paper we use the former concept. Musgrave and Thin (1948) discuss three progressivity measures that are based on this concept of progressivity: The change in average rate progression, the liability progression, and the residual income progression. Table 1 gives an overview and summarises their main properties.

Change in Average Rate Progression

The first measure we consider is the change in average rate progression. It is the first derivative of the average tax rate with respect to income:

$$\frac{\partial ATR(y)}{\partial y} \tag{3}$$

The change in average rate progression is a function of income y and can therefore vary across income levels. A tax schedule is called progressive (regressive) at income y if the average tax rate is increasing (decreasing) in income, i.e. the first derivative is positive (negative). If the change in average rate progression is positive for all levels of income,

Table 1: Progressivity Measures

		progressive	regressive
Change in Average Rate Progression	$\frac{\partial ATR(y)}{\partial y}$	> 0	< 0
Liability Progression	$\frac{\partial T(y)}{\partial y}\frac{y}{T(y)}$	> 1	< 1
Residual Income Progression	$\frac{\partial [y-T(y)]}{\partial y}\frac{y}{[y-T(y)]}$	< 1	> 1

the tax schedule is unambiguously redistributive, i.e. the after tax income distribution lorenz dominates the pre tax income distribution (Jakobsson, 1976).

Liability Progression
The second measure is called liability progression. It is the elasticity of the tax liability $T(y)$ with respect to income:

$$\frac{\partial T(y)}{\partial y}\frac{y}{T(y)} \tag{4}$$

This measure is also a function of income y and might vary across income levels. A tax schedule is progressive (regressive) at income y if the the liability progression is greater (smaller) than one. 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.

Residual Income Progression
The third measure is called residual income progression. It is the elasticity of the after-tax income $z(y) = y - T(y)$ with respect to income:

$$\frac{\partial \left[y - T(y)\right]}{\partial y}\frac{y}{\left[y - T(y)\right]} \tag{5}$$

Like both measures before, the residual income progression is a function of the income y. A tax schedule is progressive (regressive) at income y if the the liability progression is smaller (greater) than one.

It is straightforward to show that all three progressivity measures are consistent in the classification into progressive and regressive[1] for strictly positive levels of taxation ($T(y) > 0$) and income ($y > 0$). Thus, if the change in average rate progression indicates that a tax schedule is progressive (regressive) at income y, so do the liability progression and the residual income progression. However, the magnitude of progressivity is different across the three measures and the order of the magnitude of progressivity is not necessarily consistent.

[1]Proof: See appendix B.1

For the remainder of the paper we will mainly focus on the change in average rate progression. As we have shown this is sufficient to classify the tax schedules as progressive or regressive. Furthermore, this measure builds upon the most basic definition of progressivity and is easy to interpret. It enables us to analyze the progressivity not only at income y but also for a certain range of income.

2.3 Level of Taxation in a Decentralized Country

In a fiscally decentralized country, the total tax payments to all levels of government depend on the residential jurisdiction because local levels of governments can set different statutory tax rates and deductions. With the standard tool set we can describe these *local* tax schedules: The average tax rate in jurisdiction j is then defined as:

$$ATR_j(y) = \frac{T_j(y)}{y} \tag{6}$$

where $T_j(y)$ is the total tax payment in jurisdiction j. This measure is location-specific, i.e. conditional on the jurisdiction j. But we seek unconditional country-wide measures of average tax rates, marginal tax rates and progressivity. A first country-wide, i.e. unconditional, measure of the average tax rate is the *mean average tax rate*:

$$MATR(y) \equiv \sum_{j=1}^{J} P(j) ATR_j(y) \tag{7}$$

Where $P(j)$ is the probability of a taxpayer living in jurisdiction j. The local average tax rates are hence weighted by the relative size of the jurisdictions. The mean average tax rate does not take into account that the income distributions may differ systematically across jurisdictions. Hodler and Schmidheiny (2006) propose therefore a country-wide measure that accounts for income sorting of the taxpayers across local jurisdictions:

$$EATR(y) \equiv E\left[ATR_j(y)|y\right] = \sum_{j=1}^{J} P(j|y) ATR_j(y) \tag{8}$$

where $P(j|y)$ is the conditional probability that a household with gross income y lives in jurisdiction j. This measure is the expected average tax rate of a taxpayer with income y. Or simply the average tax rate that taxpayers with income y pay in the country. Unlike the mean average tax rate the probability $P(j|y)$ is allowed to vary with income. Thus, this measure accounts for differences in the income distributions and therefore for systematic income sorting. We call this measure the *effective country-wide average tax rate* (EATR).

In the absence of income sorting, i.e. if the local income distributions are the same, $f(y|j) = f(y) \ \forall \ j = 1, ..., J$, the effective average tax rate exactly equals the mean average tax rate[2], $EATR(y) = \sum_{j=1}^{J} P(j) f(y|j)/f(y) ATR_j(y) = \sum_{j=1}^{J} P(j) ATR_j(y) =$

[2] This even holds when the local income distributions differ but without systematic relation to the taxation across municipalities. Thus, all deviation of both measures must come from systematic income sorting with respect to taxation.

$MATR(y)$. Hence, the mean average tax rate can be used as a benchmark in the absence of income sorting.

Analogously, we can also define the *effective marginal tax rate*:

$$EMTR(y) = \frac{\partial E\left[T_j(y)|y\right]}{\partial y} = \sum_{j=1}^{J} \left\{ P(j|y) MTR_j(y) + \frac{\partial P(j|y)}{\partial y} T_j(y) \right\} \tag{9}$$

Note that the effective marginal tax rate is not just a weighted mean of the local marginal tax rates, $MTR_j(y)$, but also accounts for the change in the location choice probabilities $\partial P(j|y)/\partial y$.

Note that our proposed country-wide measure of the effective marginal tax rate is also the relevant marginal tax rate for the behavior of taxpayers in the absence of moving costs. A utility maximizing taxpayer simultaneously decides on how much to work (intensive margin) and where to live (extensive margin). He or she may decide to work much resulting in a high income or not so much resulting in a low income. Suppose that jurisdiction A is the optimal (utility maximizing) location given the low income and jurisdiction B is the optimal location given the high income. The marginal tax payment for the additional income is then the tax liability at B for the high income minus the tax liability at A for the low income. This difference divided by the income difference is the relevant marginal tax rate the taxpayer faces. This marginal tax rate is given by our proposed effective marginal tax rate if the observed location pattern given income, $P(j|y)$, is assumed the optimal location pattern.

2.4 Progressivity in a Decentralized Country

Applying the definition of the change in average rate progression to the effective average tax rate $EATR(y)$, we get a measure for the effective progressivity. We define the country-wide *change in average rate progression* in a fiscally decentralized country as:

$$\frac{\partial EATR(y)}{\partial y} = \sum_{j=1}^{J} \frac{\partial P(j|y)}{\partial y} ATR_j(y) + P(j|y) \frac{\partial ATR_j(y)}{\partial y} \tag{10}$$

Thus, an income tax schedule of a fiscally decentralized country is *progressive* (*regressive*) whenever the effective country-wide average tax rate *increases* (*decreases*). Note that the change in the tax rate comes from two sources: First, the change of the average tax rate at a given jurisdictions, $\partial ATR_j(y)/\partial y$. Second, the change in the probability of living in a certain jurisdiction, $\partial P(j|y)/\partial y$. Unlike for fiscally centralized progressivity measures this measure is not sufficient to make any statement about utility redistribution (see e.g. Atkinson (1970), Jakobsson (1976), Rothschild and Stiglitz (1973)) because the nominal income redistribution does not directly lead to real income redistribution due to possible capitalisation of tax differentials e.g. into housing prices.

The *country-wide* liability progression is:

$$\frac{\partial ET(y)}{\partial y} \frac{y}{ET(y)} = \frac{\partial EATR(y)}{\partial y} \frac{y}{EATR(y)} + 1 \tag{11}$$

where $ET(y) = E[T_j(y)|y] = \sum_{j=1}^{J} P(j|y)T_j(y)$.

The *country-wide* residual income progression is defined as follows:

$$\frac{\partial\left[y - ET(y)\right]}{\partial y}\frac{y}{y - ET(y)} = 1 - \frac{\partial EATR(y)}{\partial y}\frac{y}{1 - EATR(y)} \tag{12}$$

It is easy to see that the property of consistent classification into progressive/regressive also holds for these new progressivity measures for decentralized countries.

2.5 Statutory vs. Effective Tax Rates

The literature distinguishes between statutory and effective tax rates. While the former is clearly defined there is no unique definition of the latter. The *statutory tax rate* is the percentage rate applied to the respective statutory tax base. In the case of income taxation it is the rate applied to taxable income after all deductions. This tax rate is typically implicitly or explicitly stated in the tax code. Statutory tax rates are insufficient to compare the level of taxation across jurisdictions because the tax base might also vary systematically across jurisdictions (Oates, 1969). This issue is solved by the average tax rate: It is defined as the ratio of the income tax liability and the gross income before any deductions or benefits. The cross income as tax base measure and the tax liability are both comparable across jurisdictions, thus is the ratio. In principle, transfer payments and means-tested benefits should also be accounted for in order to measure the overall burden. We will abstract from these payments and benefits in our empirical application because they almost impossible to quantify in Switzerland.

There is no unique definition of *effective tax rates* in the context of income taxation.[3] In Mendoza et al. (1994) effective refers to an effective income that contains all sources of income and an effective tax burden that aggregates all different tax burdens from different income sources. This is related to the use by Gouveia and Strauss (1994) that estimate a effective tax function that relates the actual tax burden to economic income. Oates (1969) refers to effective taxation in context of property taxation as considering the correct tax base which relates to the concept of the average tax rate. Schaltegger et al. (2011) simply call the overall income taxation in Switzerland effective. Thus, there is no unique concept of effective taxation. However, most of them aim to measure a comparable tax burden across jurisdictions.

We define the effective tax rate in fiscally decentralized countries as the ratio of the tax liabilities and gross incomes actually observed. Thus, our definition contains a comparable tax base and tax payments as inGouveia and Strauss (1994) and Mendoza et al. (1994). In addition, however, our proposed country-wide effective average ($EATR$) and marginal tax rate ($EMTR$) takes into account the systematic spatial sorting of the population by income. If, for example, high income households systematically locate in low-tax jurisdictions then the tax burden of these low-tax jurisdictions will have more weight in the effective tax burden given income.

[3]The corporate taxation literature refers to effective tax rates as well. Effective refers to the correct specification of the tax base/corporate income measure (see e.g. Fullerton (1984), Mendoza et al. (1994))

3 Theoretical Background

Starting with Tiebout (1956) a large theoretical literature has been developed on fiscal decentralization and segregation across jurisdictions. Comprehensive reviews of this literature are provided in Ross and Yinger (1999), Epple and Nechyba (2004) and most recently in Brülhart et al. (2015). Typically, these models consist of a metropolitan area that is divided into several jurisdictions offering different taxation and public good bundles[4] to mobile households that differ in income. First, households choose their residential jurisdiction and thereafter vote on the the taxation and the public good provision. In most of these models local taxation is a *property* tax. Early models (see Westhoff (1977), Epple et al. (1984) and Epple and Romer (1991)) provide only equilibria with perfect segregation, i.e. all jurisdictions are exclusively populated with people in a certain income range. Epple and Platt (1998) and Epple and Sieg (1999) add heterogenous household preferences resulting in asymmetric equilibria with imperfect income segregation across jurisdictions. Calabrese et al. (2007) show that adding zoning, i.e. minimum housing size requirements, to the model magnifies the income sorting. Calabrese et al. (2012) perform a welfare analysis and show that the sorting reduces welfare on average and for most of the population.

Only few papers model local *income* taxation. Goodspeed (1989) finds that high income households sort into jurisdictions with low income tax rates and high housing prices. This result holds in Schmidheiny (2006a), Schmidheiny (2006b) and Hodler and Schmidheiny (2006). Schmidheiny (2006a) introduces progressive income taxation into the model and identifies it as a source of income segregation because the marginal gain of living in low tax jurisdictions is higher for high income households than for low income households. Gravel and Oddou (2014) show that the clear segregation pattern highly depends on assumptions made about the demand function for the public good as well as on the separability of the households' preferences.

The common feature of above theoretical literature is the existence of segregated equilibria in which high-income households tend to live in low-tax jurisdictions with high housing prices and low-income households tend to live in high-tax jurisdictions with low housing prices. The provision of public goods may be higher or lower in low-tax jurisdictions depending on the details of the models.

Hodler and Schmidheiny (2006) theoretically show that income segregation lowers the country-wide effective progressivity if local jurisdictions can set the *level* of tax rates but not the entire tax schedule. We generalize this result in Appendix A by allowing that local jurisdictions can set different tax schedules. However, this more general results still relies on the assumption that the order of tax rates across jurisdictions is the same for all income levels, i.e. a low-tax jurisdiction offers low tax rates compared to other jurisdictions for all levels of income. In reality, local tax schedules may be crossing, i.e. some jurisdictions may offer relatively low tax rates for high-income households but relatively high rates

[4]The public good is usually a publicly provided private good.

for low-income households. It remains therefore an empirical question how the observed local tax schedules and the observed income sorting affects the country-wide effective progressivity.

4 Econometric Method

This section describes the estimation of the country wide effective average tax rate function and its progressivity. The effective average tax rate is defined as:

$$EATR(y) = E\left[ATR_j(y)|y\right] = \sum_{j=1}^{J} P(j|y)ATR_j(y)$$

Because the gross income y is a continuous variable, the probability of observing a taxpayer with exact income y is zero. Consequently, we cannot directly observe the expected average tax rate. Since the local average tax rates $ATR_j(y)$ are observed, Hodler and Schmidheiny (2006) follow the strategy of estimating the probability $P(j|y)$ that a taxpayer with income y lives in jurisdiction j using Bayes' rule:

$$P(j|y) = \frac{P(j)f(y|j)}{f(y)} \tag{13}$$

where $P(j)$ is the probability that a taxpayer of any income lives in jurisdiction j, $f(y|j)$ is the gross income density in jurisdiction j and $f(y)$ is the country-wide gross income density. Hodler and Schmidheiny (2006) estimate $f(y|j)$ and $f(y)$ from data on the municipal income distribution assuming that income is log-normally distributed both at the municipal and the country level. This approach can easily be implemented as it uses publicly available aggregate data only. However, it rests on highly parametric assumptions about the income distributions.

In this paper, we suggest to directly estimate the effective average tax rate $EATR(y)$ without explicitly estimation of$P(j|y)$ by a local polynomial regression. In the neighbourhood of $\ln(y_0)$ we can approximate the country-wide average tax function with a taylor polynomial of order p:

$$\begin{aligned} ATR(\ln(y)) \approx ATR(\ln(y_0)) &+ \frac{\partial ATR(\ln(y_0))}{\partial \ln(y)}(\ln(y) - \ln(y_0)) \\ &+ \frac{\frac{\partial^2 ATR(\ln(y_0))}{\partial \ln(y)^2}}{2!}(\ln(y) - \ln(y_0))^2 \\ &+ \ldots \\ &+ \frac{\frac{\partial^p ATR(\ln(y_0))}{\partial \ln(y)^p}}{p!}(\ln(y) - \ln(y_0))^p \end{aligned} \tag{14}$$

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

$$\min_{\beta} \sum_{i}^{N} \left\{ ATR_i \sum_{j=0}^{p} \beta_j (\ln(y_i) - \ln(y_0))^j \right\} K_h(\ln(y_i) - \ln(y_0)) \tag{15}$$

where $K_h(.)$ are Epanechnikov kernel weights with bandwidth h. Using individual taxpayers' data on gross income[5] y_i and average tax rate ATR_i. Choosing polynomials of degree $p > 0$ allows us to simultaneously estimate the effective average tax rate and the change in average rate progression at a given level of gross income y_0:

$$\widehat{ATR(y_0)} = \hat{\beta}_0 \tag{16}$$

$$\frac{\partial \widehat{ATR_j(y)}}{\partial y} = \frac{\hat{\beta}_1}{y_0} \tag{17}$$

This method results in consistent[6] estimates of the effective average tax rate, $\hat{\beta}_0$, and the change in average rate progression, $\hat{\beta}_1/y_0$. This non-parametric approach requires a country-wide sample of individual data with information on the gross income and the tax schedule in the residential jurisdiction. It also requires sufficient observations around each income level y_0 including top incomes in order to get precise estimates over the entire income range. In fact, we will use the universe of taxpayers in Switzerland.

The benchmark, the mean average tax rate $MATR(y)$, can easily be calculated as

$$\widehat{MATR}(y) = \sum_{j=1}^{J} \widehat{P(j)} ATR_j(y) = \sum_{j=1}^{J} \frac{N_j}{\sum_{j=1}^{J} N_j} ATR_j(y) \tag{18}$$

where the probability $P(j)$ of living in jurisdiction j is estimated using the observed population size of the single jurisdictions N_j. As discussed above in section 3, the mean average tax rate is equal to the effective average tax rate in the absence of income sorting. Thus, the difference between the two measures must be due to systematic differences in the income distributions. For the effective marginal tax rate one follows exactly the same strategy but using the tax liability T_i as dependent variable in eq. (15). The estimator for the effective marginal tax rate is then $\hat{\beta}_1/y_0$. The liability progression and the residual income progression can be estimated in a similar way.

5 Institutional Background

Switzerland taxes income on three different levels: federal, cantonal and municipal. This highly decentralized income tax system with its large variation in the local tax rates makes Switzerland especially interesting for an empirical application of our set of measures.

The federal state charges the gross income y with tax $T_f(y)$ independent of the residential location within Switzerland. The federal tax amount $T_f(y)$ is a function of gross income and the household type, i.e. deductions and statutory tax rates depend on household characteristics: Married couples are jointly taxed and treated as a one taxpayer. Tax deductions depend e.g. on the marital status and number of children. There are two different statutory federal tax rates: one for married couples and one for unmarried individuals.

[5]The logarithmic transformation allows us to work with a fixed bandwidth and symmetric kernels for the transformed variable and therefore growing bandwidths for the original gross income.

[6]The local polynomial regression estimator is generally biased but consistent. Including higher order polynomials reduces this bias especially at the boundaries (see e.g. Fan and Gijbels (1996)).

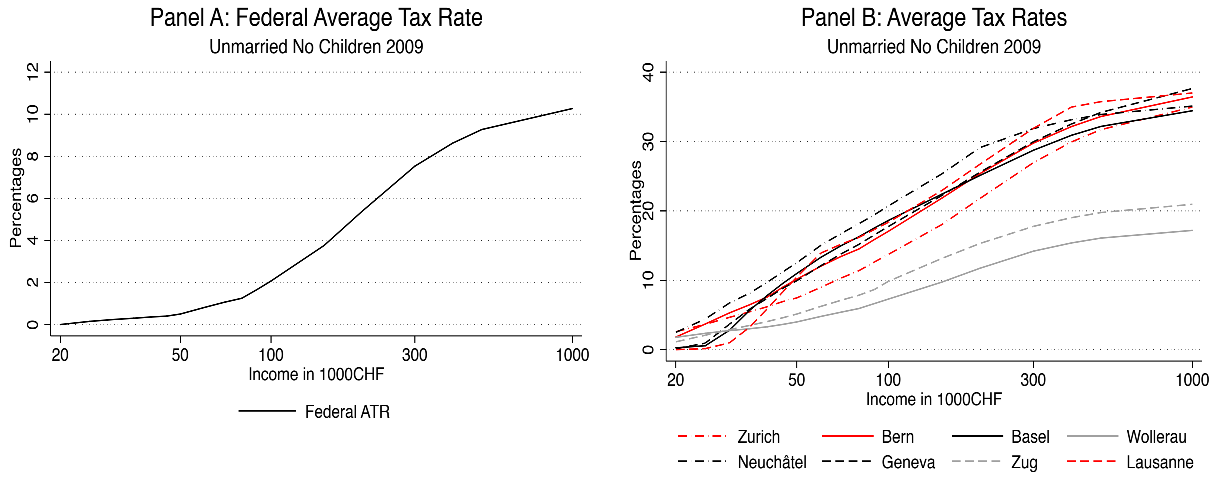

Figure 1: Panel A: Federal average tax rates for unmarried taxpayers in Switzerland in 2009. Panel B: Average tax rates (including federal, cantonal and municipal tax) for unmarried taxpayers in 8 Swiss municipalities in 2009.

The left panel in Figure 1 shows the federal average tax rate, $ATR_f(y) = T_f(y)/y$, for unmarried individuals in 2009. This average tax rate is the *effective* rate in the sense that it includes household-type specific deductions, is defined over a comparable tax base and is unconditional on the residential jurisdiction. However, it does not consider means-tested benefits.[7] The federal average tax rate increases steeply with gross income and is therefore highly progressive for all levels of income.

The 26 ccantons set their own income tax schedules $T_c(y)$ where c denotes the residential canton which is uniquely identified by the residential municipality j. The cantonal tax schedules differ in terms of both deductibles and statutory tax rates. The cantonal tax schedules are typically less progressive than the federal one. In some cantons, Obwalden and Uri, it is a flat-rate tax.

The 2,624 municipalities also tax income. The municipal income tax is a multiple of the cantonal tax liability: $M_jT_c(y)$, i.e. within cantons, municipal tax rates only vary with the tax multiplier M_j.[8]

The total income tax liability of an individual with gross income y living in municipality j is the sum of the income taxes collected by the three governmental layers:

$$T_j(y) = (1 + M_j)T_c(y) + T_f(y). \tag{19}$$

The average tax rate in municipality j at gross income y is the ratio of the income tax

[7] see section 2.5 for discussion

[8] There is also a local church tax defined as a multiple of the cantonal tax amount. Church taxes are collected by the government for individuals who are members of either the catholic, protestant or jewish community. The geographic borders of the municipalities and the religious communities are typically but not always congruent. We add the tax multiplier of the local majority church to the municipal tax multiplier in the empirical study. The church tax is small and does not drive our results.

liability $T_j(y)$ and the gross income y:

$$ATR_j(y) = \frac{T_j(y)}{y} \tag{20}$$

The federal and all 26 cantonal tax schedules are progressive, i.e. $\partial ATR_f(y)/\partial y > 0$ and $\partial ATR_c(y)/\partial y > 0$. Hence, the combined local tax schedules are also progressive in each municipality, i.e. $\partial ATR_j(y)/\partial y > 0$.

This tax system results in 2,624 tax schedules in Switzerland. The right panel in Figure1 shows the combined average tax rate, $ATR_j(y) = T_j(y)/y$, for unmarried individuals in 8 selected municipalities in the year 2009. It illustrates the large variation in these tax schedules across municipalities, both in the level and the progressivity. Additionally, it reveals that tax schedules in municipalities located in different cantons can intersect. Thus, a municipality might be a high tax municipality for low incomes and a low tax municipality for high incomes. Section 6 provides more descriptive results about the spatial distribution of the tax schedules.

6 Data

As described in the methodology section, we need data on taxpayers' gross income and average tax rates. We use administrative data on the universe of all Swiss taxpayers and the universe of Swiss municipalities. We focus on three household types: Unmarried taxpayers without children, married taxpayers without children and married taxpayers with children. The following section describes the data set and its construction in detail.

6.1 Individual Income Data

We use confidential administrative data on the universe of Swiss taxpayers from the Swiss Federal Tax Administration (ESTV) for the years 1974 to 2009. Note that married couples are treated as one taxpayer in Switzerland. We observe their federal taxable income, their residential location (municipality) and their tax relevant household characteristics (marital status and number of children that qualify for tax deductions). Federal taxable income is reported income from wages, self-employment, pension benefits, interest, dividends, etc. minus social security contributions and household-type specific deductions for the federal income tax.[9] We approximate gross income y_i, i.e. income before social security contributions and tax deductions, for each taxpayer from taxable income and household characteristics applying the federal tax code (see Appendix C for details).

In 2009, we observe 1,649,167 single households, 427,377 married couples without children and 303,283 married couples with children. Table 2 reports descriptive statistics for the gross income distribution of these three different types of taxpayers. Unmarried taxpayers have the lowest mean income of the three types and married couples with children the highest. This also holds for all percentiles up to the 99% percentile. The

[9]Capital gains are not taxed in Switzerland and therefore not reported as income in our data.

Table 2: Descriptive Statistics of Individual Gross Income in 2009.

	Sample	Unmarried No Children	Married Couple No Children	Married Couple 2 Children
Nobs.	2,379,827	1,649,167	427,377	303,283
Mean	75,509	56,173	114,498	125,730
SD	155,670	93,820	268,202	185,400
Percentiles				
50%	62,198	51,382	93,179	101,531
90%	131,649	94,870	177,590	194,793
95%	169,246	117,394	230,891	256,930
99%	319,446	199,575	463,273	506,930
99.9%	961,823	556,622	1622,840	1434,418

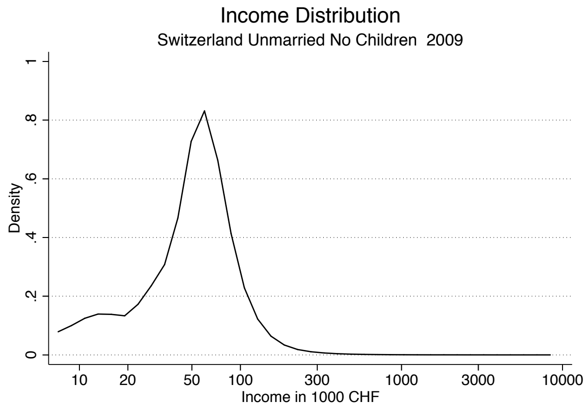

Figure 2: Kernel density estimate of the gross income distribution of unmarried taxpayers without children in Switzerland in 2009.

income distribution for the married taxpayers without children seems to have a fat right tail compared to the other household types.

Figure 2 shows the a kernel density estimate of the income distribution of unmarried taxpayers without children in Switzerland. Note that the income scale is logarithmic since the distribution is extremely skewed to the right. The mode is close the median and mean. The density estimates for the other household types look similar but shifted to the right.

6.2 Tax Burden

The Swiss Federal Tax Administration (ESTV) annually publishes the tax burden for a sample of Swiss municipalities.[10] Reported is the combined cantonal, municipal, and

[10]Swiss Federal Tax Administration, Steuerbelastung in der Schweiz, Neuchâtel: Swiss Federal Statistical Office. Data for the years before 1996 is only available as scanned images. This data was digitized by Kurt Schmidheiny and his team within the Sinergia Project 130648 founded by the Swiss National

Table 3: Descriptive Statistics: Average Tax Rates of Swiss Municipalities at Respective Income Percentiles

	Percentile	Nobs.	Mean	Std. Dev.	Min	Max
	50%	2624	9.95	1.49	4.06	13.7
Unmarried	80%	2624	13.93	1.81	5.7	18.81
No Children	90%	2624	16.19	1.94	6.93	21.42
	99%	2624	24.5	2.59	11.66	30.82
	50%	2624	10.94	1.79	4.79	15.37
Married	80%	2624	15.04	1.83	6.94	19.76
No Children	90%	2624	18.49	1.99	9.03	23.69
	99%	2624	29.46	2.73	15.28	34.97
	50%	2624	8.25	1.74	3.23	11.89
Married	80%	2624	12.95	1.79	5.93	17.61
2 Children	90%	2624	17.05	1.91	8.42	22.29
	99%	2624	28.89	2.66	15.01	34.55

church tax burden as well as the federal tax burden. Tax burdens are reported for different household types (unmarried taxpayer without children, married couple without children, married couple with two underage children, retired married couple) and for different levels of gross income.[11]

The reported tax burden in Swiss Francs (CHF) is calculated using the federal and cantonal tax codes as well as the municipal multipliers. The calculation first subtracts social security contributions as well as federal and canton-specific tax deductions for e.g. underage children and applies the respective statutory tax rates for unmarried and married taxpayers.

For the year 2009, the ESTV reports the tax burden for 813 municipalities. This data is extended to all 2,624 Swiss municipalities by collecting tax multipliers for all municipalities (see Parchet (2014) for details). Table 3 reports mean and standard deviation for selected household types (unmarried, married two children) and incomes (median, top-20%, top-10%, top-1%). Especially the spreads between the highest and lowest tax rates indicate a large variation in the tax rates across municipalities in Switzerland. For an unmarried taxpayer with median incomes this spread is already more than 9 percentage points. The tax burden in the highest tax municipality is three times the one of the lowest tax municipality. For higher incomes the spread increases even more in absolute but not in relative terms for all household types. 10pt10pt

Science Foundation.

[11]For example in 2009: CHF 20,000; 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.

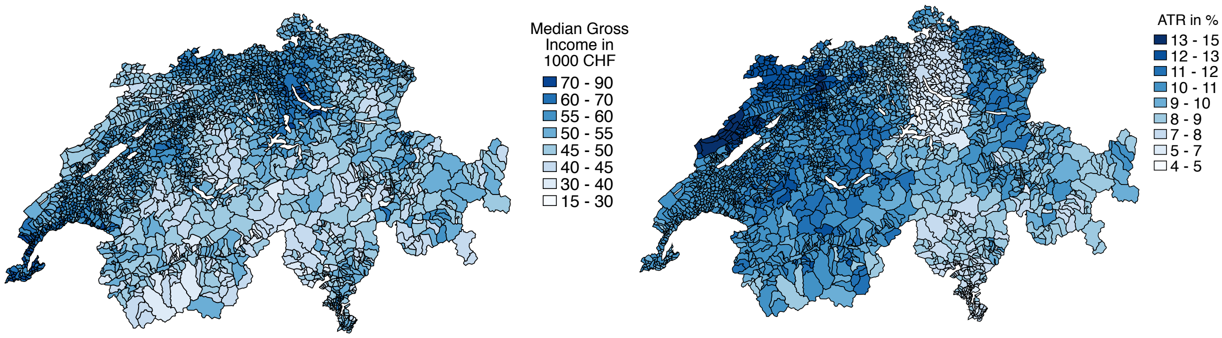

Figure 3: Left map: median income of unmarried taxpayers across Swiss municipalities. Right map: Average tax rate of an unmarried taxpayers without children and median income.

6.3 Individual Tax Rates

For our analysis we need the individual average tax rates and not only those at the grid points provided by the ESTV. The individual taxpayers' data does only contain the tax information needed for the federal income tax, i.e. it does not contain any information about the cantonal or municipal taxes. Thus, we have to calculate them by using the data on the tax burden in their residential municipalities. We calculate the individual tax burden, $T_i(y_i)$, for each taxpayer using the approximated gross income, y_i, and the data on the combined tax burden for different household types and income levels in all Swiss municipalities. We interpolate the tax burden for income levels which are not reported in our tax burden data (see Appendix C for details). The average tax rate for each taxpayer in our individual data is simply $ATR_i = T_i(y_i)/y_i$.

6.4 Income Sorting

The left map in Figure3 depicts the spatial distribution of the incomes of unmarried taxpayers in Switzerland. The median income differs substantially across municipalities. This also holds for smaller geographic regions. The right map depicts the average tax rates in the municipalities that a taxpayer with median income faces. Also the tax rates vary a lot across Switzerland and smaller geographic regions. If we compare both maps it seems that areas with low tax rates have higher median incomes. This could be a first evidence for systematic income sorting. The evidence becomes even stronger if we compare the income distribution of those municipalities with the highest marginal tax rate for top incomes with those for the lowest. Figure 4 plots these income distributions. There is clearly more mass at higher incomes in the low tax municipalities. Thus, income sorting seems to play a role in the case of Switzerland.

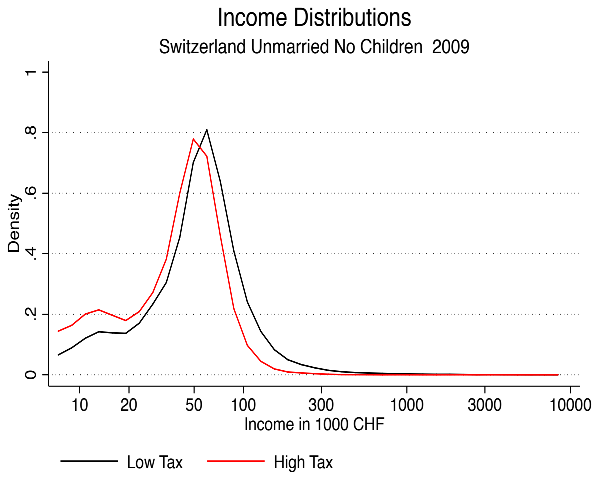

Figure 4: Kernel density estimates of income distributions in the lowest 5% marginal tax rate municipalities and the highest 5%. Unmarried taxpayers without children.

7 Results

We first report country-wide results for the whole of Switzerland for the year 2009 in Subsection 7.1. We then restrict our analysis to a single metropolitan area in Subsection 7.2. The historic development from 1974 to 2009 is shown in Subsection 7.3.

7.1 Switzerland

The left panel of Figure 5 presents the results from the non-parametric estimation of the effective average tax rate for unmarried taxpayers in 2009. The solid red line is the country-wide effective average tax rate $EATR(y)$. The grey band is the pointwise 95% confidence interval. The black dashed lines represent the average tax rate in the municipality with the highest and lowest tax burden at the given gross income, respectively. The red dash-dotted line is our benchmark, the mean average tax rate $MATR(y)$. It is monotonically increasing with gross income. This reflects the fact that the combined federal, cantonal and municipal tax schedules are monotonically increasing in each Swiss municipality as discussed in Section 5.

Starting from 100,000 CHF (approximately 96,000 USD end of 2009), the *effective* average tax rate is significantly lower than the mean average tax rate. The gap is constantly increasing up to about ten percentage points for an income of 10M CHF. This difference arises from income sorting. Thus, taxpayers with incomes above 100,000 CHF live systematically in municipalities with comparably low tax rates. Starting from 1M CHF, the effective average tax rate is even decreasing. Also our estimates for the *effective* country-wide change in average rate progression shown in the right panel of Figure 5 are lower for high incomes than the benchmark and negative for incomes above 1M CHF which would indicate that the tax system in this range is even regressive. However, all negative estimates are not significantly different from zero. For incomes above 3M CHF,

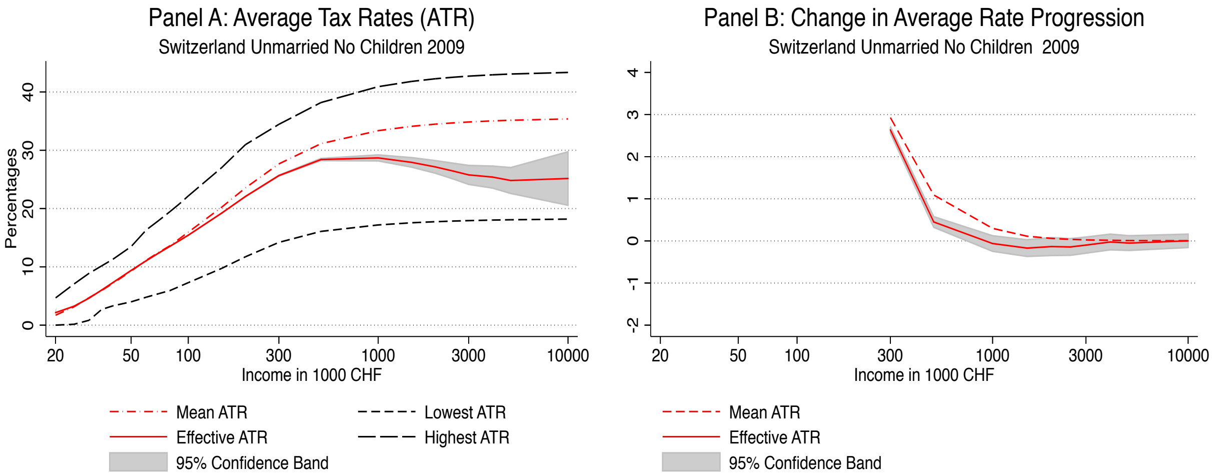

Figure 5: Country-wide average tax rate and progressivity of unmarried taxpayers. 1,649,167 observations, local polynomial regression (order 3) with Epanechnikov kernel and half-log-bandwidth 0.343. Change in average rate progression is multiplied by 100,000.

Table 4: Discrete Differences between the effective average tax rate with income y and with an income of 500,000 CHF for three household types in Switzerland in 2009

Income	Unmarried No Children	Married No Children	Married 2 Children
Effective average tax rate EATR			
500,000	28.41	26.96	26.93
Difference of EATR compared to 500,000			
1,000,000	0.279	-	-
	(0.277)	-	-
2,000'000	-1.266**	2.891***	4.497***
	(0.559)	(0.229)	(0.277)
3'000,000	-2.638***	2.190***	4.280***
	(0.836)	(0.365)	(0.495)
5,000,000	-3.601***	1.822***	3.796***
	(1.138)	(0.660)	(0.824)
10,000,000	-3.253	-0.819	3.687
	(2.339)	(1.340)	(1.75)

Standard Errors in Parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

there is not even a significant difference between the effective average tax rate and the benchmark anymore. This is due to the fact that marginal tax rates are constant in all cantons as of an income of around 500,000 CHF which gives a progressivity of zero in the limit by construction.

The estimated *effective* progressivity shown in the right panel is a *local* estimate in the sense that it is the *marginal* change of the effective average tax rate at a given level of

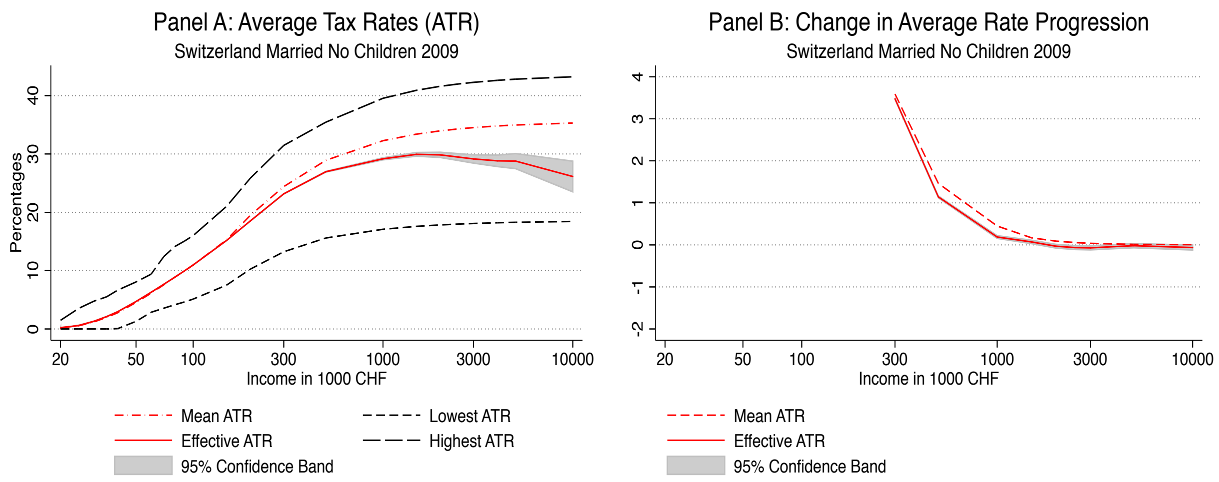

Figure 6: Country-wide average tax rate and progressivity of married couples without children. 427,377 observations, local polynomial regression (order 3) with Epanechnikov kernel and half-log-bandwidth 0.571. Change in average rate progression is multiplied by 100,000.

gross income. The term progressivity can in principle also be applied to discrete changes of the effective average tax rate. The left panel of Figure 5 shows that the confidence band of the effective average tax rates falls after an income of about 1M CHF indicating that the effective country-wide tax schedule is in fact *regressive* for large changes in income. Table 4 reports statistical tests for discrete changes in income form 500,000 CHF to incomes above 1M CHF. Moving from e.g. 500,000 CHF gross income to 5M CHF significantly ($p < 0.05$) reduces the effective average tax rate by 3.6 percentage points from 28.4 to 24.8 percent for unmarried taxpayers. The results for the married taxpayers read as follows: At all income levels from 500,000 CHF households pay significantly higher tax rates up to an income of 5M CHF. Households with an income of 10M CHF do not face significantly higher tax rates than households with an income of 500,000 CHF.

The results for married couples without children are similar as Figure 6 shows. The deviation of the effective average tax rate from the mean average tax rate starts at about 150,000 CHF and also increases up to ten percentage points. The progressivity for high incomes is also significantly lower. In the range of 2M up to 3M CHF it gets even slightly but significantly negative.

For families with two children the results differ from the previous types. Figure 7 depicts these results. We immediately see that the difference between the effective tax rate and the benchmark as well as the difference between the effective progressivity and its benchmark are smaller than in the previous cases. The effective average tax rate is only significantly lower than the mean average tax rate for incomes above 200,000 CHF and the difference only reaches six percentage points for households with an annual income of 10M CHF. The progressivity is also significantly lower but the difference of the effective progressivity to its benchmark is smaller than for unmarried taxpayers and married couples without children. These results are not surprising since families with children are likely to differ from the other types in there preferences regarding the location

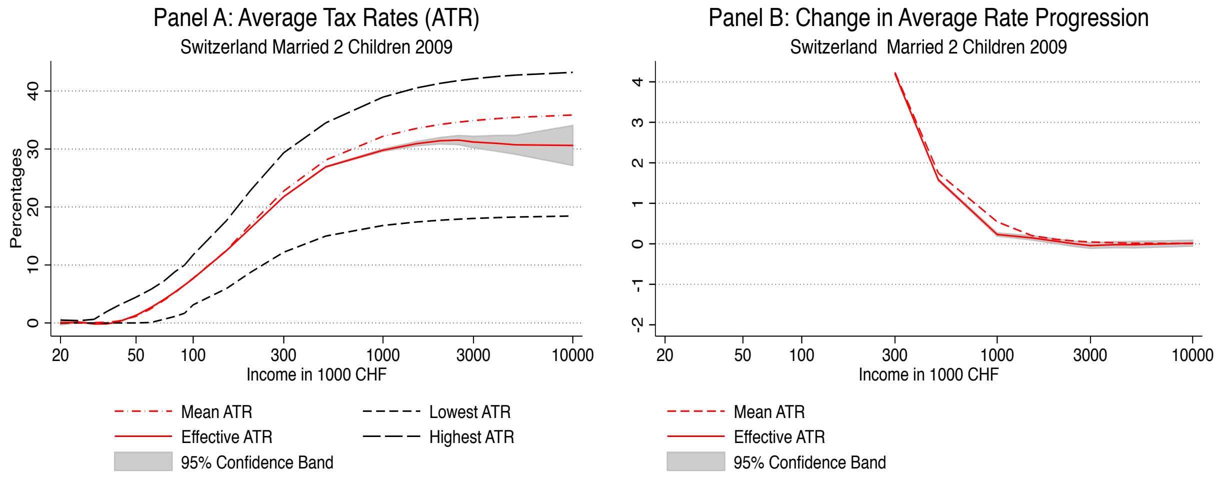

Figure 7: Country-wide average tax rate and progressivity of married couples with two children. 303,283 observations, local polynomial regression (order 3) with Epanechnikov kernel and half-log-bandwidth 0.516. Change in average rate progression is multiplied by 100,000.

choice. Unlike the other household types they care e.g. about the quality of schooling (Epple et al. (2012)). In Switzerland, elementary schooling is indeed exclusive and only accessible for children living within a school district. Black (1999) shows that parents are indeed willing to pay for higher schooling quality. Also Bayer et al. (2007) find the same result but of minor size. They point out that this difference is explained by neighbourhood quality effects. Following Basten et al. (2014) this is likely to be the case in Switzerland as well because schools are to a large extent financed and regulated[12] by the cantons. Thus, we would not expect that school provision drives our results but rather neighbourhood composition effects. Thus, families might care a lot about their peers and their children's peers in school which effects ex post school quality. Bayer and McMillan (2012) also finds that families care more about the availability of large houses compared to households without children. Overall, the differences in our results between households with and without children are more likely to be driven by housing availability and peer effects than differences in the quality of supplied schooling.

We also estimate the effective marginal tax rate, $EMTR(y)$, using a non-parametric regression. Figure 8 presents the results of the estimations for unmarried taxpayers (left panel) and for married couples without children (right panel). For unmarried taxpayers, the effective marginal tax rate is lower than the mean marginal tax rate (benchmark) for incomes above 80,000 CHF. For married couples the deviations starts not before 100,000 CHF. While for both household types the difference between the effective and the mean marginal tax rate is comparably small for incomes below 300,000 CHF it starts increasing a lot for higher incomes. This increase is driven by a decrease in the effective marginal tax rate while the mean marginal tax rates for high are almost constant in both cases. The latter is due to the fact that also the statutory marginal tax rates are constant for most of the cantons in this income range. Similar results but of smaller size also hold for

[12]e.g. class sizes.

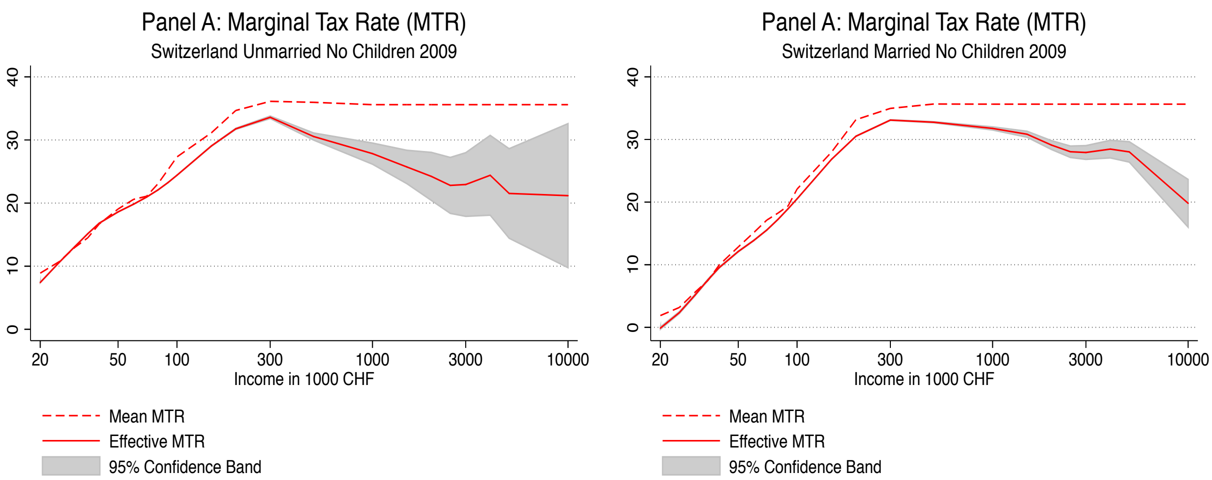

Figure 8: Country-wide marginal tax rate of unmarried taxpayers and married couples without children. Unmarried taxpayers: 1,649,167 observations, local polynomial regression (order 3) with Epanechnikov kernel and half-log-bandwidth 0.343. Married couples without children: 427,377 observations, local polynomial regression (order 3) with Epanechnikov kernel and half-log-bandwidth 0.571.

families with two children.

For all three household types, the effective average tax rate, the effective marginal tax rate and the effective progressivity are lower than the benchmark that does not account for income sorting. We even find that for unmarried taxpayers, the income tax system gets regressive for a certain range of high incomes.

The results so far include the highly progressive federal income tax which is constant across the country. In order to asses the role of the federal tax, we repeat the estimation

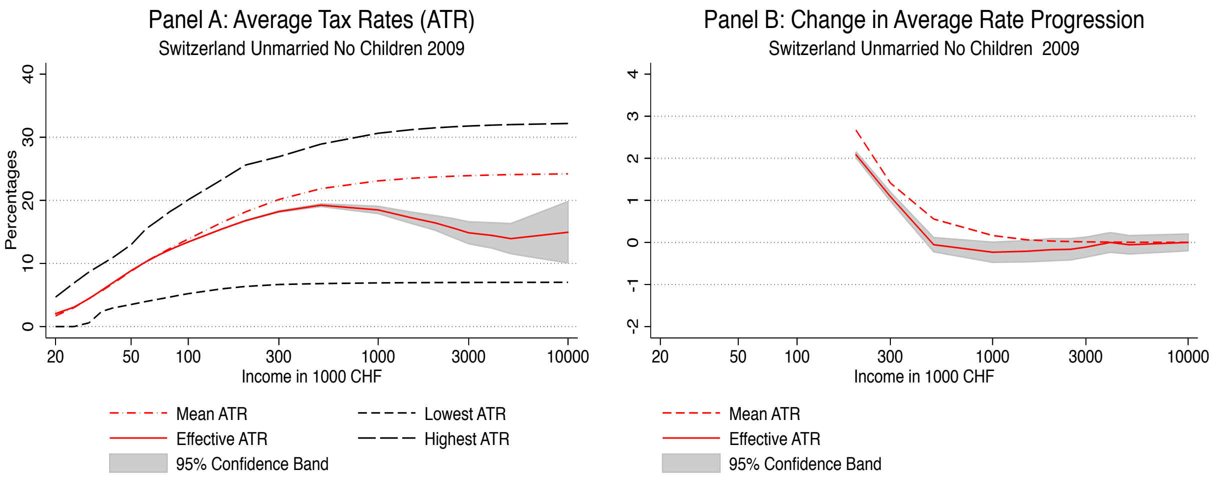

Figure 9: Country-wide average tax rate and progressivity of sub-central (cantonal and municipal) income taxes for unmarried taxpayers. 1,649,167 observations, local polynomial regression (order 3) with Epanechnikov kernel and half-log-bandwidth 0.299. Change in average rate progression is multiplied by 100,000.

considering only sub-central, i.e. cantonal and municipal taxes. Figure 9 shows these results. Just considering local tax rate differentials strengthens our previous findings: the effective average tax rates stop rising after an income of 500,000 CHF and fall after about 1M CHF.

7.2 Zurich Metropolitan Area

Theoretical models of spatial income sorting from local tax rate differentials typically model a single city or metropolitan area (see Section 3). In such a setting it is reasonable to assume that households earn some given income at some workplace (e.g. the central business district) and choose their residential location within the city given their income. However, this assumption is not reasonable when studying a whole country, because choosing a metropolitan area or region within the country will likely affect the workplace and hence income. In order to get our empirical results closer to the theoretical motivation, we repeat the analysis considering only municipalities within one metropolitan area.[13] We choose the Zurich metropolitan area, the economically most important area in Switzerland.[14]

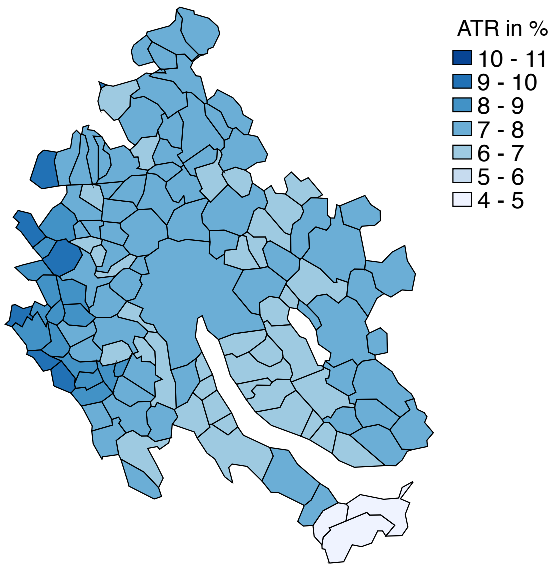

Figure 10: Average tax rates for an unmarried taxpayer with median income - Zurich.

Even within this smaller area, we observe large tax differential as Figure 10 indicates. Because the tax differentials are comparable to those of Switzerland and the mobility within a smaller area should be higher than across Switzerland, we expect the effects within this metropolitan are to be lager than for the entire country.

Figure 11 presents the results of the estimations for single households. We observe an increasing gap between the average tax rates between the lowest and highest possible

[13]Note that our empirical results for the entire country are valid but only harder to explain theoretically.

[14]The metropolitan area is defined by the Swiss Federal Statistical Office. Note that is not congruent with the canton of Zurich.

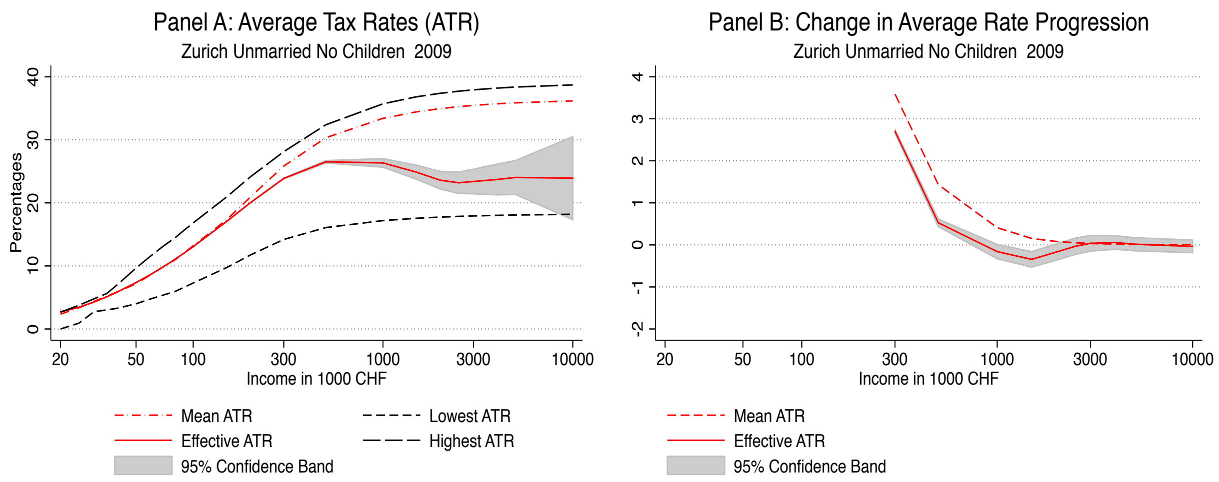

Figure 11: Average tax rate and progressivity of unmarried taxpayers - Zurich. 286,831 observations, local polynomial regression (order 3) with Epanechnikov kernel and half-log-bandwidth 0.459. Change in average rate progression is multiplied by 100,000.

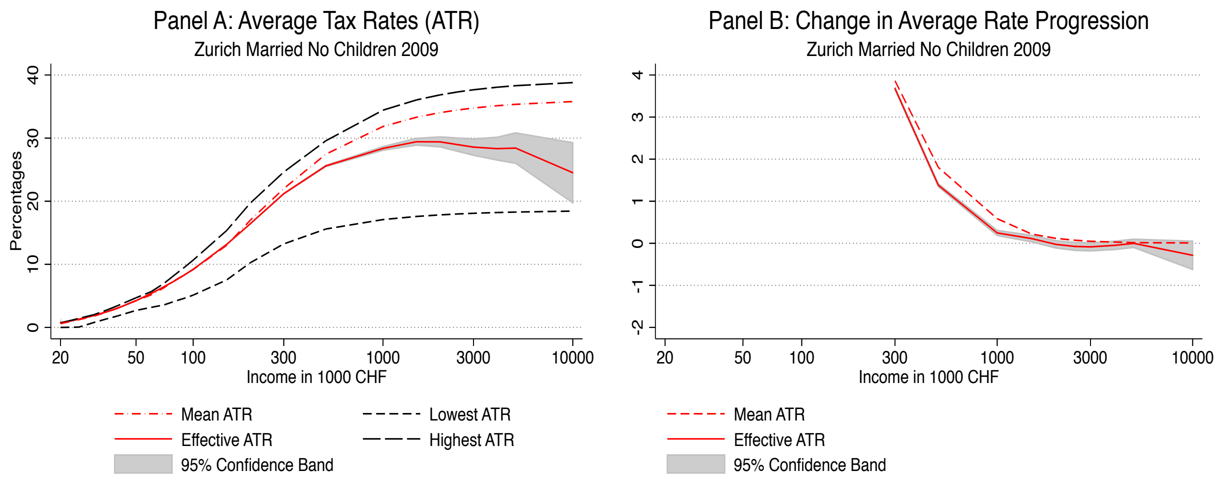

Figure 12: Average tax rate and progressivity of married couples without children - Zurich. 66,721 observations, local polynomial regression (order 3) with Epanechnikov kernel and half-log-bandwidth 0.534. Change in average rate progression is multiplied by 100,000.

average tax rates up to 20 percentage points for an income of 10M CHF which generates large possible gains of moving to the lower tax municipalities. The 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 effective 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 3M CHF. As of an income of 1M CHF the effective average tax rate starts declining up to an income of 3M CHF. The progressivity is also significantly lower and in this case even negative for incomes between 1M CHF and 2M CHF as panel B of Figure 11 indicates.

For married couples the picture looks similar. However, there is no range of income where our estimates of the progressivity are significantly negative (see Figure 12). For

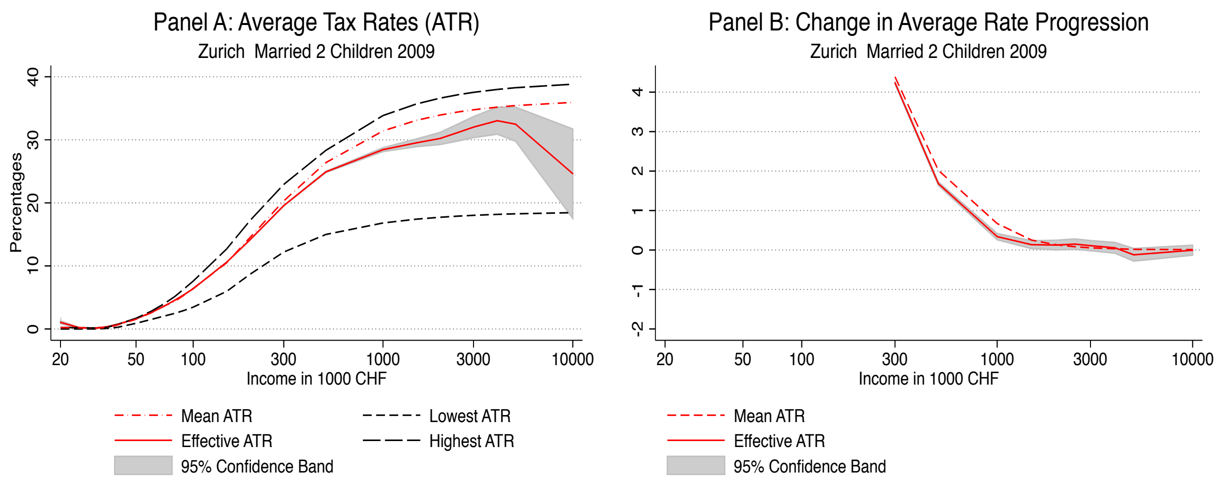

Figure 13: Average tax rate and progressivity of married couples with two children - Zurich. 41,857 observations, local polynomial regression (order 3) with Epanechnikov kernel and half-log-bandwidth 0.461. Change in average rate progression is multiplied by 100,000.

families with children the the difference between the effective and the mean average tax rate is again smaller. This is consistent with our explanation from above that families with children differ substantially in their location preferences from the other household types.

Hence, we have shown that results also hold for smaller functional urban areas and are not driven by cultural or structural differences between different regions of Switzerland. For singles the difference between the effective average tax rate and the mean average tax rate is even higher than in Switzerland.

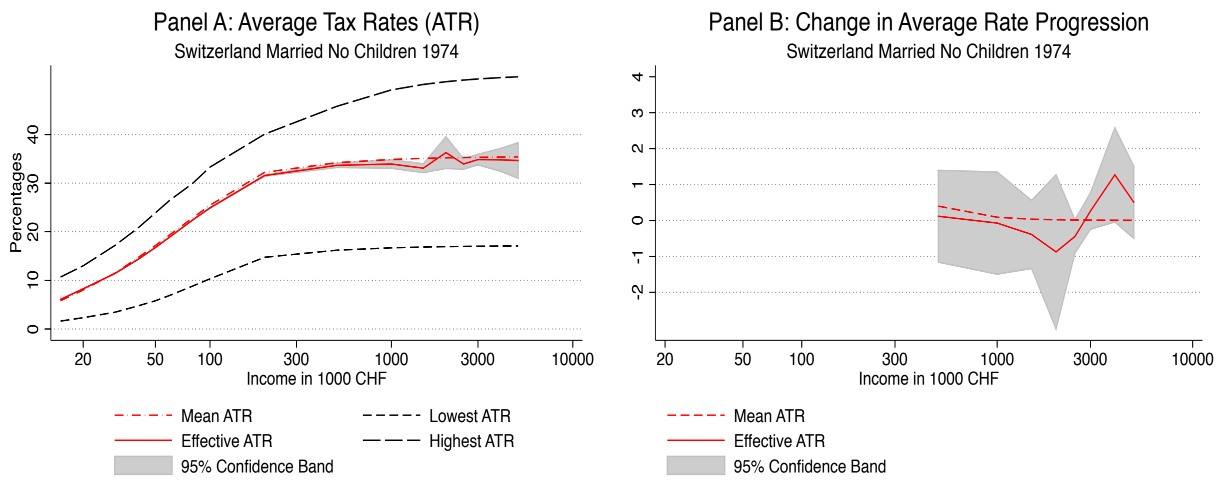

Figure 14: Country-wide average tax rate and progressivity of married couples without children in 1974. 377,745 observations, local polynomial regression (order 1) with Epanechnikov kernel and half-log-bandwidth 0.057. Change in average rate progression is multiplied by 100,000.

7.3 Historic Development

All observed results depend on two factors: First people need to have the possibility to avoid taxes, this means there must be a spread in the tax rates between the municipalities. Second people must be mobile, i.e. their must be low frictions in choosing the residential municipality. These factors probably changed over time. We therefore repeated our analysis for the years 1974 and 1986. Figure 14 presents our estimates for the year 1974 while figure 15 presents the results for 1986. We clearly see that in both years married couples without faced a large spread in tax rates (even higher in 1974 than today). Yet, in 1974 we do not find strong evidence for systematic income sorting. The effective average tax rate and the progressivity almost equal their benchmarks. This changes

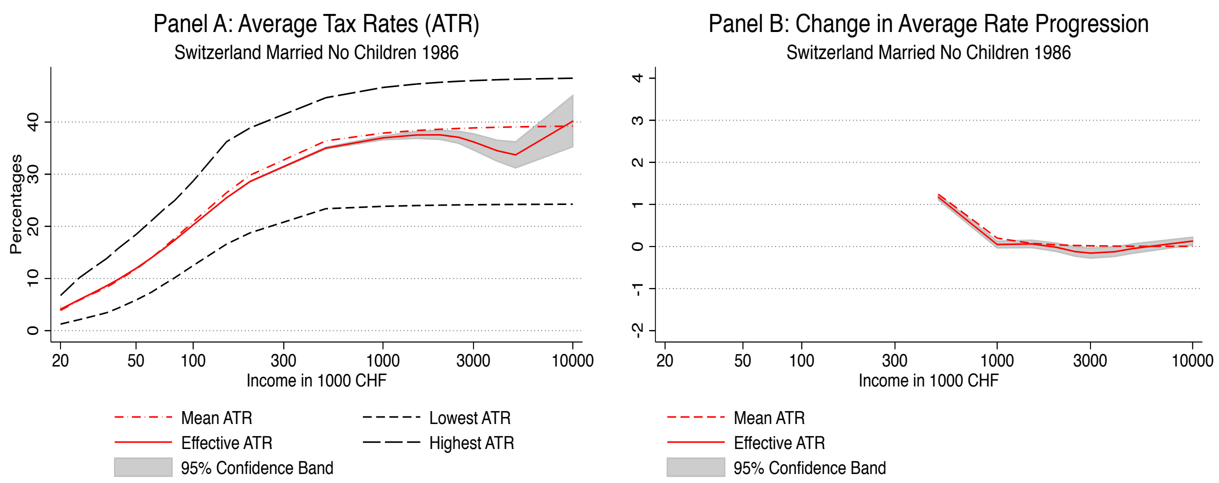

Figure 15: Country-wide average tax rate and progressivity of married couples without children in 1986. 348,876 observations, local polynomial regression (order 3) with Epanechnikov kernel and half-log-bandwidth 0.511. Change in average rate progression is multiplied by 100,000.

until 1986, where we find clear evidence for systematic income sorting. The size of the difference between the effective measures and their benchmarks however is much smaller than in 2009. Because the possibility of avoiding high tax rates was even higher in 1974, we conclude that there was less systematic sorting of high income taxpayers to low-tax municipalities in 1974. This might be explained by generally lower residential mobility in 1974. Furthermore, these results suggests that it might not have been the rich people living in certain municipalities allowing the municipalities to lower their taxes but rather rich people moving to low tax municipalities that led to the strong income sorting we observe today.

8 Considering Housing Price Capitalization

In spatial equilibrium, low local tax rates should capitalize into high local housing prices. The benefits from lower tax rates are (at least partly) offset by higher local prices (Basten et al., 2014; Morger, 2013; Schmidheiny, 2006a). Figure 16 shows on the left a map of the average tax rate a single household with median income faces in the metropolitan area of Zurich. On the right it shows a map with the median per square meter rents in this area. The two maps indicate that there is at least some capitalization of the tax rates into housing prices in the Zurich metropolitan area.

Our measure of the effective average tax rate in a fiscally decentralized country does not account for differences in real income due to such housing price differences. From the perspective of the state, this is the relevant measure as it measures the degree to which households contribute to the common good. However, from the perspective of the individual households, we should also account for the additional burden paid indirectly

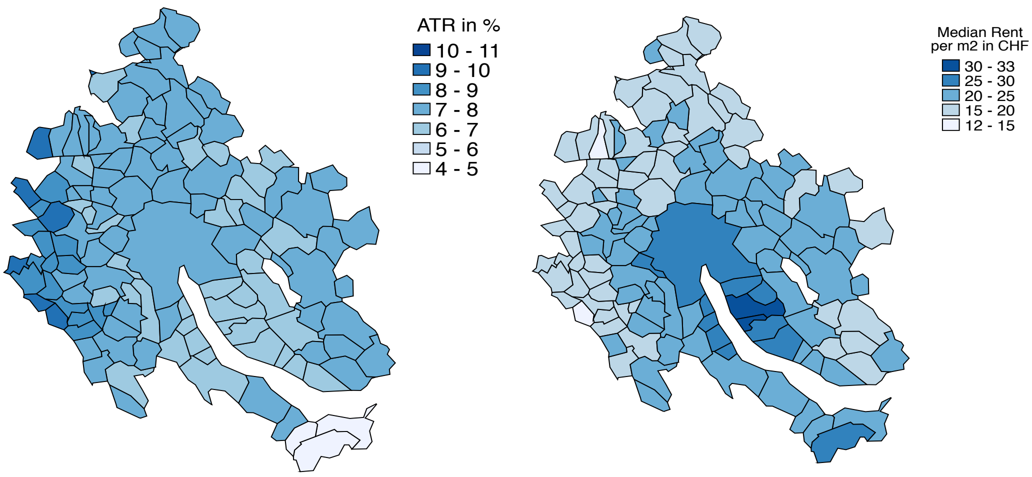

Figure 16: Average Tax Rates for Umarried Taxpayers with Median Gross Income and Median Rents - Zurich

through house prices.

The household perspective is the motive behind the concept of the *residual income progression* discussed in section 2. This concept allows to naturally account for the disutility of higher local housing prices. We propose the following new measure: The *equivalent residual income progression*. The basic idea behind this measure is to compare *residual utility* instead of residual income. Residual utility accounts for both different tax rates and different local prices. However, using an ordinal concept, utility is only identified up to a monotone transformation. Measuring the progressivity of residual utility would depend on some arbitrary transformation. We therefore use the utility function to calculate an equivalent income which a household would need to obtain in a reference municipality in order to be indifferent to its actual place of residence. This is similar in spirit to the concept of the equivalent variation used for policy evaluations in microeconomics.

Consider a larger geographic are that is divided into jurisdictions $j = 1, .., J$. A household i maximises his utility over a bundle of goods x and housing h subject to his budget constraint:

$$\max_{x,h} U(h, x) = h^{\alpha} x^{(1-\alpha)} \tag{21}$$

$$\text{s.t. } [1 - ATR_j(y)]y \geq px + r_j h \tag{22}$$

Where the p and r denote the price vector of all other goods and housing respectively. Note that the price vector p is independent of the residential jurisdiction j. This is a reasonable assumption within a metropolitan area where prices of mobile goods should not vary across jurisdictions. Also the nominal income is exogenous in this model. Considering a metropolitan area this is a reasonable assumption but likely not for a whole country where the occupation choice might influence the set of possible location choices. Hence, we can apply our measure only to metropolitan areas where it is reasonable to assume that people having an occupation within the metropolitan area will chose their residential jurisdictions amongst the jurisdictions in the metropolitan area without having to change their occupation.

Solving this model, it is easy to show that the indirect utility of an individual with income y living in jurisdiction j is:

$$V_j(y) = \left[(1-\alpha)\frac{1}{p}\right]^{(1-\alpha)} \left[\alpha\frac{1}{r_j}\right]^{\alpha} [1 - ATR_j(y)]y \tag{23}$$

Let the residual income denote as: $R_j(y) = [1 - ATR_j(y)]y$. Then we can rewrite the indirect utility as:

$$V_j(y) = \gamma \left[\frac{1}{r_j}\right]^{\alpha} R_j(y) \tag{24}$$

The equivalent residual income R^e in the main city m as reference location gives the residual income an individual with income y would need to earn in the main city of the

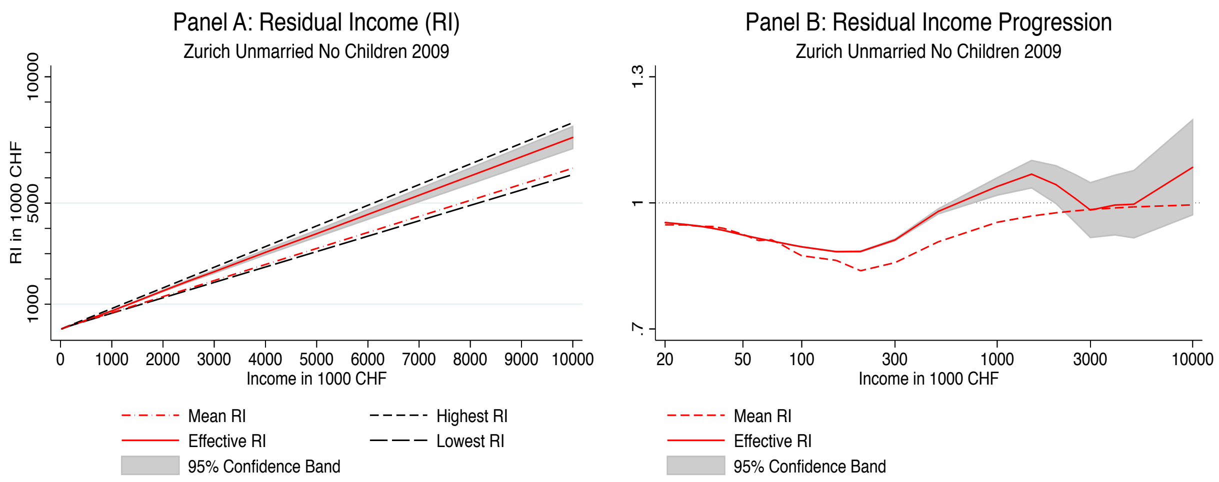

Figure 17: Residual income and residual income progression of unmarried taxpayers without children - Zurich. 286,831 observations, local polynomial regression (order 3) with Epanechnikov kernel and half-log-bandwidth 0.459.

metropolitan area to gain the same utility as in his residential jurisdiction j.

$$\left[\frac{1}{r_m}\right]^\alpha R^e = \left[\frac{1}{r_j}\right]^\alpha R_j(y) \tag{25}$$

$$R^e = \left[\frac{r_m}{r_j}\right]^\alpha R_j(y) \tag{26}$$

Thus, we have a measure that enables us to compare real net-of-tax incomes within a metropolitan area and therefore we can also calculate the progressivity of the tax system in terms of real net-of-tax income. All we need in addition is the relative amount of residual income spent on housing α and the housing prices in the jurisdictions.

In Switzerland, the share of residual income spent on housing is 21%.[15] The housing prices are median rents per square meter in 2009 of the single municipalities.[16]

We first present estimates of the residual income progression for the Zurich area without accounting for housing price differences: The left panel of Figure 17 shows a plot of the nominal residual income against the nominal gross income. In analogue to above, the solid red line depicts the effective residual income where the dashed red line depicts the mean residual income (benchmark). For high incomes the effective residual income is clearly above the benchmark. The right panel depicts the effective residual income progression. As shown above it is supposed to deliver the same result as the change in average rate progression regarding the classification into regressive and progressive parts of the income tax system. Here this is indeed the case. Also this measure clearly shows significant regressivity for incomes between 1 and 2.5M CHF.

Figure 18 shows the same measures for the *equivalent* residual income progression, i.e. accounting for the disutility of local house prices. There are no remarkable changes if we account for housing price differences in the Zurich metropolitan area. The regressivity

[15]Source: Bundesamt für Statistik Haushaltshaltsbudgeterhebung 2011

[16]Source: Rent prices from the housing advertisements of all major Swiss online housing rental platforms

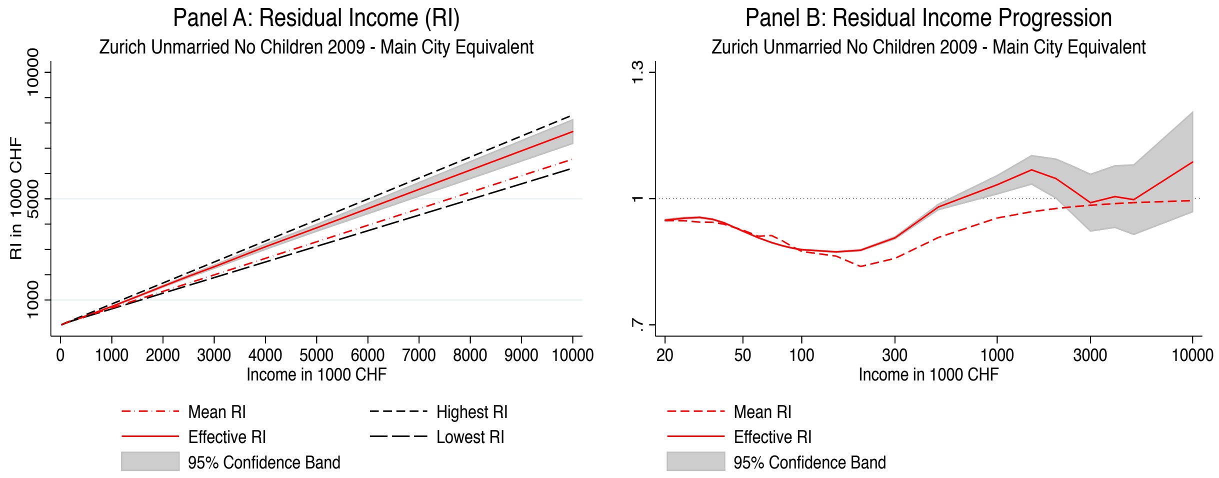

Figure 18: Residual income and residual income progression - main city equivalent - of unmarried taxpayers without children - Zurich. 286,831 observations, local polynomial regression (order 3) with Epanechnikov kernel and half-log-bandwidth 0.459.

survives for the respective income range. Consequently, those people do not only contribute relatively less to the public good but also suffer relatively less in terms of their utility. Thus, for the range of 1 up to 2.5M CHF we have a redistribution of utility to higher incomes. Furthermore, these findings show that the low tax rates do not capitalise into high housing prices for those very high income people. This is in line with the results of Morger (2013) who also shows that the capitalization of tax differentials into housing prices is far from 100% in Switzerland.

9 Conclusion

We propose a set of measures that allows to quantify the overall level and the progressivity in a fiscally decentralized country. We define two measures for the level of taxation: the *effective average tax rate* and the *effective marginal tax rate*. Additionally, we define three progressivity measures: the *country-wide change in average rate progression*, *country-wide liability progression*, and the *country-wide residual income progression*. All our measures take the relative size of the jurisdictions and potential income sorting into account. Furthermore, we provide an estimation strategy for our measures together with a benchmark that allows to determine the roll of systematic income sorting.

Exploiting the universe of Swiss income taxpayers, we find high income households face lower tax rates than in the benchmark case. This is because high income households sort systematically into low tax municipalities. The same holds for the effective progressivity. The direction of this effect is robust across different household types but the size varies. For families with two children, the difference between the effective tax rates and the benchmark is fairly small. In contrast, for married couples without children, we even find a regressive part of the income tax system for very high incomes. Our results also hold within metropolitan areas. Historically, the size of the effect increased over time which is likely due to higher mobility of taxpayers. Accounting for housing price capitalization does not alter our results.

The contribution of this paper is twofold: Fist, it contributes to the basic taxation literature by providing a consistent set of measures that allow to quantify the effective tax rates and the effective progressivity of decentralized tax systems. Second, it contributes to the empiric literature on income segregation showing that systematic income sorting leads to lower effective tax rates and to a lower effective progressivity of the income tax system for high income households in Switzerland.

References

Atkinson, A. B. (1970). On the measurement of inequality. *Journal of economic theory 2*(3), 244–263.

Basten, C. C., M. Von Ehrlich, and A. Lassmann (2014). Income taxes, sorting, and the costs of housing: Evidence from municipal boundaries in switzerland.

Bayer, P., F. Ferreira, and R. McMillan (2007). A unified framework for measuring preferences for schools and neighborhoods. *Journal of Political Economy 115*(4), 588–638.

Bayer, P. and R. McMillan (2012). Tiebout sorting and neighborhood stratification. *Journal of Public Economics 96*(11), 1129–1143.

Black, S. E. (1999). Do better schools matter? parental valuation of elementary education. *Quarterly journal of economics*, 577–599.

Boadway, R. and M. Keen (2000). Redistribution. *Handbook of income Distribution 1*, 677–789.

Brülhart, M., S. Bucovetsky, and K. Schmidheiny (2015). Taxes in cities: Interedependence, asymmetrey and agglomeration. *Handbook of Regional and Urban Economics 5*.

Calabrese, S., D. Epple, and R. Romano (2007). On the political economy of zoning. *Journal of Public Economics 91*(1), 25–49.

Calabrese, S. M., D. N. Epple, and R. E. Romano (2012). Inefficiencies from metropolitan political and fiscal decentralization: Failures of tiebout competition. *The Review of Economic Studies 79*(3), 1081–1111.

Conesa, J. C. and D. Krueger (2006). On the optimal progressivity of the income tax code. *Journal of Monetary Economics 53*(7), 1425–1450.

Epple, D., R. Filimon, and T. Romer (1984). Equilibrium among local jurisdictions: Toward an integrated treatment of voting and residential choice. *Journal of public economics 24*(3), 281–308.

Epple, D. and T. Nechyba (2004). Fiscal decentralization. *Handbook of regional and urban economics 4*, 2423–2480.

Epple, D. and G. J. Platt (1998). Equilibrium and local redistribution in an urban economy when households differ in both preferences and incomes. *Journal of Urban Economics 43*(1), 23–51.

Epple, D., R. Romano, and H. Sieg (2012). The intergenerational conflict over the provision of public education. *Journal of Public Economics 96*(3), 255–268.

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

Epple, D. and H. Sieg (1999). Estimating equilibrium models of local jurisdictions. *Journal of Political Economy 107*(4), 645–681.

Fan, J. and I. Gijbels (1996). Local polynomial modelling and its applications. *Chapman, Hall, London.*

Fullerton, D. (1984). Which effective tax rate? *National Tax Journal*, 23–41.

Goodspeed, T. J. (1989). A re-examination of the use of ability to pay taxes by local governments. *Journal of Public Economics 38*(3), 319–342.

Gouveia, M. and R. P. Strauss (1994). Effective federal individual income tax functions: An exploratory empirical analysis. *National Tax Journal*, 317–339.

Gravel, N. and R. Oddou (2014). The segregative properties of endogenous jurisdiction formation with a land market. *Journal of Public Economics 117*, 15–27.

Gruber, J. and E. Saez (2002). The elasticity of taxable income: evidence and implications. *Journal of public Economics 84*(1), 1–32.

Hodler, R. and K. Schmidheiny (2006). How fiscal decentralization flattens progressive taxes. *FinanzArchiv: Public Finance Analysis 62*(2), 281–304.

Jakobsson, U. (1976). On the measurement of the degree of progression. *Journal of public economics 5*(1), 161–168.

Mendoza, E. G., A. Razin, and L. L. Tesar (1994). Effective tax rates in macroeconomics: Cross-country estimates of tax rates on factor incomes and consumption. *Journal of Monetary Economics 34*(3), 297–323.

Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. *The review of economic studies*, 175–208.

Morger, M. (2013). Heterogeneity in income tax capitalization and its effects on segregation within switzerland. *Arbeitspapiere - Eidgenössische Steuerverwaltung.*

Musgrave, R. A. and T. Thin (1948). Income tax progression, 1929-48. *The Journal of Political Economy 56*(6), 498–514.

Oates, W. E. (1969). The effects of property taxes and local public spending on property values: An empirical study of tax capitalization and the tiebout hypothesis. *The Journal of Political Economy*, 957–971.

Parchet, R. (2014). *Are local tax rates strategic complements or strategic substitutes?* IdEP Economic Papers 2014/07.

Pigou, A. C. (1951). *A study in public finance*. London: Macmillan.

Ross, S. and J. Yinger (1999). Sorting and voting: A review of the literature on urban public finance. *Handbook of regional and urban economics 3*, 2001–2060.

Rothschild, M. and J. E. Stiglitz (1973). Some further results on the measurement of inequality. *Journal of Economic Theory 6*(2), 188–204.

Saez, E. (2001). Using elasticities to derive optimal income tax rates. *Review of Economic Studies 68*(1), 205–29.

Saez, E., J. Slemrod, and S. H. Giertz (2012). The elasticity of taxable income with respect to marginal tax rates: A critical review. *Journal of Economic Literature*, 3–50.

Schaltegger, C. A., F. Somogyi, and J.-E. Sturm (2011). Tax competition and income sorting: Evidence from the zurich metropolitan area. *European Journal of Political Economy 27*(3), 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.

Slitor, R. E. (1948). The measurement of progressivity and built-in flexibility. *The Quarterly Journal of Economics 62*(2), 309–313.

Tiebout, C. M. (1956). A pure theory of local expenditures. *The journal of political economy*, 416–424.

Westhoff, F. (1977). Existence of equilibria in economies with a local public good. *Journal of economic Theory 14*(1), 84–112.

10 Appendix

A Theoretical Model

We build up on this model of Schmidheiny (2006a) except that we relax some assumptions made on the tax schedule. Instead of assuming that the tax functions in both jurisdictions are a multiple of each other, we only assume that the tax functions in both jurisdictions do not cross, i.e. the order of jurisdictions with respect to the level of taxation is perfect and consistent over gross income and that higher tax rates imply higher progressivity.

There are two jurisdictions $j = 1, 2$ with different progressive tax functions $ATR_j(y)$: $\partial ATR_j(y)/\partial y > 0 \ \forall y$. Without loss of generality we assume that $ATR_1(y) > ATR_2(y)$ and $\partial ATR_1(y)/\partial y > \partial ATR_2(y)/\partial y$ 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 jurisdiction 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 jurisdiction. This is the primary result we build on.

As defined above, the *effective average tax rate* $ATR(y)$ is then given as:

$$\begin{aligned} ATR(y) &= \sum_{j=1}^{2} P(j|y) ATR_j(y) \qquad (27) \\ &= P(1|y) ATR_1(y) + [1 - P(1|y)] ATR_2(y) \end{aligned}$$

And the *mean average tax rate* is:

$$MATR(y) = \sum_{j=1}^{2} P(j) ATR_j(y) \qquad (28)$$

Let the difference of both level measures denote as:

$$\begin{aligned} \Delta(y) &= ATR(y) - MATR(y) \qquad (29) \\ &= [P(1|y) - P(1)][ATR_1(y) - ATR_2(y)] \qquad (30) \end{aligned}$$

Then we can derive the following proposition:

Proposition 1. *If both jurisdictions are populated and if the housing prices p_j differ and the tax rates in jurisdiction 1 are strictly larger than in jurisdiction 2, then there exists an gross income y^* such that for all gross incomes below y^* the effective average tax rate is higher (strictly higher for some y) than the mean average tax rate and that for all gross incomes above y^* the effective average tax rate is lower (strictly lower for some y) than the mean average tax rate.*

Proof: $[ATR_1(y) - ATR_2(y)] > 0$ 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, $\Delta(\underline{y}) > 0$ 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 $\partial P(1|y)/\partial P(1) \leq 0$ 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 effective average tax rate of low income people is higher than the mean average tax rate, while it is lower for people with high incomes.

In this two jurisdiction model we can write the effective marginal tax rate as:

$$MTR(y) = \frac{\partial T(y)}{\partial y} = \frac{\partial E\{T_j(y)|y\}}{\partial y} \tag{31}$$

$$= \underbrace{\frac{\partial P(1|y)}{\partial y}[T_1(y) - T_2(y)]}_{\leq 0} + \underbrace{P(1|y)\frac{\partial T_1(y)}{\partial y} + [1 - P(1|y)]\frac{\partial T_2(y)}{\partial y}}_{E\{\partial T_j/\partial y|y\}} \tag{32}$$

From equation 32 it follows directly that the effective marginal tax rate is even lower than a weighted mean of the local marginal tax rates that takes the income sorting into account $E\{\partial T_j/\partial y|y\}$.

The mean marginal tax rate is:

$$MMTR(y) = \frac{\partial MT(y)}{\partial y} \tag{33}$$

$$= P(1)\frac{\partial T_1(y)}{\partial y} + [1 - P(1)]\frac{\partial T_2(y)}{\partial y} \tag{34}$$

Let the difference of both denote as $D(y)$:

$$D(y) = MTR(y) - MMTR(y) \tag{35}$$

$$= \underbrace{\frac{\partial P(1|y)}{\partial y}[T_1(y) - T_2(y)]}_{\leq 0} + [P(1|y) - P(1)] \underbrace{\left[\frac{\partial T_1(y)}{\partial y} - \frac{\partial T_2(y)}{\partial y}\right]}_{\leq 0} \tag{36}$$

Proposition 2. *If both jurisdictions are populated and if the housing prices p_j differ and the tax rates in jurisdiction 1 are strictly larger than in jurisdiction 2, then there exists an gross income y^{**} such that for all gross incomes above y^{**} the effective marginal tax rate is lower (strictly lower for some y) than the mean marginal tax rate.*

Proof: The proof directly follows as a corollary from proposition 1 and the negativity of the first addend.
Thus, we can show that in this two jurisdiction model also the effective marginal tax rate is lower than the mean marginal tax rate for high incomes. Note that we cannot make a general statement about the relationship for incomes lower than y^{**}.

Beyond the level predictions our model has also some interesting implications for the progressivity of the tax system:

$$\frac{\partial ATR(y)}{\partial y} = \sum_{j=1}^{2} \frac{\partial P(j|y)}{\partial y} ATR_j(y) + \sum_{j=1}^{2} P(j|y) \frac{\partial ATR_j(y)}{\partial y} \tag{37}$$

$$= \underbrace{\frac{\partial P(1|y)}{\partial y} [ATR_1(y) - ATR_2(y)]}_{\leq 0} + \underbrace{P(1|y) \frac{\partial ATR_1(y)}{\partial y} + [1 - P(1|y)] \frac{\partial ATR_2(y)}{\partial y}}_{E\{\partial ATR_j(y)/\partial y\}} \tag{38}$$

Because $\partial P(1|y)/\partial y < 0$ and $ATR_1(y) - ATR_2(y) > 0$, the first addend is negative. The second addend is the expected value of the local change in 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 change in 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 change in average rate progression like in our model[17] of Hodler and Schmidheiny (2006), we can derive the following relationship: Let the change in average rate progression of the mean average tax rate be:

$$\frac{\partial MATR(y)}{\partial y} = \sum_{j=1}^{2} P(j) \frac{\partial ATR_j(y)}{\partial y} = P(1) \frac{\partial ATR_1(y)}{\partial y} + [1 - P(1)] \frac{\partial ATR_2(y)}{\partial y} \tag{39}$$

Then the difference of this progressivity measures is:

$$\delta(y) = \underbrace{\frac{\partial P(1|y)}{\partial y} [ATR_1(y) - ATR_2(y)]}_{\leq 0} + [P(1|y) - P(1)] \left[\frac{\partial ATR_1(y)}{\partial y} - \frac{\partial ATR_2(y)}{\partial y} \right] \tag{40}$$

Proposition 3. *If both jurisdictions are populated and if the housing prices p_j differ and the tax rates in jurisdiction 1 are strictly larger than in jurisdiction 2, then there exists an gross income y^{***} such that for all gross incomes above y^{***} the effective progressivity is lower (strictly lower for some y) than the progressivity of the mean average tax rate.*

Proof: The proof directly follows as a corollary from proposition 1, the fact that $\partial ATR_1(y)/\partial y > \partial ATR_2(y)/\partial y$, and the negativity of the first addend.

Therefore, the two jurisdiction model does not only imply that the level of the effective average tax rate is below the mean average tax rate for high incomes but also that its progressivity is lower than the one of the benchmark. For low incomes we cannot derive a similar statement.

[17] This is the case in Switzerland for the local taxes within a canton.

B Proofs

B.1 Consistency of Progressivity Measures

Liability Progression For strictly positive income, $y > 0$, and strictly positive average tax rates, $ATR(y) > 0$, the liability progression and the change in average rate progression are consistent in their classification of tax schedules into regressive or progressive. Proof:

$$\frac{\partial T(y)}{\partial y} \frac{y}{T(y)} > 1 \tag{41}$$

$$\left(\frac{\partial ATR(y)}{\partial y} y + ATR(y) \right) \frac{1}{ATR(y)} > 1 \tag{42}$$

$$\frac{\partial ATR(y)}{\partial y} \frac{y}{ATR(y)} > 0 \tag{43}$$

$$\frac{\partial ATR(y)}{\partial y} > 0 \tag{44}$$

Residual Income Progression For strictly positive income, $y > 0$, and strictly positive average tax rates, $ATR(y) > 0$, and average tax rates of less than 100%, $ATR(y) < 1$, the residual income progression and the change in average rate progression are consistent in their classification of tax schedules into regressive or progressive. Proof:

$$\frac{\partial \left[y - T(y)\right]}{\partial y} \frac{y}{\left[y - T(y)\right]} < 1 \tag{45}$$

$$\left(1 - \frac{\partial ATR(y)}{\partial y} y - ATR(y) \right) \frac{1}{1 - ATR(y)} < 1 \tag{46}$$

$$\frac{\partial ATR(y)}{\partial y} \frac{y}{1 - ATR(y)} > 0 \tag{47}$$

$$\frac{\partial ATR(y)}{\partial y} > 0 \tag{48}$$

Thus, all three progressivity measures are consistent in their classification of tax schedules into regressive or progressive under regular conditions.

C Data

C.1 Gross Income

Because we only observe the taxable income and the average tax rates are defined over the gross income, we need to reconstruct the gross income of all Swiss taxpayers. The annual publication of the Swiss Federal Tax Administration also contains average tax rates on the federal level $ATR^f(y)$. From this, we calculate the tax liabilities at all published income levels:

$$T_f(y) = ATR_f(y)y \tag{49}$$

For all household types we can calculate the implicit deductions by comparing this tax liability with the statutory tax liability. The statutory tax[18] liability $S(z)$ is defined over the taxable income $z = y - d$. Because the deductions d are the only unknown in the following equation, we can reconstruct the assumed deductions of the Swiss Federal Tax Administration for the respective household type with gross income y.

$$T_f(y) = S(y^t) = S(y - d) \tag{50}$$

$$d(y) = y - S^{-1}\left(T_f(y)\right) \tag{51}$$

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

$$d = a + by \tag{52}$$

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 5 represents the results. While singles are assumed to deduct a fix amount of 4,679 CHF, couples to deduct 9,063 CHF, and families 22,912 CHF. Additionally, all household types are assumed to deduct 10% of their gross income.

Table 5: Deductions

	Unmarried No Children	Married Couple No Children	Married Couple 2 Children
Fix amount	4679	9063	22912
Fraction of income	0.1032	0.1032	0.1036

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 liabilities by the ESTV that we use to calculate the average tax 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 z_i as follows:

$$y_i = \frac{z_i + a}{1 - b} \tag{53}$$

C.2 Individual Tax Rates

In a 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

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

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:

$$ATR_i = \frac{\theta(m_i)(y_i - 1,000,000) + T(1,000,000, m_i)}{y_i} \tag{54}$$

$$\theta(m_i) = \frac{T(1,000,000, m_i) - T(500,000, m_i)}{500,000} \tag{55}$$

This method assumes that the maximum marginal tax rate $\theta(j)$ is already reached at a gross income of 500,000 CHF. This assumption holds exactly for 21 cantons and is a very close approximation for the remaining 5. Furthermore, the deductions must grow linearly in the gross income which is already assumed in the average tax rate data.