

How Fiscal Decentralization Flattens Progressive Taxes*

Roland Hodler
University of Bern

Kurt Schmidheiny
Tufts University

This revision: April 6, 2006

Abstract

We study the tension between fiscal decentralization and progressive taxation. We present a multi-community model in which households differ in incomes and housing preferences and in which the local income tax rate is a function of an exogenous progressive tax schedule and an endogenous local tax multiplier. The progression of the tax schedule induces a self-sorting process that results in substantial though imperfect income sorting. Considering this income sorting, the resulting tax schedule is less progressive than the exogenous tax schedule. Empirical evidence from the largest Swiss metropolitan area supports the predictions of our model.

JEL: H73, R23

Key words: Progressive Taxation, Fiscal Decentralization, Income Segregation

*We would like to thank John Ashworth, Bruno Heyndels, Mark Schelker and participants at the Public Choice Society Meeting, the European Public Choice Society Meeting and Norwegian-German Seminar in Public Economics for helpful comments.

1 Introduction

In this paper, we study the tension between fiscal decentralization and progressive taxation. We investigate to what extent fiscal decentralization reduces the progression of a common tax schedule in a federation in which communities can set the level of taxation. We find that progressive taxation and fiscal decentralization lead to income sorting, which reduces the progression of the tax schedule.

We base our analysis on a multi-community model, in which the income tax rate is a function of an exogenous progressive tax schedule and an endogenous local tax multiplier. Local tax revenue is used to finance a local good. In the basic version of the model, the mobile individual households differ only in their incomes. In equilibrium, no household wants to move, local housing markets clear and the communities' budgets are balanced. It follows that the local tax multipliers must be higher in communities in which housing prices are lower. The progression of the tax schedule then induces a self-sorting process that results in perfect income sorting. Poor households locate in high-tax communities while rich households locate in low-tax communities. Different from the previous literature,¹ sorting is explained by the progression of the tax schedule and does not require strong assumptions on the preferences for either public goods or housing. While some degree of income sorting is observed in reality, this sorting is never perfect (see e.g. Epple and Sieg, 1999, Hardman and Ioannides, 2004, and Ioannides, 2004). A potential reason for *imperfect* income sorting is that households differ in their preferences. We therefore add heterogeneity in the households' preferences to our model. In particular, we assume heterogeneity in the preferences for housing. As each household's location choice depends now on its income and its preferences, the income sorting is imperfect in equilibrium: households with the same income are found in different communities, though rich households are still more likely to locate in low-tax communities than poor households.

Tax progression - measured as liability progression - is in our setup independent of the tax multiplier and therefore constant across communities. In equilibrium, rich households avoid high tax rates by moving to low-tax communities. Considering income sorting, the *mean tax schedule* is therefore shown to be less progressive than the tax schedule in any given community. The mean tax schedule depicts the tax rates as faced by the households as a function of income.

To illustrate the implications of this model, we calibrate a fully-specified version to the Zurich area, the largest Swiss metropolitan area. Swiss metropolitan areas offer an excellent laboratory

¹See e.g. Ellickson (1971), Westhoff (1977), Epple and Romer (1991) and the literature surveyed in Ross and Yinger (1999).

for the analysis of fiscal decentralization. In Switzerland, each community can individually set the *level* of income taxes by a local tax multiplier while the cantons (states) fix the (progressive) *schedule* of income taxes. The equilibrium values of this simulation show the same pattern across communities as we observe in the Zurich area: Rich households locate mainly, but not exclusively, in communities with low tax multipliers and poor households mainly in communities with high tax multipliers.

We then use data on the spatial distribution of incomes in the Zurich area to estimate the mean tax schedule in this area. We find that the mean tax schedule is less progressive than the tax schedule implemented by the canton because rich households are more likely to live in low-tax communities than poor households. This finding is in line with the predictions of our theoretical model. The observed differences in taxes and housing prices are however relatively small compared to the results of our calibrated model. We argue that the tax equalization programs of the canton of Zurich could account for some of these differences.

This paper is most closely related to Feldstein and Wrobel (1998), Epple and Platt (1998) and Schmidheiny (2006a). Feldstein and Wrobel (1998) show that a shift in a single US state's tax progression has no redistributive effects since migration leads to an adjustment of the net wages and the employment structure. Complementary, we show that location choices undermine the redistributive effect of progressive taxation in presence of fiscal decentralization even if wages do not adjust. Epple and Platt (1998) include preference heterogeneity as we do but study local property taxation rather than local income taxation. Schmidheiny (2006a) estimates the location choice part of our model but does not analyze its equilibrium properties.

More generally, our paper relates to the literature on fiscal federalism.² It is a well known normative principle of this literature that income redistribution should be centralized. As progressive income taxes are a particular mean to redistribute income, it directly follows that "progressive income taxation [...] - if substantial in scope - must be uniform within the entire area over which there is a high degree of capital and labor mobility" (Musgrave 1971, p. 7). In this paper, we provide some support for this view by showing that fiscal decentralization does indeed undermine the progression of the tax schedule. However, we also find theoretical and empirical evidence that the income sorting of the population does not completely offset progression; hence leaving room for substantial redistribution through progressive income taxation at the local level.

The paper is organized as follows: Section 2 briefly informs about fiscal decentralization and

²For a recent survey of this literature see Oates (1999).

progressive taxation in Switzerland and some other countries with a comparable tax system. Section 3 presents the theoretical model and some results concerning the agents' location choice. It further proves that an (asymmetric) equilibrium exists. Section 4 presents the simulation of a fully-specified version of our model, which is calibrated to the Zurich metropolitan area. Section 5 estimates the mean tax schedule faced by the households in this area. Section 6 concludes.

2 Fiscal Decentralization and Progressive Taxation in Switzerland (and elsewhere)

Switzerland is an exemplary federal fiscal system. The Swiss federation comprises 26 states, the so-called cantons. The cantons are divided into roughly 3000 municipalities of varying size and population. All three state levels finance their expenditures essentially by their own taxes and fees. 46% of the total tax revenue are imposed by the federation, 32% by the cantons and 22% by the municipalities.³ While the federal government is mainly financed by indirect taxes (61% of federal tax revenue) such as the VAT, the cantons and municipalities largely rely on direct taxes. Income taxes account for 60% of cantonal and 84% of municipal tax revenue.

The cantons organize their tax systems autonomously. For example, they decide upon the level of income and corporate taxes and the degree of tax progression. The individual municipalities in turn can set a tax multiplier for income and corporate taxes. The municipal tax is then the cantonal tax rate multiplied by the municipal tax multiplier. Federal and cantonal systems of fiscal equalization limit the tax differences across cantons and across municipalities within the same canton to some extent, but still leave room for considerable variation.

The above outlined federal system leads to ample differences of income taxes across Swiss municipalities. For example, for a two-child family with a gross income of 80,000 Swiss francs (CHF) combined cantonal and municipal income taxes ranged from 3.6% to 11.3% in the year 1997 (and its federal income tax was 0.7%). With an income of 500,000 CHF a two-child family faced much higher tax rates due to the progression of the tax schedules. Combined cantonal and municipal income taxes ranged from 10.9% to 28.7%⁴ for this household (and its federal income tax was 9.4%). Within metropolitan areas the (municipality) tax differences are smaller but still differ by a factor of 1.5 in e.g. the Zurich area.

While local taxation of *property* is widespread, especially in the United States, local taxation

³All figures in this paragraph apply to 2001. Source: Swiss Federal Tax Administration (2002), Öffentliche Finanzen der Schweiz 2001, Neuchâtel: Swiss Federal Statistical Office.

⁴Note, however, that this high tax rate is rather hypothetical as it is very unlikely to find a household with such high income in one of the very high-tax municipalities.

of *income* is rarer. Local income taxation at municipal level is e.g. observed in four U.S. states (Indiana, Maryland, Ohio, and Pennsylvania) and in Denmark. Different from the progressive local tax scheme in Switzerland, these states and countries apply a *flat* local tax. Belgium is to our knowledge the only country with a similar system of fiscal decentralization at municipal level as Switzerland. In Belgium, each of the three regions, i.e. Flanders, Wallonia and the Brussels Region, collects progressive income taxes. Furthermore, each municipality can generate its own income tax revenue by adding a fixed percentage surcharge on the (progressive) regional income tax. Canada had a similar system at provincial level between 1977 and 1996 (see Boadway and Kitchen, 1980): Personal income taxes in Canadian provinces (except Quebec) were a percentage of the (progressive) federal tax.

Tax competition is often constrained by tax equalization programs. This is also the case in Switzerland. In the canton of Zurich, as an example, there are horizontal and vertical tax equalization programs that limit the tax (multiplier) competition among communities: First, there is a program that redistributes tax revenues from the communities with the highest per capita tax bases to the communities with the lowest tax bases. Second, the canton of Zurich subsidizes the communities with the highest tax multipliers. The aggregate support for communities with low tax bases and high tax multipliers was about 200 million Swiss francs in 1999. In addition, the city of Zurich, which does not participate in these tax equalization programs, received cantonal payments of about 130 million Swiss francs.⁵ As described above, the observed tax differentials are despite these equalization programs substantial in the Zurich area.

3 The Model

In this section, we introduce and solve the model. After presenting the general setting, we characterize the preferences and derive the resulting allocation of households across distinct communities. We then prove the existence of an asymmetric equilibrium, and we introduce heterogeneity in the households' preferences. Finally, we discuss how the resulting segregation affects the tax schedule and its progression.

3.1 The Setting

There is a metropolitan area with two communities $i = 1, 2$ of given land area L_i .⁶ The area is populated by a continuum of households, which differ in their income $y \in [\underline{y}, \bar{y}]$. Income follows

⁵Source: Canton of Zurich (2000), Bestandesaufnahme des Finanzausgleichs des Kantons Zürichs.

⁶The restriction to two communities is made for simplicity.

a distribution function $f(y) > 0$.

There are three goods in the economy: private consumption b , housing h and a publicly provided local good g . The housing h is provided by absentee landlords, and housing markets are competitive. In each community i , the price for housing p_i equates housing supply HS_i with aggregate housing demand HD_i . We assume that housing supply $HS_i = HS(L_i, p_i,)$ is a non-decreasing function of land area L_i and housing price p_i .

Each community i spends the amount $n_i g$ to provide the local good g , where n_i is the measure of households living in community i . The communities levy income taxes to finance the local good. In each community i , the tax rate consists of two parts, a local tax multiplier t_i and a tax schedule $r(y) > 0$, which is continuous and progressive, i.e. $r'(y) > 0$. Total taxes are thus $T_i(y) = t_i r(y) y$ for a household with income y , the average tax rate is $T_i(y)/y = t r(y) \in [0, 1)$ and the marginal tax rate $\partial T_i(y)/\partial y = t_i[r(y) + y r'(y)] \in [0, 1)$. The tax schedule $r(y)$ is exogenous (to the communities) and equal in both communities.

The progression of the tax rate can be measured as *liability progression* (see Musgrave and Thin, 1948). Liability progression is the elasticity of total taxes with respect to income:

$$\rho_i(y) := \frac{\partial T_i(y)}{\partial y} \frac{y}{T_i(y)} = \frac{r(y) + y r'(y)}{r(y)} > 1. \quad (1)$$

Note that liability progression is independent of the tax multiplier t_i and therefore equal in both communities: $\rho_i(y) = \rho(y)$ for $i = 1, 2$.⁷

We assume that the local good g is fixed and identical across communities. In each community i , the tax multiplier t_i is therefore determined by budget balance. There are two reasons for assuming exogeneity of the local good. First, the local good g can be thought as a locally provided, locally financed but centrally decided good. We think that many locally financed goods, particularly in Switzerland, satisfy this description. Schooling, for example, accounts for the largest item in municipal budgets in Switzerland; local neighborhood schools are locally provided and locally financed, however, cantonal regulation leave little discretionary power for financially relevant decisions.⁸ Our model focuses therefore on the revenue side of fiscal decentralization. Second, we think that the progression of income taxes is a very important factor for income sorting in Switzerland. However, our model would become intractable allowing for both progressive taxation and endogenous provision of local goods.⁹

⁷Of course, liability progression would also be equal to $\rho(y)$ in case of full tax equalization across communities.

⁸In particular, teachers' salaries and class size are regulated by cantonal law. Furthermore, cantonal courts ruled, based on equity considerations, that schools (in rich neighborhoods) are not allowed to provide additional tutoring or extra classes for extraordinary strong or weak pupils.

⁹Schmidheiny (2006b) studies endogenous local goods determined in municipal majority votes, but financed by

Further, we assume that each household can move costlessly and chooses the community maximizing its utility as place of residence.

3.2 Preferences and Location Choice

The preferences of the households are described by the utility function

$$U(h, b) , \quad (2)$$

where h is the consumption of housing and b the consumption of the private good.¹⁰ We assume the utility function to be strictly increasing, strictly quasi-concave, and twice continuously differentiable in h and b .

Households face the budget constraint (omitting community indices)

$$ph + b \leq y_d = y[1 - t \cdot r(y)] , \quad (3)$$

where p is the price of housing; the price of the private good is set to unity. Disposable income y_d depends on the local tax multiplier t and the tax rate schedule $r(y)$.

Maximization of the utility function (2) with respect to h and b subject to constraint (3) yields housing demand $h^* = h(p, y_d) = h(t, p, y)$, demand for the private good $b^* = y(1-t) - ph(t, p, y)$, and indirect utility

$$V(t, p, y) = U(h^*, b^*) . \quad (4)$$

For later use note that V is continuous in t , p and y .

We assume that the elasticity of housing with respect to the disposable income is smaller or equal to unity, i.e.,

$$\varepsilon_{h, y_d} := \frac{\partial h^*}{\partial y_d} \frac{y_d}{h^*} \leq 1 \quad \text{for all } y_d \text{ and } p. \quad (5)$$

The assumption of income elasticity for housing demand below one is well supported by the large empirical literature on housing demand. Mayo's (1981) seminal survey of empirical studies using microdata reports consistent income elasticity below one. This result is robust controlling for housing prices, demographic household variables (e.g. Mayo 1981, Hansen, Formby and Smith, 1998), tenure choice (e.g. Henderson and Ioannides 1986, Hansen, Formby and Smith, 1998) and functional form (Hansen, Formby and Smith, 1996). The assumption is also supported by the Swiss data used in the calibration of our model.

flat local income taxes. His model predicts qualitatively the same equilibrium situation as the model presented in this paper: rich households locate in low-tax communities with high housing prices while the poor prefer low-price communities with high taxes. The provision of local public goods is slightly higher in the rich low-tax community.

¹⁰Since the local good g is constant across communities and not of primary interest for our considerations, we assume for simplicity that it does not enter the utility function. Equivalently, we could assume that it enters separably.

Next, we present two properties of the households' indifference curves that will lead to segregation of the population by incomes:

Property 1

$$M(t, p, y) := \left. \frac{dt}{dp} \right|_{dV=0} = -\frac{\partial V / \partial p}{\partial V / \partial t} = -\frac{h^*}{y \cdot r(y)} < 0$$

Property 1 follows from the strictly increasing utility function after applying the implicit function theorem and the envelope theorem. It implies that a household can be made indifferent towards an increase in the tax multiplier t when it is compensated by decreased housing prices p , and vice versa.

Property 2

$$\frac{\partial M}{\partial y} = \left[1 - \varepsilon_{h, y_d} \frac{\partial y_d}{\partial y} \frac{y}{y_d} \right] \frac{h^*}{y^2 r(y)} + \frac{\partial r(y)}{\partial y} \frac{h^*}{y r^2(y)} > 0 \text{ for all } y, t \text{ and } p.$$

Proof: By assumption, $\varepsilon_{h, y_d} \leq 1$. Our assumptions about the bounds of the average and the marginal tax rate guarantee $(\partial y_d / \partial y)(y / y_d) = [1 - tr - ty r'(y)] / [1 - tr(y)] \in [0, 1]$. The assumption that $r(y)$ increases in y , implying $\partial r(y) / \partial y > 0$, concludes the proof. \square

Property 2, which is key for the subsequent results, implies that the decrease in housing prices p which compensates a household for a higher tax multiplier t has to be larger for rich households than for poor ones. The reason is that poor households care weakly more about housing prices p than rich households because $\varepsilon_{h, y_d} \leq 1$ while rich households care more about the tax multiplier t because of progression of the tax schedule, i.e. $r'(y) > 0$. Note that due to the progression of the tax schedule, property 2 even holds with homothetic preferences, i.e. in case of $\varepsilon_{h, y_d} = 1$.¹¹

Given a set of community characteristics, (p_i, t_i) for $i = 1, 2$, a household with income y prefers to live in community i if and only if

$$V(p_i, t_i, y) \geq V(p_j, t_j, y). \tag{6}$$

From this, the following proposition directly follows. We assume with no loss of generality that $t_1 > t_2$.

Proposition 1 (Order of community characteristics)

If both communities are populated and if $(p_1, t_1) \neq (p_2, t_2)$ with $t_1 > t_2$, then community 1 with the higher tax rate $t_1 > t_2$ must have lower housing prices $p_1 < p_2$ than community 2.

¹¹Moreover, property 2 would also hold if $\varepsilon_{h, y_d} < 1$ and $r'(y) = 0$.

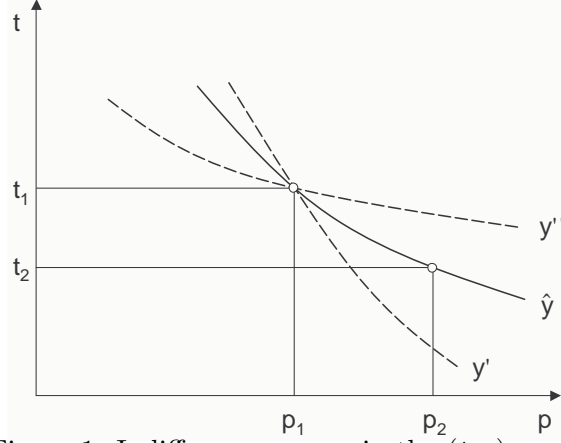


Figure 1: Indifference curves in the (t, p) space

Proof: Suppose the opposite, i.e., $p_1 \geq p_2$ and $t_1 > t_2$. In this case, no household would prefer to live in 1 (for the same reason that leads to property 1). This is a contradiction. \square

In the remaining part of this section, we show how households allocate themselves across the two communities when the communities differ in their housing prices p_i and their tax multipliers t_i .

Lemma 1 (Boundary indifference)

There is a ‘border’ household that is indifferent between the two communities. That is, if a household with income y' prefers to live in community i and a household with income $y'' > y'$ in community j , then there exists a border household with income \hat{y} , $y' \leq \hat{y} \leq y''$ who is indifferent between the two communities.

Proof: A household with y' prefers i to j , hence $V(p_i, t_i, y') - V(p_j, t_j, y') \geq 0$. The opposite is true for a household with y'' : $V(p_i, t_i, y'') - V(p_j, t_j, y'') \leq 0$. From the continuity of V in y follows the continuity of $V(p_i, t_i, y) - V(p_j, t_j, y)$ in y . The intermediate value theorem then proves that there is at least one \hat{y} between y' and y'' such that $V(p_i, t_i, \hat{y}) - V(p_j, t_j, \hat{y}) = 0$. \square

Proposition 2 (Income segregation)

If both communities are populated and if $(p_1, t_1) \neq (p_2, t_2)$ with $t_1 > t_2$, then any household in community 2 is richer than the border household and any household in community 1 is poorer than the border household.

Proof: The proof uses Figure 1, which shows the indifference curves in the (t, p) -space for three different income levels $y' < \hat{y} < y''$. The indifference curves represent all (t, p) combinations that households consider as good as community 1’s (p_1, t_1) -pair. Each household prefers pairs south-west of its indifference curve. It follows from property 1 that the indifference curves decrease in

the (t, p) -space and from property 2 that they become flatter as income rises. Imagine now a community 2, characterized by $t_2 < t_1$ and $p_2 > p_1$, where household \hat{y} is indifferent to 1. All poorer households, e.g. y' , prefer 1 to 2 and all richer households, e.g. y'' , prefer 2 to 1. \square

3.3 Equilibrium

An equilibrium requires that each household is located in the community that maximizes its utility, that local housing markets clear, that each community has a balanced public budget and that each community has a positive population.

There always exists a *symmetric* equilibrium in which both communities have identical characteristics $(p_1, t_1) = (p_2, t_2)$ and in which the households allocate themselves such that all communities show the same income distribution.¹² The following proposition shows that an *asymmetric* equilibrium, i.e. an equilibrium in which $(p_1, t_1) \neq (p_2, t_2)$, can exist as well. In addition, we assume:¹³

- (i) Housing supply $HS(L_i, p_i)$ is positive at price zero: $HS(L_i, 0) = \underline{L}_i > 0$ for $i = 1, 2$.
- (ii) There is a minimum income $\underline{y} > g$.
- (iii) There is no upper bound on utility from housing consumption: $h_i \rightarrow \infty, b_i > 0, h_j < \infty$ and $b_j < \infty$ implies $U(h_i, b_i) > U(h_j, b_j)$.

Proposition 3 (Existence of an asymmetric equilibrium)

Given the additional assumptions (i) to (iii), there exists an equilibrium in which the communities 1 and 2 exhibit different characteristics, i.e. $(p_1, t_1) \neq (p_2, t_2)$.

The proof is given in appendix 1.

3.4 Adding Heterogeneous Preferences

So far, we have assumed that households differ in their incomes y only. This has led to perfect income sorting, which is a very unrealistic prediction. In this section, we extend the model by assuming that the households' preferences are also heterogeneous, which will lead to imperfect income sorting.

The household preferences are now represented by the utility function $U(h, b; \alpha)$, where the parameter α describes the taste for housing. The higher α , the more a household is, ceteris

¹²Other equilibria in which $(p_1, t_1) = (p_2, t_2)$ may exist as well.

¹³As it will become evident in section 4, these assumptions are sufficient, but not necessary for the existence of an asymmetric equilibrium.

paribus, willing to spend on housing. Hence, the housing demand increases in α , i.e.

$$\frac{\partial h^*}{\partial \alpha} = \frac{\partial h(t, p; y, \alpha)}{\partial \alpha} > 0 \text{ for all } t, p, y \text{ and } \alpha. \quad (7)$$

We assume that h^* is continuous in α , and that preferences are distributed independently of income with density function $f(\alpha) > 0$ for $\alpha \in [\underline{\alpha}, \bar{\alpha}]$.

It follows that perfect income segregation still holds, but only within the subpopulation of households with identical preferences α and if both communities are populated by households with preferences α . In addition to segregation by incomes, there is segregation by preferences:

Lemma 2 (Boundary indifference with heterogeneous preferences)

Consider a subpopulation of households with identical income y . If a household with taste α' prefers to live in community i and a household with taste $\alpha'' > \alpha'$ in community j , then there exists a border household with taste $\hat{\alpha}(y)$, $\alpha' \leq \hat{\alpha}(y) \leq \alpha''$ who is indifferent between the two communities.

The proof is analogous to the proof of lemma 1. The continuity of $V(p_i, t_i, y, \alpha) - V(p_j, t_j, y, \alpha)$ in α follows from the continuity of h^* in α .

Proposition 4 (Preference segregation)

If both communities are populated by households with income y and if $(p_1, t_1) \neq (p_2, t_2)$ with $t_1 > t_2$, then among the subpopulation of households with identical income y any household in community 1 has a stronger taste for housing than the border household and any household in community 2 has a weaker taste for housing than the border household.

The proof is analogous to the proof of proposition 2 using $\partial M / \partial \alpha = -(\partial h^* / \partial \alpha) / [y r(y)] < 0$ as counterpart to property 2.

Corollary 1

The locus of border households $\hat{\alpha}(y)$ is strictly increasing in y .

Proof: Suppose $\hat{\alpha}(y)$ is non-increasing at some border household $(\tilde{y}, \tilde{\alpha} = \hat{\alpha}(\tilde{y}))$. Then there exists a border household with $y' > \tilde{y}$ and $\alpha' = \hat{\alpha}(y') \leq \tilde{\alpha}$. Proposition 4 implies that all households with y' and $\alpha \geq \alpha'$ weakly prefer community 1. The household with y' and $\tilde{\alpha}$ therefore weakly prefers community 1. However, proposition 2 implies that all households with $\tilde{\alpha}$ and $y > \tilde{y}$ including $(y', \tilde{\alpha})$ strictly prefer community 2. This is a contradiction. \square

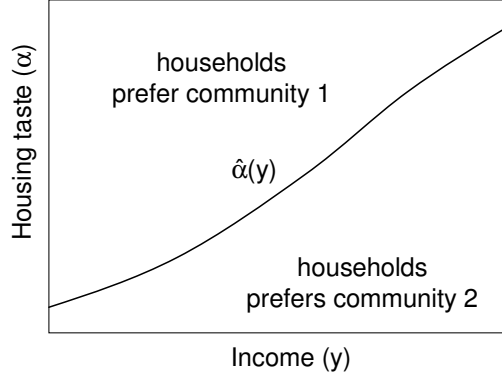


Figure 2: Simultaneous income and preference segregation.

Simultaneous heterogeneity in incomes and tastes leads to a more realistic pattern of household segregation. Figure 2 shows the locus of households which are indifferent between communities 1 and 2. Community 1 with a high tax multiplier t_1 and low housing prices p_1 is populated by households which are relatively poor or which have a relatively strong taste for housing. Community 2 with $t_2 < t_1$ and $p_2 > p_1$ is populated by households which are relatively rich or which have a relatively weak taste for housing. Hence, even though income segregation is no longer perfect, income groups still tend to gather.

Proposition 4 and Lemma 2 allow to calculate the probability that a household with income y lives in community 1,

$$P(1|y) = \int_{\hat{\alpha}(y)}^{\infty} f(\alpha) d\alpha = 1 - F(\hat{\alpha}(y)), \quad (8)$$

where $F(\cdot)$ denotes the cdf of the preference parameter α . The following lemma shows that this probability decreases in y :

Lemma 3 (Population shares by income)

If both communities are populated and if $(p_1, t_1) \neq (p_2, t_2)$ with $t_1 > t_2$, then the probability that a household with income y lives in community 1 is weakly decreasing in y for all y and strictly decreasing in y for some y .

Proof: Equation (8) implies $\partial P(1|y)/\partial y = -\partial F(\hat{\alpha}(y))/\partial y = -\partial F(\hat{\alpha}(y))/\partial \hat{\alpha}(y) \cdot \partial \hat{\alpha}(y)/\partial y$, and corollary 1 implies $\partial \hat{\alpha}(y)/\partial y > 0$. Since $f(\alpha) \geq 0$ for all α , it holds for all y that $\partial F(\hat{\alpha}(y))/\partial \hat{\alpha}(y) \geq 0$ and, consequently, that $\partial P(1|y)/\partial y \leq 0$. The border household $\hat{\alpha}(y)$ exists, i.e. $f(\hat{\alpha}(y)) > 0$, at least for some y as otherwise one community would not be populated. Therefore $\partial F(\hat{\alpha}(y))/\partial \hat{\alpha}(y) > 0$ and $\partial P(1|y)/\partial y > 0$ for some y . \square

3.5 Tax Progression

In this section, we derive how the described equilibrium sorting of the population affects liability progression of the tax schedule.

Recall that the average tax rate a household faces depends on the tax multiplier t_i of the community he locates in. We introduce the concept of the *mean average tax rate* to describe the expected average tax rate a household faces in equilibrium unconditional on his location

$$E[tr(y)|y] = \sum_{i=1,2} [P(i|y) t_i] r(y) = [P(1|y) t_1 + [1 - P(1|y)] t_2] r(y). \quad (9)$$

Hence, a household with income y pays in equilibrium on average total taxes equal to $E[T(y)|y] = E[tr(y)|y]y$. The liability progression of the mean tax schedule is thus

$$\rho^*(y) := \frac{\partial E[T(y)|y]}{\partial y} \frac{y}{E[T(y)|y]} = \frac{r(y) + y r'(y)}{r(y)} + \frac{[\partial P(1|y)/\partial y] (t_1 - t_2) y}{P(1|y) t_1 + [1 - P(1|y)] t_2}. \quad (10)$$

Proposition 5 (Tax progression in equilibrium)

If both communities are populated and if $(p_1, t_1) \neq (p_2, t_2)$ with $t_1 > t_2$, then the liability progression of the mean tax schedule is weakly lower than the liability progression of the tax schedule in each community for all incomes y and strictly lower for some incomes y .

Proof: The first addend of equation (10) equals the liability progression in any given community as defined in equation (1). The second addend is thus the difference between the liability progression of the mean tax schedule and the liability progression in any community. Since $t_1 > t_2$, this difference is weakly negative if $\partial P(1|y)/\partial y \leq 0$, and strictly negative if $\partial P(1|y)/\partial y < 0$. Lemma 3 completes the proof. \square

4 A Specified Version of the Model

We investigate the qualitative and quantitative properties of the model in a fully specified example in this section. The specification is kept as simple as possible but still captures all mechanisms of the model. The example is calibrated to the Zurich area, the largest Swiss metropolitan area.

The tax schedule is taken from Young (1990)

$$r(y) = r_0 \{1 - [1 + r_2 (y - d)^{r_1}]^{-1/r_1}\}$$

with parameters $r_0 > 0$, $r_1 > 0$ and $r_2 > 0$. In addition to Young, we also include a deductible $d > 0$. The average local tax rate $tr(y)$ and the local marginal tax rate $t[y_t \partial r(y)/\partial y + r(y)]$

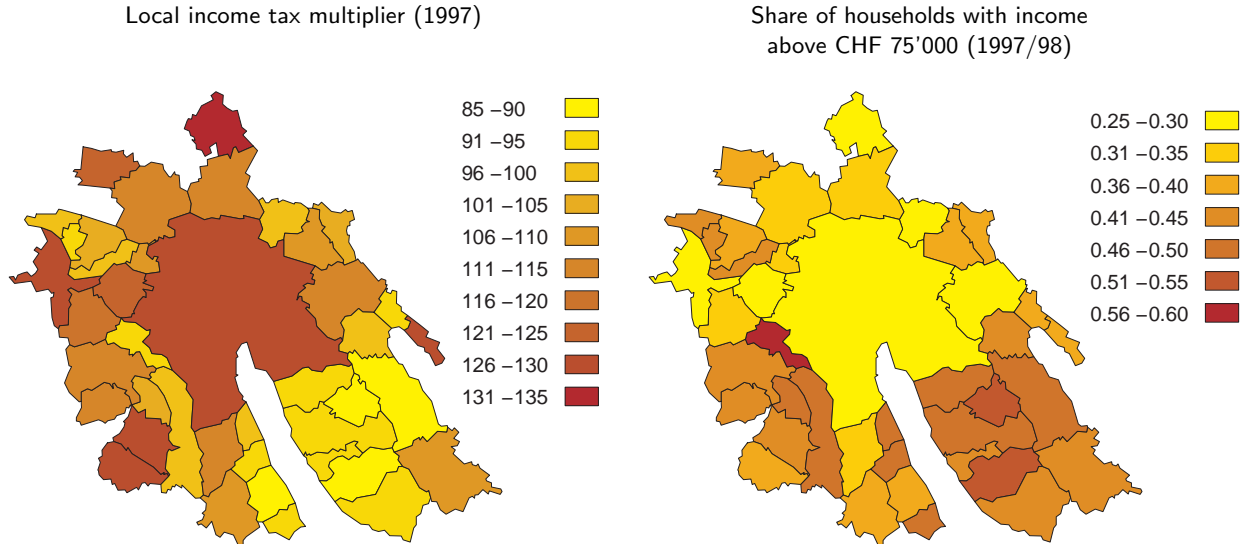


Figure 3: Taxes and incomes in the Zurich metropolitan area.

are increasing in income y . The marginal tax rate is above the average tax rate for all incomes exceeding the deductible d , and both asymptotically approach a maximum tr_0 .

Household preferences are described by a Cobb-Douglas utility function:

$$U = h^\alpha b^{1-\alpha},$$

where $0 < \alpha < 1$ stands for the taste parameter of the general model. Utility function and tax schedule satisfy properties 1 and 2.

We adopt the housing supply function

$$HS_i = L_i(p_i)^\theta$$

from Epple and Romer (1991).¹⁴

We calibrate the above outlined model to the Zurich metropolitan area, the largest metropolitan area in Switzerland. The city of Zurich has about 330 thousand inhabitants and is the capital of the canton (state) of Zurich. The canton of Zurich counts 1.2 million inhabitants in 171 individual communities. As described in section 2, each of these communities can choose its own tax multiplier. We restrict the analysis to the city of Zurich and a ring of the most integrated communities around the center.¹⁵ This agglomeration of 41 communities is modelled in a highly

¹⁴Epple and Romer derive this housing supply function from an explicit production function, where $0 \leq \theta \leq 1$ is the ratio of non-land to land input.

¹⁵This ring is formed by all communities in the canton of Zurich with more than 1/3 of the working population commuting to the center. The number of commuters to the city of Zurich and the size of the working population in the communities is based on the 1990 Census. This definition of the urban area is chosen to justify the model's assumption that households income is exogenous, i.e. that they choose their place of residence independent of where they work. It results in a set of communities closest to the central business district.

Table 1: Equilibrium values of the specified model.

	[1]	[2]	[3]	[4]	[5]	[6]
	model simulation			data metropolitan area of Zurich		
	high-tax	low-tax	unified	high-tax ^a	low-tax ^b	whole
L : area	0.50	0.50	1	0.496	0.504	1
p : rent	10.2	12.2	11.3	233 ^c , 204 ^d	228 ^c	232 ^c
t : tax shifter	2.11	0.63	0.86	1.19-1.31	0.85-1.15	0.85-1.31
n : inhabitants	0.38	0.62	1	0.650	0.350	1
Ey : mean income	43,829	84,039	68,656	61,700	76,553	68,656

The calibrated model parameters: $g = 2500$, $E(\ln y) = 11.0$, $SD(\ln y) = 0.517$, $y_{min} = 15300$, $y_{max} = 500000$, $E(\alpha) = 0.25$, $S.A.(\alpha) = 0.11$, $\theta = 3$, $r_0 = 0.2$, $r_1 = 0.5$ and $r_2 = 0.00065$.

^a 9 communities with highest taxes, ^b 32 communities with lowest taxes,

^c population weighted mean, ^d mean excluding the city of Zurich

stylized way as two distinct jurisdictions with equal land size. We refer to these two communities as the *low-tax* and the *high-tax community*, respectively.

The details of the calibration are described in the appendix 2 and summarized at the bottom of Table 1. Figure 3 shows the local tax multipliers (left map) and the spatial income distribution (right map) in the calibrated area.¹⁶

4.1 Simulated Equilibrium

The equilibrium values p_i and t_i in both communities satisfy equations (11) and (12) and guarantee that no households wants to move. As there is no closed form solution to this nonlinear system of four equations and four unknowns, we solve numerically for the equilibrium values of the model.¹⁷

Table 1 shows in columns 1 and 2 the equilibrium values of the simulated model with heterogeneous tastes. There are substantial differences in mean incomes, taxes and housing prices between the two communities: The high-tax community exhibits about three times higher taxes and about 15 % lower housing prices than the low-tax community. The left graph in Figure 4 shows the segregation pattern in the income-taste space. The population is imperfectly sorted by incomes. While most rich households are found in the low-tax community, some rich households with a strong taste for housing prefer the high-tax low-price community and poor households with a low taste for housing prefer the low-tax high-price community. The right graph in Fig-

¹⁶Data from the following sources: Housing prices: Wüest & Partner, Zürich. Tax rates: Statistisches Amt des Kantons Zürich, Steuerfüsse 1997. Income distribution: Swiss Federal Tax Administration. Considered are all communities where more than 1/3 of the working population is commuting to the center community.

¹⁷The aggregation of individual behavior requires double integrals over the community population. These integrals cannot be calculated analytically. We use Gauss-Legendre Quadrature with 40 nodes in each dimension to approximate the various double integrals. We numerically solve for the equilibrium values by minimizing the sum of squared deviations from the equilibrium conditions with the Gauss-Newton method.

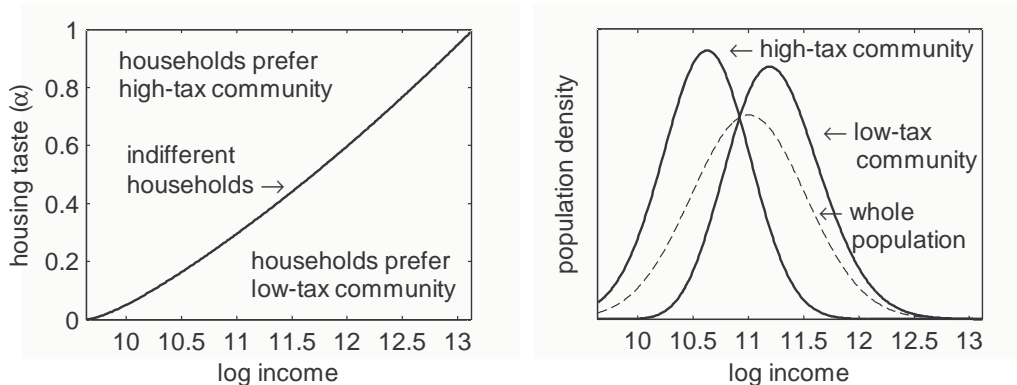


Figure 4: Income and taste segregation in equilibrium. The left Figure shows the preferred community for all household types. The right Figure shows the resulting income distributions in both communities.

Figure 4 shows the resulting income distributions in the two communities. The mean income in the high-tax community is now slightly more than half the one in the low-tax community. Column 3 in table 1 gives the equilibrium values for the hypothetical case that the two communities merged or harmonized their taxes.

The last three columns in table 1 report the actual housing rents, tax multipliers and mean income for three sets of communities: Column 4 shows the values for the 9 communities with highest taxes and column 5 for the 32 communities with lowest taxes. Note that these two groups have equal land area dedicated to development. Column 6 shows averages for the whole metro area. A comparison of our simulation results with the actual data shows that our simple two-community model correctly predicts that rich households tend to live in low-tax communities. The striking relationship between income taxation and spatial income distribution can also be seen by comparing the two maps in Figure 3. The predicted higher housing prices in the low-tax communities are supported by the data when the city of Zurich is excluded.¹⁸

The predicted magnitude of the differences is much higher than observed. This is not unexpected as our theoretical model abstracts from the horizontal and vertical tax equalization programs in Zurich (see Section 2). Imperfect tax equalization reduces the magnitude of tax and price differentials across communities. Note however, that the order of community characteristics found in proposition 1 remains unchanged. Imperfect tax equalization does not change the sorting mechanism in the model as long as (substantial) tax and price differentials remain.

¹⁸The city of Zurich is the central business district (CBD) where both housing prices and taxes are relatively high. This is most likely due to the center's intrinsic attractiveness (e.g. its cultural life, low commuting costs) which is not captured in our model. Multi-community models usually abstract from geography, i.e. physical distance, and exclude the CBD in the empirical test (see e.g. Epple and Sieg, 1999).

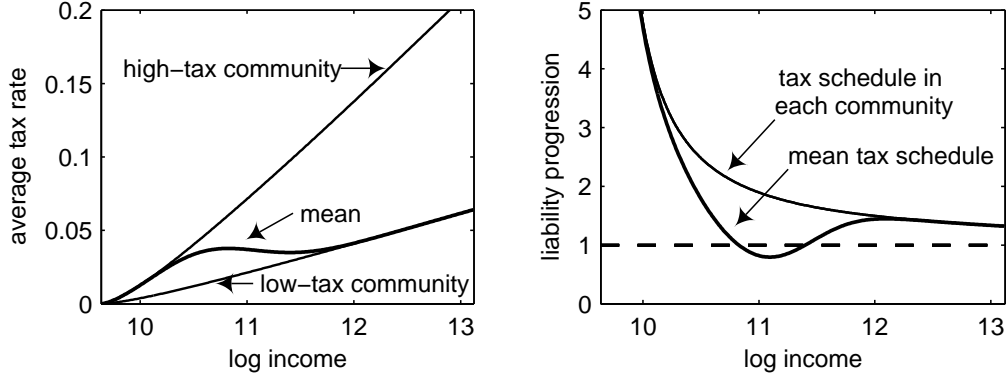


Figure 5: Average tax rate (left) and liability progression (right) by income.

4.2 The Resulting Tax Schedule

The average tax rate $t_i r(y)$ depends not only on the individual household's income but also on its place of residence. The left panel in Figure 5 shows the average tax rate in the high-tax and in the low-tax community of our specified model. The right panel in Figure 5 shows the liability progression of the tax schedules in each community. As shown before, liability progression is independent of the tax multiplier and therefore equal in both communities.

The place of residence is not random and rich households are more likely to reside in low-tax communities. The mean average tax rate $E[t_i r(y)|y]$ as defined in equation (9) describes the tax schedule that is realized after considering the sorting of the population. In other words, it gives the tax rate a household with income y pays on average. The left panel in Figure 5 also shows the mean average tax rate. For very poor households the mean tax schedule follows the one in the high-tax community while for very rich households it follows the one in the low-tax community. Middle income households face average tax rates in between. The resulting liability progression is shown in the right panel in Figure 5. The mean tax schedule follows the liability progression of each community's tax schedule for very poor and very rich household, while it is substantially lower for middle income households. For some middle incomes, liability progression is even below 1, indicating that the mean tax schedule is regressive.

5 Evidence

In this section, we *estimate* the mean average tax rates that households with a given income face in the Zurich metropolitan area as well as the liability progression of the mean tax schedule. We then compare our estimates to the results obtained in the previous section.

5.1 Method

In principle, the mean average tax rate can be estimated from a random sample of households in the studied area. Knowing each households' income and community tax rate allows to directly estimate the mean average tax rate with e.g. a kernel regression. The random sampling automatically accounts for the sorting of the population by incomes. Unfortunately, we do not have such microdata with tax information. Furthermore, available survey data suffers from small sample sizes and stratified sampling over communities.

We therefore follow an alternative estimation strategy. The mean average tax rate of a household with income y can be estimated from equation (9):

$$\hat{E}[tr(y)|y] = \sum_i [\hat{P}(i|y) \cdot t_i r(y)]$$

As the canton sets the tax schedule $r(y)$ and the individual communities their tax multipliers t_i , the average tax rate $t_i r(y)$ for any income y in any community i is known.

The estimated probability that a household with income y lives in community i can be expressed as function of the overall income distribution $f(y)$, the income distribution in each community $f(y|i)$ and the population share of each community $P(i)$:

$$\hat{P}(i|y) = \frac{\hat{f}(y|i)\hat{P}(i)}{\hat{f}(y)} = \frac{\hat{f}(y|i)n_i}{\sum_j [\hat{f}(y|j)n_j]},$$

where n_i is the known number of households living in community i .

It remains, therefore, estimating the income density $\hat{f}(y|i)$ of each community i in the area. We estimate $\hat{f}(y|i)$ from publicly available local income distribution data. The federal tax administration publishes the number of households with taxable income in seven different income classes.¹⁹ We assume that incomes are log-normally distributed and estimate mean and variance of this distribution using maximum likelihood.²⁰ We estimate a truncated log-normal distribution as the first reported income interval is empty for technical reasons. The log likelihood function for any community i is

$$\log \mathcal{L}_i = \sum_{k=1}^6 s_k \cdot \log \left[\frac{\Phi\left(\frac{c_{k+1}-\mu_i}{\sigma_i}\right) - \Phi\left(\frac{c_k-\mu_i}{\sigma_i}\right)}{1 - \Phi\left(\frac{c_1-\mu_i}{\sigma_i}\right)} \right],$$

where μ_i and σ_i^2 are mean and variance of log income in community i . s_k is the number of households in income class k with lower interval limit $c_k \in \{\log(15000), \log(20000), \log(30000)\}$,

¹⁹Swiss Federal Tax Administration, Steuerbelastung in der Schweiz, Natürliche Personen nach Gemeinden 1997, Neuchâtel: Swiss Federal Statistical Office.

²⁰Note that this maximum likelihood estimator corresponds to an ordered probit with known thresholds.

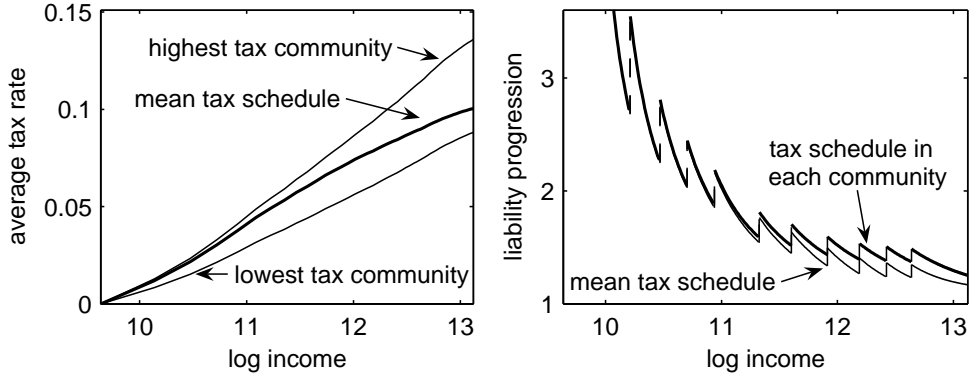


Figure 6: Estimated average tax rate (left) and liability progression (right) by income.

$\log(40000), \log(50000), \log(75000), \infty\}$. $\Phi(\cdot)$ is the cdf of the standard normal distribution. The income density in community i is then estimated as

$$\hat{f}(y|i) = \frac{1}{\hat{\sigma}_i y \sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{\log(y) - \hat{\mu}_i}{\hat{\sigma}_i} \right)^2 \right].$$

5.2 Results and Discussion

The left panel in Figure 6 shows the mean average tax rate in the Zurich metropolitan area. The top line is the average tax rate $t_i r(y)$ of households living in the community with the highest tax multiplier t_i . The bottom line is the average tax rate of households in the community with the lowest tax multiplier t_i . The middle line is the estimated mean average tax rate that households with income y face in this area, $\hat{E}[tr(y)|y]$. As one can see, the mean average tax rate is for poor households almost the one of the highest-tax communities, which include the city of Zurich.²¹ This is, of course, because most poor households live in high-tax communities. As households become richer, they are more likely to live in low-tax communities. The mean average tax rate for rich households is therefore closer to the one in lowest-tax communities.

Liability progression of the tax schedule is shown in the right panel in Figure 6. Note that the canton of Zurich sets constant marginal tax rates within 12 income intervals. At the interval limits, the average tax rate therefore shows kinks and the liability progression shows discontinuities. Liability progression is the same in each given community. Considering the estimated income sorting, liability progression of the mean tax schedule is lower than the one of each community. This applies in particular for households with high incomes y .

The results from the estimation (Figure 6) are qualitatively similar to the predictions from the simulation of our model (Figure 5). There are though some noteworthy differences: First,

²¹The tax multiplier is 131 in the highest-tax community and 130 in the city of Zurich.

the difference between the highest and the lowest tax multipliers is in reality smaller than our simulation predicts. A major reason is certainly the horizontal and vertical tax equalization programs in the canton of Zurich, which lower the tax differentials across communities (see Section 2). Moreover, public expenditures are not exactly the same in all communities (as our model assumes). They are on average 10 % higher in the communities that constitute the low-tax community in our simulation than in the communities constituting the high-tax community.

Second, the mean average tax rates for very rich households remain in reality above the average tax rates in the lowest-tax community while they approach the average tax rates in the low-tax community in our simulation. The reason for this difference is that in reality not all very rich households choose to live in the lowest tax communities. In particular, many high income households choose to live in the high-tax city of Zurich. We think the main reason is that the location choice also depends on preferences beyond the taste for housing. For example, residents in the city of Zurich might also value its cultural life or low commuting costs.

Third, the difference between the liability progression of the mean tax schedule and the liability progression of each community's tax schedule peaks in the estimation at a higher income than in our simulation. The reason is that the share of rich households living in the high-tax community is very small in our simulation while it is substantial and thus more sensitive to changes in income in the estimation.

6 Conclusions

This paper studies the tension between fiscal decentralization and progressive taxation. We present a multi-community model in which the local income tax rate depends on an exogenous progressive tax schedule and a tax multiplier that can differ across communities. The progression of the tax schedule induces a self-sorting process that results in income sorting. This income sorting is however imperfect if households differ in their preferences for housing. But rich households are still more likely to locate themselves in communities with low tax multipliers than poor households. As a consequence, the mean tax schedule becomes less progressive than the exogenous tax schedule. Empirical evidence from the largest Swiss metropolitan area, the Zurich area, supports our predictions: Rich households are more likely to live in communities with low tax multipliers than poor households, and the mean tax schedule is thus less progressive than the tax schedule of the canton of Zurich.

These findings suggest, in line with the literature on fiscal federalism, that progressive taxes

should indeed be implemented at the state or national level if one wants them to unfold their full redistributive effect. But they also show that substantial redistribution through progressive income taxation is even possible at the community level.

Appendix 1: Proof of Proposition 3

We define $\Delta V(\hat{y}) := V_1(p_1(\hat{y}), t_1(\hat{y}), \hat{y}) - V_2(p_2(\hat{y}), t_2(\hat{y}), \hat{y})$, where $p_i(\hat{y})$ and $t_i(\hat{y})$ are the equilibrium housing price and the equilibrium tax multiplier, respectively, in i given that households with $y < \hat{y}$ live in 1 and households with $y > \hat{y}$ in 2. (Remember from proposition 2 that any household in 2 must be richer than any household in 1.) Hence, $V_i(p_i(\hat{y}), t_i(\hat{y}), \hat{y})$ is the indirect utility of a household with \hat{y} in i given this allocation of households.

We prove proposition 3 by showing (1) that $\Delta V(\hat{y})$ is continuous and (2) that $\Delta V(\hat{y}) > 0$ as $\hat{y} \rightarrow \underline{y}$ and that $\Delta V(\hat{y}) < 0$ as $\hat{y} \rightarrow \bar{y}$. It follows then from the intermediate value theorem that there is at least one \hat{y} , $\underline{y} < \hat{y} < \bar{y}$, such that $\Delta V(\hat{y}) = 0$. This implies - from the definition of ΔV - that households with \hat{y} are indifferent between the two communities, that p_1 and p_2 clear the housing markets, and that t_1 and t_2 balance the community budgets.

(1) In equilibrium, p_i is determined by $HS(L_i, p_i) = HD_i$. It follows from proposition 2 that

$$HD_i = \int_{\underline{y}_i}^{\bar{y}_i} h(p_i, t_i; y) f(y) dx, \quad (11)$$

where \underline{y}_i and \bar{y}_i are the highest and lowest incomes in i . The hereby implicitly defined p_i is continuous in \bar{y}_i and \underline{y}_i given continuity of $HS(\cdot)$, $h(\cdot)$ and $f(\cdot)$. The balanced budget requirement and proposition 2 imply

$$t_i = \frac{n_i g}{\int_{\underline{y}_i}^{\bar{y}_i} r(y) f(y) dx}, \quad (12)$$

where

$$n_i = \int_{\underline{y}_i}^{\bar{y}_i} f(y) dx. \quad (13)$$

Given continuity of $r(\cdot)$ and $f(\cdot)$, t_i is continuous in \bar{y}_i and \underline{y}_i . Since V_i is continuous in p_i , t_i and y and since p_i and t_i are continuous in y , $\Delta V(\hat{y})$ is continuous in \hat{y} .

(2) It follows from equations (11) and (13) that $HD_1 \rightarrow 0$ and $n_1 \rightarrow 0$ as $\hat{y} \rightarrow \underline{y}$. Since assumption (i) guarantees that $HS(L_1, 0) = \underline{L}_1 > 0$ (and since $\partial HS(L_i, p_i) / \partial p_i \geq 0$), it holds that $h^*(p_1, t_1; y) \rightarrow \infty$ and $p_1 \rightarrow 0$ as $\hat{y} \rightarrow \underline{y}$. Hence, $b^* \rightarrow \underline{y} - g > 0$, where the strict inequality follows from assumption (ii). Assumption (iii) then guarantees that $\Delta V(\hat{y}) > 0$ as $\hat{y} \rightarrow \underline{y}$. Analogously, it can be shown that $\Delta V(\hat{y}) < 0$ as $\hat{y} \rightarrow \bar{y}$.²² \square

Appendix 2: Calibration

Tax schedule: The parameters $r_0 = 0.2$, $r_1 = 0.5$ and $r_2 = 0.00065$ closely approximate the tax scheme of the canton of Zurich for a married couple. The deductible $d = 15,300$ is based on a

²²The only difference is that $b^* \rightarrow \bar{y} - g$, which exceeds $\underline{y} - g$.

family with one child.²³

Income Distribution: The income distribution is calibrated with data from the Swiss Federal Tax Administration. We use a log-normal distribution to approximate this right-skewed distribution. The estimation of the mean $E(\ln y) = 11.0$ and standard deviation $SD(\ln y) = 0.517$ from the observed income bins is described in the empirical section 5.1. For numerical tractability, the model distribution is truncated at a minimum income equal to the deductible $y_{min} = 15,300$ and a maximum income $y_{max} = 500,000$.

Taste Distribution: The distribution of the taste for housing is calibrated with data from the Swiss labor force survey.²⁴ The Swiss labor force survey contains the monthly housing expenditure of renters.²⁵ Note that the taste parameter α in the Cobb-Douglas utility function is the share of housing in a utility maximizing household. We therefore estimate each household taste parameter as $\alpha = (ph)/y_d$, where ph is observed households housing expenditure and y_d is observed household income minus federal, state and communal taxes. A beta distribution with mean $E(\alpha) = 0.25$ and standard deviation $SD(\alpha) = 0.11$ describes the distribution of the so calculated taste parameter well. Taste and income are assumed to be uncorrelated.

Housing and Public Good Production: The price elasticity of housing supply is $\theta = 3$ as in Epple and Romer (1991) and Goodspeed (1989). The targeted public goods provision $g = 2,500$ is such that the (population) weighted average tax multiplier in the calibrated model with heterogeneous tastes equals the observed one.

²³Tax scheme according to Steuergesetz vom 8. Juni 1997, Tarif a.

²⁴Swiss Federal Statistical Office, Schweizerische Arbeitskrafterhebung (SAKE) 1995.

²⁵Of course, there is a selection bias by only considering renters. This seems nevertheless justified because the proportion of renters is very high in Switzerland (65% in the data set used).

References

- [1] Boadway, Robin W. and Harry M. Kitchen (1980), *Canadian Tax Policy*, Toronto, Canada: Canadian Tax Foundation.
- [2] Ellickson, Brayn (1971), Jurisdictional Fragmentation and Residential Choice, *American Economic Review*, 61, 334-339.
- [3] Epple, Dennis and Glenn J. Platt (1998), Equilibrium and Local Redistribution in an Urban Economy when Households Differ in both Preferences and Income, *Journal of Urban Economics*, 43, 23-51.
- [4] Epple, Dennis and Holger Sieg (1999), Estimating Equilibrium Models of Local Jurisdictions, *Journal of Political Economy*, 107, 645-681.
- [5] Epple, Dennis and Thomas Romer (1991), Mobility and Redistribution, *Journal of Political Economy*, 99, 828-858.
- [6] Feldstein, Martin and Marian V. Wrobel (1998), Can State Taxes Redistribute Income?, *Journal of Public Economics*, 68, 369-396.
- [7] Goodspeed, Timothy J. (1989), A Re-Examination of the Use of Ability to Pay Taxes by Local Governments, *Journal of Public Economics*, 38, 319-342.
- [8] Hansen, Julia L., Formby, John P., and Smith, W. James (1998), Estimating the Income Elasticity of Demand for Housing: A Comparison of Traditional and Lorenz-Concentration Curve Methodologies, *Journal of Housing Economics*, 7, 328-342.
- [9] Hansen, Julia L., Formby, John P., and Smith, W. James (1996), The Income Elasticity of Demand for Housing: Evidence from Concentration Curves, *Journal of Urban Economics*, 39, 173-192.
- [10] Hardman, Anna and Yannis M. Ioannides (2004), Neighbors income distribution: economic segregation and mixing in US urban neighborhoods, *Journal of Housing Economics*, 13, 368-382.
- [11] Ioannides, Yannis M. (2004) Neighborhood income distributions, *Journal of Urban Economics*, 56, 435-457.
- [12] Mayo, Stephen K. (1981), Theory and Estimation in the Economics of Housing Demand, *Journal of Urban Economics*, 10, 95-116.
- [13] Musgrave, Richard M. (1971), Economics of Fiscal Federalism, *Nebraska Journal of Economics and Business*, 10, 3-13.
- [14] Musgrave, Richard M. and Tun Thin (1948), Income Tax Progression, 1929-48, *Journal of Political Economy*, 56, 498-514.
- [15] Oates, Wallace E. (1999), An Essay on Fiscal Federalism, *Journal of Economic Literature*, 37, 1120-1149.
- [16] Ross, Stephen and John Yinger (1999), Sorting and Voting: A Review of the Literature on Urban Public Finance. In: Paul Ceshire and Edwin S. Mills (eds.), *Handbook of Regional and Urban Economics*, Vol. 3, Amsterdam: North-Holland.
- [17] Schmidheiny, Kurt (2006a), Income Segregation and Local Progressive Taxation: Empirical Evidence from Switzerland, *Journal of Public Economics*, 90, 429-458.
- [18] Schmidheiny, Kurt (2006b), Income Segregation from Local Income Taxation When Households Differ in Both Preferences and Incomes, *Regional Science and Urban Economics*, 36, 270-299.
- [19] Westhoff, F. (1977), Existence of Equilibria in Economies with a Local Public Good, *Journal of Economic Theory*, 14, 84-112.
- [20] Young, Peyton H. (1990), Progressive Taxation and Equal Sacrifice, *American Economic Review*, 80, 253-266.