Income Segregation from Local Income Taxation: When Households Differ in Both Preferences and Incomes

Kurt Schmidheiny[‡§]
Department of Economics, Tufts University, Medford, USA
and CESifo, Munich, Germany

October 26, 2005
Forthcoming in *Regional Science and Urban Economics*

Abstract

This paper presents a model of an urban area with local income taxes used to finance a local public good. Households differ in both incomes and their taste for housing. The existence of a segregated equilibrium is shown in a calibrated two-community model assuming single-peaked distributions for both income and housing taste. The equilibrium features income segregation of the population across the communities. The segregation is, however, imperfect: some rich households can be found in poor communities and vice-versa. The calibrated model is able to explain the substantial differences in local income tax levels and average incomes across communities as observed in e.g. Switzerland. The numerical investigation reveals that the ordering of community characteristics critically depends on the substitutability between the public and the private good. The numerical investigation also suggests that taste heterogeneity reduces the distributional effects of local tax differences. The numerical investigation furthermore suggests that the rich community can set lower taxes when it is small.

Key Words: Income Segregation, Income Sorting, Fiscal Decentralization, Income Taxation, Local Public Goods

JEL-classification: H71, H73, R13

[‡]Kurt Schmidheiny, Department of Economics, Tufts University, Braker Hall, Medford, MA 02155. Tel.: +1 617 627-6394. Email: kurt.schmidheiny@tufts.edu.

[§]I would like to thank Veronika Aegerter, Frank Bodmer, Stephen Calabrese, Dennis Epple, Timothy Goodspeed, Yannis Ioannides, Anke Kessler, Klaus Neusser, and the participants of the brown-bag seminars in Bern, the 2003 annual meetings of the Swiss Society of Economics and Statistics in Bern and the Verein für Socialpolitik in Zürich. I am grateful to Wüest und Partner, and especially to Jürg Mosimann for the provision of the housing price data. Heather Murray helped me find the right words. Finally, I wish to thank the editor, Konrad Stahl, and two anonymous referees for improving the quality of this paper.

1 Introduction

Decentralized financing of local public goods is a natural counterpart of decentralized decision about their provision. As Oates (1972) argued, local units deciding upon public programs are more likely to trade off costs against benefits if these programs are financed by local taxes. Fiscal Federalism is intensively debated in the European Union. On the one hand there are attempts to coordinate fiscal policies across EU member states. On the other hand, increased regional self-government, as implied by the subsidiarity principle, calls for some regional fiscal autonomy.

I develop a model of an urban area with local income taxes used to finance a local public good. Households differ in both incomes and their taste for housing. The existence of an asymmetric equilibrium is shown in a calibrated two-community model assuming single-peaked distributions for both income and housing taste. The equilibrium features income segregation of the population across communities. This segregation is, however, imperfect: some rich households can be found in on average poor communities and some poor households can be found in on average rich communities. The calibrated model is able to explain the substantial differences in local income tax levels and average incomes across communities as observed in e.g. Switzerland. In accordance with the empirical findings, the rich community shows lower taxes and both higher housing prices and a higher public goods provision than the poor community. This order of community characteristics depends, however, on the preferences for the local public good. The above ordering of community characteristics holds for low degrees of substitutability between public and private goods. When the public good is easily substituted by private goods, the rich community exhibits higher housing prices and higher public goods provision as well as higher taxes. The numerical investigation also suggests that taste heterogeneity reduces the distributional effects of local tax differences. The numerical investigation furthermore suggests that the ability of the rich community to set low taxes is higher when it is physically small. However, a tax haven need not be small.

Following Tiebout's (1956) seminal work, there is a long tradition of modelling fiscal decentralization at community level. The consideration of heterogeneous household incomes by Ellickson (1971) and Westhoff (1977) moved the focus away from seeking optimal community size to the study of urban areas with given community borders. While this strand of research was followed by a large number of studies investigating local *property* taxation (surveyed in Ross and Yinger 1999, and Epple and Nechyba, 2004), there have been few contributions on local *income* taxation. Hansen and Kessler (2001a) elegantly study a local transfer financed by local income taxes in a model with inelastic housing demand and an exogeneous Laffer curve. Calabrese (2001) studies local income taxation in a model similar to Hansen and Kessler's but with price responsive housing demand. Konishi (1996) provides an existence proof for equilibria in models with income taxation under weak assumptions. He does not study the extent of income sorting in the established equilibria.

Multi-community models with agents that differ in income typically predict *perfect* segregation of the population by income, i.e. households of the same income group live in the same community. However, recent literature on spatial income sorting (Epple and Sieg, 1999, Hardman and Ioannides, 2004, Ioannides, 2004 and Bayer, McMillan and Rueben, 2004) forcefully demonstrate that the sorting is very imperfect. Rhode and Strumpf (2003) show a long term trend of decreasing income sorting despite falling costs of moving. Schmidheiny (forthcoming) shows significant but imperfect sorting among movers in a metropolitan area. This clear empirical finding is almost completely missing in the theoretical literature. A notable exception are Epple and Platt (1998) who study a model with property taxation and show that the introduction of (continuous) heterogeneous tastes for housing indeed predicts a more realistic incomplete segregation of the population. Kessler and Lülfesmann (2005) introduce two types of households with high and low taste for the public good in a model with local income taxation to establish equilibria with imperfect income sorting. Keely (2004) attributes imperfect income sorting to the role of housing developers. Hindriks (2001) creates income mixing by assuming that households have an intrinsic preference for either of two locations.

This paper follows Epple and Platt (1998) but introduces heterogeneous tastes in a multi-community model with local *income* taxation and a partly substitutable public good. A similar model has been investigated by Goodspeed (1989). This study generalizes Goodspeed's analysis both by introducing heterogeneous tastes and by using a realistic single-peaked distribution of the population. Not only does this single-peakedness capture a realistic feature of urban economies, but it also challenges the existence of equilibria in multi-community models with income taxation. The possible non-existence of segregated equilibria under a non-uniform income distribution is shown by Hansen and Kessler (2001b). My model can also be seen as a generalization of Kessler and Lülfesmann (2005) and consequently shares many of their findings. The main difference is the introduction of a housing market which allows me to establish equilibria without the assumption of an exogeneous Laffer curve of tax income (caused by e.g. losses from distortive taxation). Equilibrium housing prices provide a further mechanism for income sorting which is studied in this paper. Rather than assuming heterogenous taste for the public good I introduce heterogeneous taste for housing, for which an empirical counterpart can easily be found. Furthermore I introduce a continuum of housing tastes rather than just two types; this allows me to study the role of variance in preferences.

While local taxation of *property* is widespread, especially in the United States, local taxation of *income* is rarer. Local *income* taxation at municipal level is e.g. observed in four U.S. states (Indiana, Maryland, Ohio, and Pennsylvania), in Denmark, Belgium and Switzerland. The numerical simulations in this paper are based on a Swiss metropolitan

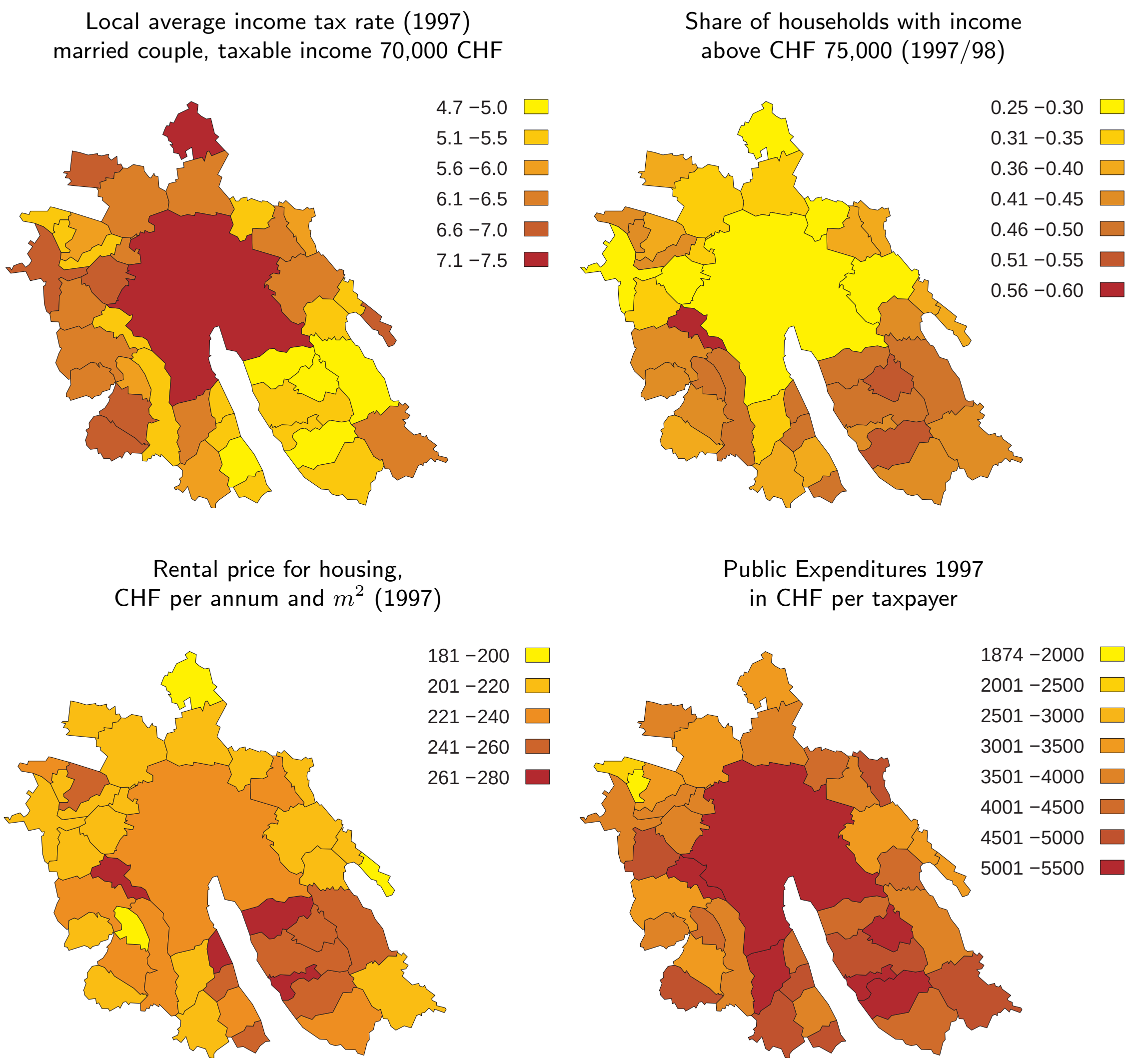

Figure 1: Community characteristics in the metropolitan area of Zurich.

area. Switzerland is an exemplary case of a federal fiscal system and offers a laboratory for the study of tax decentralization. Switzerland is a federation of 26 states, the so-called cantons. The cantons are divided into individual communities of varying size and population. The roughly 3000 communities form individual jurisdictions with great autonomy in terms of providing local public goods such as school services or infrastructure. These local public goods are financed mainly by local income taxes. While cantons autonomously organize the whole tax system, e.g. the degree of tax progression or the split between income and corporate taxes, the communities can generally only set a tax shifter in a given cantonal tax scheme. There is considerable variation in income taxes across Swiss communities. For example, for a married couple with a gross income of

70,000 Swiss francs (CHF) the totalled cantonal and community average income tax rate ranged from 4.7% (city of Zug) to 13.2 % (Lauterbrunnen, canton of Bern) in the year 2001. Within metropolitan areas the (community) tax differences are smaller but may still differ by a factor of 1.5 across the Zurich area for example. Figure 1 shows the substantial differences in local tax rates, income levels, housing prices and the provision of public goods in this community system.[1] The top-right map visualizes the considerable segregation by incomes in the Zurich area. The two top maps demonstrate a striking relationship between income taxation and spatial income distribution: the local share of rich households is almost an inverted picture of the local tax levels. It is particularly interesting to see whether multi-community models are able to explain the observed tax differentials. Note that municipality financing by local *income* taxation has been applied for more than 150 years in Switzerland. This system of more than 3000 municipalities thus empirically contradicts Nechyba's (1997) claim that "local income taxes play no empirically important role" and that their virtual inexistence in the United States proofs "property taxes a dominant tax strategy for local governments".

The paper is organized as follows: Section 2 introduces the formal model and derives the properties of the household utility function which induce segregation of the population. In the first part of Section 3 the calibration used for the numerical investigation of the equilibrium is described. In the middle part of Section 3 the numerical equilibrium is presented and the welfare implications of the decentralized decision making are discussed. The remainder of Section 3 discusses the robustness of the result to changes in the calibrated parameters. Section 4 draws conclusions.

2 The Model

The model economy is divided into J distinct communities. The area is populated by a continuum of heterogeneous households, which differ in both income $y \in [\underline{y}, \overline{y}]$, $0 < \underline{y}, \overline{y} < \infty$, and a parameter $\alpha \in [0,1]$ describing their taste for housing. Incomes and tastes are jointly distributed according to the density function $f(y, \alpha) > 0$. There are three goods in the economy: a composite private consumption b, housing h and a local public good g.[2] The latter is local in the sense that it is only consumed by the residents of a community.

A household can move costlessly and chooses the community in which its utility is maximized as place of residence. Each community indexed by j can individually set the amount of the local public good g_j and the local income tax rate $t_j \in [0,1]$. These decisions are made in a majority rule vote by the residents respecting budget balance in

[1] See data sources in section 3.1. Considered are all communities where more than 1/3 of the working population is commuting to the center community.

[2] See Section 2.4 for a discussion of the nature of the public good.

the community. Each community has a fixed amount of land L_j from which housing stock is produced. All households are renters and the housing stock is owned by an absentee landlord. The price for housing p_j in community j is determined in a competitive housing market. The private good is considered the numeraire. A community j is fully characterised by the triple (t_j, p_j, g_j). The set of all possible community characteristics is given by $\Gamma = [0, 1] \times \mathbb{R}^{++} \times \mathbb{R}^{+}$. Location choice, voting and the resulting community characteristics are simultaneously determined in the general equilibrium.

2.1 Households

The preferences of the households are described by a Stone-Geary utility function (Stone, 1954)

$$U(h, b, g; \alpha) := \alpha\, ln(h - \beta_h) + (1 - \alpha)\, ln(b - \beta_b) + \gamma\, ln(g - \beta_g)\,,$$

where h is the consumption of housing, b the consumption of the private good and g the consumption of the publicly provided good. $\beta_h > 0$, $\beta_b > 0$ and β_g are sometimes referred to as existential needs for housing, private good and public good, respectively. The parameter $\alpha \in [0, 1]$ describes the households' taste for housing, as will become apparent below.

Households face a budget constraint:

$$ph + b \leq y(1 - t)\,,$$

where p is the price of housing and t the local income tax. Note that the price of the private good is set to unity. Maximisation of the utility function with respect to h and b subject to the budget constraint yields the housing demand function

$$h^* := h(t, p, g; y, \alpha) = \frac{\alpha[y(1 - t) - p\beta_h - \beta_b]}{p} + \beta_h$$

and the demand for the the private good $b^* = y(1 - t) - ph^*$. Both demand functions are linear functions of after-tax income $y(1 - t)$, reflecting the fact that Stone-Geary utility implies a linear expenditure system (LES) and vice-versa (see e.g. Deaton and Muellbauer 1980). Housing demand is increasing in α as long as the household can satisfy its existential needs, i.e. $\partial(h)/\partial(\alpha) > 0$ iff $y(1-t) > p\beta_h + \beta_b > 0$. $\alpha = 0$ implies that the housing demand is equal to the existential needs and hence does not change with household income. $\alpha = 1$ denotes a household which spends all extra income on housing after paying his existential need.

The indirect utility

$$V(t, p, g; y, \alpha) := U(h^*, b^*, g; \alpha)$$

gives the utility of a household with income y and preference parameter α in a community with income tax t, housing prices p and a public good provision g.

2.2 Location Choice

Households take the community characteristics as given when they choose their place of residence. They value the different communities by their local tax rates t_j, housing prices p_j and public good provision g_j. A household chooses to locate in the community in which its utility is maximal. A household chooses j if and only if

$$V(t_j, p_j, g_j; y, \alpha) \geq V(t_i, p_i, g_i; y, \alpha) \quad \text{for all } i\,. \tag{1}$$

An important feature of this model is that households have to value the communities with respect to 3 dimensions (p, g, and t) of community characteristics. In previously studied multi-community models, communities are fully described by 2 dimensions only. In models with local property taxes (e.g. Epple and Platt, 1998), the households only care about after-tax housing prices $p(1-t)$ and not about p and t independently. Income tax models (e.g. Kessler and Lülfesmann, 2005) usually abstract from a housing market and fully characterize the communities by local taxes t and public goods or transfers g.[3] The 3-dimensional characteristics space allows for a much richer structure of potential equilibria. E.g. it is conceivable that a community with both higher taxes and higher housing prices than any other community can attract residents by offering more public goods.[4]

I will now show how the assumed household preferences lead to spatial segregation of the population. Before describing the allocation of households across communities in propositions 1 and 2, I explicitly state the specific properties of the assumed utility function that generate spatial segregation. Property 1 is trivial; Properties 2 and 3 are a direct consequence of the specified utility function assuming that existential needs are strictly satisfied, i.e. $y(1-t) > p\beta_h + \beta_b > 0$ and $g > \beta_g$. The derivation is provided in Appendix 1.

Property 1 (Relative preferences)
For all $(t, p, g, y, \alpha) \in \Gamma \times \mathbb{R}^+ \times [0, 1]$

$$M_{g,t}(t, p, g, y, \alpha) := \left.\frac{dg}{dt}\right|_{dV=0, dp=0} > 0\,,$$

$$M_{g,p}(t, p, g, y, \alpha) := \left.\frac{dg}{dp}\right|_{dV=0, dt=0} > 0\,,$$

$$M_{t,p}(t, p, g, y, \alpha) := \left.\frac{dt}{dp}\right|_{dV=0, dg=0} < 0\,.$$

[3]The reduction to a 2-dimensional characteristics space simplifies the analysis fundamentally. An exception is Goodspeed (1986, 1989) who studies the 3-dimensional characteristics space. However, Goodspeed seems not to detect the implied technical difficulties (see footnote 7).

[4]Examples of such equilibria are shown in section 3.6.

Property 1 signs the marginal rate of substitution $M_{\cdot,\cdot}$ between each pair of community characteristics. Property 1 states that a household can be compensated for a tax increase either by more public good provision or by lower housing prices. Westhoff (1977) calls this trade-off the *relative preference* for the public good. Property 1 also states that a household is compensated for higher housing prices by more public good provision. Property 1 holds under the standard assumption about the influence of prices, taxes and public goods on the household's well-being and is not specific to Stone-Geary utility.

Property 2 (Monotonicity of relative preferences)

(a) *For all* $(t, p, g, y) \in \Gamma \times \mathbb{R}^+$ *and any* $\alpha \in [0, 1]$,

$$\frac{\partial M_{g,t}}{\partial y} < 0 \quad \text{and} \quad \frac{\partial M_{g,p}}{\partial y} < 0\,.$$

(b) *For all* $(t, p, g, \alpha) \in \Gamma \times [0, 1]$ *and any* $y \in \mathbb{R}^+$,

$$\frac{\partial M_{g,p}}{\partial \alpha} > 0 \quad \text{and} \quad \frac{\partial M_{t,p}}{\partial \alpha} < 0\,.$$

Property 2 shows that the relative preference for community characteristics changes *monotonically* with both income and taste. Property 2a reveals that rich households need more public good compensation than poor households for the same increase in taxes or prices.[5] Property 2b states that households with a strong taste for housing need a smaller increase in public good provision or a larger decrease in tax rates to be compensated for a price increase than households with weak taste for housing. Property 2 explains why different households choose different locations in equilibrium. It is equivalent to the Spence-Mirrless condition in information economics. Property 2 is a consequence of the non-homothetic nature of Stone-Geary preferences. $\partial M_{g,p}/\partial y < 0$ is shared with Ellikson (1971), Westhoff (1977), Epple, Filimon and Romer (1984, 1993) and Goodspeed (1986, 1989). Property $\partial M_{g,t}/\partial y < 0$ is shared with Goodspeed (1986, 1989).[6] Property 2b introduces preference heterogeneity similar to Epple and Platt (1998).

Property 3 (Proportional shift of relative preferences)

(a) *For all* $(t, p, g, y) \in \Gamma \times \mathbb{R}^+$ *and any given* $\alpha \in [0, 1]$, *both*

$$\frac{\partial M_{g,t}}{\partial y} \Big/ \frac{\partial M_{g,p}}{\partial y} \quad \text{and} \quad \frac{\partial M_{t,g}}{\partial y} \Big/ \frac{\partial M_{t,p}}{\partial y}$$

[5] Note that $\partial M_{t,p}/\partial y$ cannot be signed as the marginal rate of substitution between tax and housing price, $M_{t,p}$, decreases with income if $\varepsilon > 1$ and increases if $\varepsilon < 1$.

[6] Goodspeed (1989) shows that $\partial M_{g,p}/\partial y < 0$ is equivalent to $\varepsilon_{g,y}/\varepsilon_{g,p} > \varepsilon_{h,y}$, where ε_{g,p_g} is the (shadow) price elasticity of demand for the public good and $\varepsilon_{h,y}$ is the income elasticity of demand for housing. Goodspeed (1989) also shows that $\partial M_{g,t}/\partial y < 0$ is equivalent to $\varepsilon_{g,y}/\varepsilon_{g,p} > 1$, where $\varepsilon_{g,y}$ is the income elasticity of public goods demand. He points to empirical evidence that shows that both assumptions are reasonable.

are independent of y, where $M_{t,g} = 1/M_{g,t}$.

(b) *For all $(t, p, g, \alpha) \in \Gamma \times [0, 1]$ and any given $y \in \mathbb{R}^+$,*

$$\frac{\partial M_{g,t}}{\partial \alpha} = 0\,.$$

The first part of Property 3a states that the ratio of the income effects on the $g - t$ trade-off and on the $g - p$ trade-off is independent of the income level. It is insightful to describe this property in terms of discrete changes. Consider an increase in local taxes by dt. The tax increase is compensated by $d^t g(y)$ of public goods. An increase of local housing prices by dp is compensated by another amount $d^p g(y)$ of public goods. Both compensations decrease with a (discrete) increase in income y (property 2a): $d^y d^t g := d^t g(y+dy) - d^t g(y) < 0$, $d^y d^p g := d^p g(y+dy) - d^p g(y) < 0$. The first part of Property 3a states now that the ratio $d^y d^t g / d^y d^p g$ of the income effects on the two compensations is constant w.r.t. to income. The second part of Property 3a is explained analogously. Property 3b states that the $g - t$ trade-off is independent of housing preferences α. Property 3b is a stronger version of Property 3a. Property 3 is another feature of the assumed Stone-Geary utility but considerably less intuitive than the first two properties. It results from both the linear expenditure system and the additive separability between g and (h, b). Property 3 is important as it rules out intractable segregation patterns where the middle income class prefers one community while rich *and* poor households prefer another community (see proof in Schmidheiny 2002).[7]

The distribution of the households across communities implied by Properties 2 and 3 is described in the following paragraphs. A first observation is that all households are indifferent between all communities when the communities have identical community characteristics, i.e. $(t_i, p_i) = (t_j, p_j)$ for all j, i. In this case the households settle such that communities have the same income distribution.

The following lemma and propositions describe how the utility maximizing households will be allocated across communities.

Lemma 1 (Boundary indifference)
Consider the subpopulation with taste α. If a household with income y' prefers to live in community j and another household with income $y'' > y'$ prefers to live in community i, then there is a 'border' household with income $\hat{y}_{ji}(\alpha)$, $y' \leq \hat{y}_{ji}(\alpha) \leq y''$, which is indifferent between the two communities.

Proof: The household with income y' prefers j to i, hence $V_j(y') - V_i(y') \geq 0$, where $V_j(y) := V(t_j, p_j, g_j; y, \alpha)$. The opposite is true for a household with income y'', thus

[7] It seems difficult to justify either Property 3a or Property 3b empirically. Goodspeed seems to derive perfect income segregation without Property 3 in the same setting. However, the graphical proof he provides in Goodspeed (1986) is incomplete. Goodspeed (1989) uses the Stone-Geary utility function for numerical simulations and fails to detect the missing assumption.

$V_j(y'') - V_i(y'') \leq 0$. $V_j(y) - V_i(y)$ is continuous in y as V is continuous in y. The intermediate value theorem implies that there is at least one $\hat{y}_{ji}$ between y' and y'' s.t. $V_j(\hat{y}_{ji}) - V_i(\hat{y}_{ji}) = 0$. The existence of $\hat{y}_{ji}$ follows from $f(y, \alpha) > 0$. □

The set of 'border' households is described by the function $\hat{y}_{ji}(\alpha)$. Equivalently, the set of border households is described by the inverse function $\hat{\alpha}_{ji}(y)$, implicitly defined by $V_j(\hat{\alpha}_{ji}(y)) = V_i(y, \hat{\alpha}_{ji}(y))$.

Definition 1 (Conditional income segregation)
An allocation of households is called conditionally segregated by incomes if the J sets $I_j = \{y :$ household with income y and taste α prefers community $j\}$ satisfy

- *I_j is an interval for all j,*
- *$I_j \cap I_i = \emptyset$ for all $i \neq j$,*
- *$I_1 \cup ... \cup I_J = [\underline{y}, \overline{y}]$*

for any α and for any j: $I_j \neq \emptyset$ for at least one α.

Definition 1 means that in a subpopulation with equal tastes, any community is populated by a single and distinct income class.

Proposition 1 (Conditional income segregation)
When the household preferences are described by a Stone-Geary utility function and all J communities exhibit distinct characteristics, $(t_j, p_j, g_j) \neq (t_i, p_i, g_i)$ for all $i \neq j$, then the allocation of households is conditionally segregated by incomes.

Proof: The proof uses the fact that the utility difference $V_j - V_i = V(t_j, p_j, g_j; y, \alpha) - V(t_i, p_i, g_i; y, \alpha)$ between community j and i is strictly monotonic in y (see Appendix 1):

$$sign\,\frac{\partial(V_j - V_i)}{\partial y} = sign\left(\frac{p_j\beta_h + \beta_b}{1 - t_j} - \frac{p_i\beta_h + \beta_b}{1 - t_i}\right).$$

Consider three households with income $y' < y'' < y'''$ respectively and suppose that the allocation of households does not satisfy Definition 1: y' as well as y''' prefer community j, but y'' strictly prefers community i. Given the location preference of y' and y''', it follows from Lemma 1 that there is an indifferent household $\hat{y}$, $y' \leq \hat{y} < y''$. The above sign condition implies that all households richer than $\hat{y}$, i.e. y''', also prefer i, which is a contradiction. □

Schmidheiny (2002) shows that the Properties 1, 2a to 3a are sufficient conditions for income segregation.

Definition 2 (Conditional taste segregation)
An allocation of households is called conditionally segregated by tastes if the J sets $I_j = \{\alpha :$ household with income y and taste α prefers community $j\}$ satisfy

- *I_j is an interval for all j,*
- *$I_j \cap I_j = \emptyset$, for all $i \neq j$,*
- *$I_1 \cup ... \cup I_J = [0, 1]$*

for any y and for any j: $I_j \neq \emptyset$ for at least one y.

Definition 2 means that in a subpopulation with equal incomes, any community is populated by a single and distinct interval of tastes.

Proposition 2 (Conditional taste segregation)
When the household preferences are described by a Stone-Geary utility function and all J communities exhibit distinct characteristics, $(t_j, p_j, g_j) \neq (t_i, p_i, g_i)$ for all $i \neq j$, then the allocation of households is conditionally segregated by tastes. Households in communities with lower housing prices have stronger tastes for housing than households in communities with higher housing prices.

Proof: The proof of the first sentence is analogous to Proposition 1 using the sign condition $sign(\partial(dV_j - dV_i)/\partial\alpha) = sign(p_i - p_j)$ (derived in Appendix 1). Second sentence: Consider $p_i < p_j$ and a household $(\hat{y}, \hat{\alpha})$ which is indifferent between the two communities j and i, hence $V_j(\hat{y}, \hat{\alpha}) = V_i(\hat{y}, \hat{\alpha})$. Then any household with the same income y and taste parameter $\alpha > \hat{\alpha}$ prefers community i, i.e. $V_j(\hat{y}, \hat{\alpha}) < V_i(\hat{y}, \hat{\alpha})$, since $\partial(dV_j - dV_i)/\partial\alpha < 0$ if $p_i < p_j$. □

Propositions 1 and 2 offer two ways of calculating a community's population:

$$n_j = \int_0^1 \int_{\underline{y}_j(\alpha)}^{\overline{y}_j(\alpha)} f(y, \alpha)\, dy\, d\alpha = \int_{\underline{y}}^{\overline{y}} \int_{\underline{\alpha}_j(y)}^{\overline{\alpha}_j(y)} f(y, \alpha)\, d\alpha\, dy,$$

where $\underline{y}_j(\alpha)$ and $\overline{y}_j(\alpha)$ are the lowest and highest income in community j given the subpopulation with taste α. $\underline{y}_j(\alpha)$ is given by the locus of indifferent households $\hat{y}_{ji}$ between community j and its 'adjacent' community i with lower income households. The other boundaries $\overline{y}_j(\alpha)$, $\underline{\alpha}_j(y)$ and $\overline{y}_j(\alpha)$ are given analogously. Closed form expressions for these boundaries are given in Appendix 1. Note that the adjacent community might not be the same for all subpopulations. This is demonstrated in Figure 2 showing four examples of possible segregation patterns in the case of three populated communities.

2.3 Housing Market

Within each community housing is produced from land and non-land factors. The housing supply in each community j is assumed to be an increasing function of the housing price p_j and the land dedicated to housing L_j. The housing supply function

$$HS_j = L_j \cdot p_j^{\theta}$$

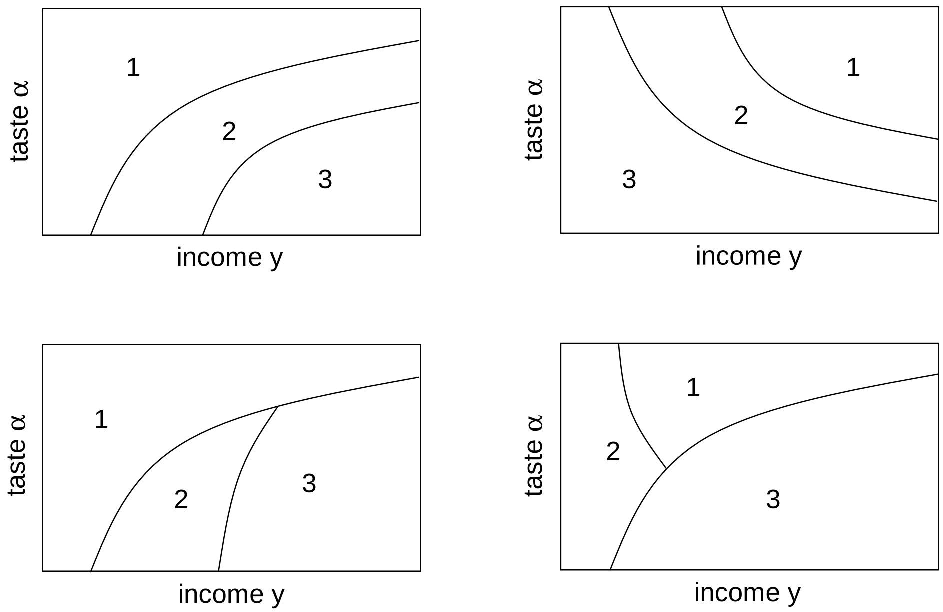

Figure 2: Examples of segregation patterns in the three-community case. The areas denoted by '1', '2' and '3' show the attributes of the households that prefer community 1, 2 or 3 respectively.

is adopted from Epple and Romer (1991), who derived it from an explicit housing production function; θ is the ratio of non-land to land input.

The aggregate housing demand in community j is

$$HD_j = \int_0^1 \int_{\underline{y}_j(\alpha)}^{\overline{y}_j(\alpha)} h(t_j, p_j, g_j; y, \alpha) f(y, \alpha)\, dy\, d\alpha.$$

In equilibrium, the price for housing in community j clears the housing market

$$HD_j = HS_j. \tag{2}$$

Definition 3 (Equilibrium Income Elasticity of Housing Price)
The elasticity of equilibrium housing prices w.r.t. to aggregate disposable income in a community j with given population is

$$\frac{dp_j}{d[(1-t_j)Ey_i]} \frac{(1-t_j)Ey_i}{p_j}\Big|_{HD_j=HS_j} = \frac{dp_j}{d(1-t_j)} \frac{1-t_j}{p_j}\Big|_{HD_j=HS_j}.$$

Definition 3 defines the reaction of equilibrium housing prices to changes in disposable income of the population and hence in the tax rate. Note that the reaction of the housing price depends not only on the housing supply function but also on the characteristics, i.e. tastes and incomes, of the local population.

2.4 Public Sector

Community j sets the amount of a local public good. It is public in the sense that it is publicly provided and that all residents consume the same amount of the good. The cost of providing this good is an increasing function of the amount provided g_j and the number of inhabitants n_j in the community. For simplicity, I assume:

$$C(g_j, n_j) = c_0 + c_1\, g_j\, n_j\,,$$

where $c_0 \geq 0$ and $c_1 > 0$. Note that there are no spillovers in the production of the good across communities. The increasing cost in the number of beneficiaries means that the good is not a pure public good since there is rivalry in consumption. One can think of e.g. schools, street construction and maintenance, city planning activities, etc. A positive constant c_0 implies increasing returns to scale in the production of the public good.

The community finances the publicly provided good by a proportional income tax. The tax revenue is

$$T_j = \int_0^1 \int_{\underline{y}_j(\alpha)}^{\overline{y}_j(\alpha)} t_j\, y\, f(y,\alpha)\, dy\, d\alpha = n_j\, t_j\, Ey_j\,,$$

where Ey_j is the mean income in community j. In equilibrium, the community's budget is balanced:

$$C(g_j, n_j) = T_j\,. \tag{3}$$

The tax rate and the amount of public goods are determined in a majority rule vote by the residents of the community. At this stage, households take the population of the community as given. Epple and Romer (1991) call voters with this behavior *myopic*, as they ignore the migration consequences from the political outcome in their community.[8]

Definition 4 (Public choice frontier)
The public choice frontier PCF_j in community j is the set of (p_j, g_j, t_j) triples, where the pair (g_j, t_j) satisfies budget balance (Eq. 3) and p_j clears the housing market (Eq. 2), given the housing demand with tax rate t_j.

[8] Voter myopia is assumed in Westhoff (1977), Rose-Ackerman (1979), Epple and Romer (1991) and Fernandez and Rogerson (1996). Fernandez and Rogerson (1996) rationalize 'myopicness' by modelling the households' decisions as a two-stage process in which households first choose their place of residence and then choose their consumption bundle and political behavior in the chosen community. Assuming perfect foresight, the political outcome in the second stage is anticipated when choosing the location. In Epple and Platt (1998) voters take the migrational effects of their voting decision into account while holding other locations' political decision constant. Kessler and Lülfesmann (2005) allow the households to relocate after voting. I was not able to find numerical equilibria without assuming myopic voters.

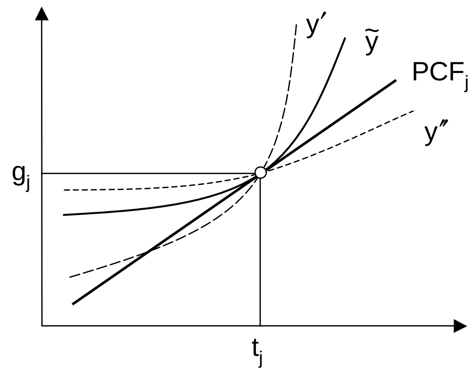

Figure 3: Voters' indifference curves in the (t, g) space.

Proposition 3 (Segregation of voters)
Consider the subpopulation of households with taste α in community j and assume that the income elasticity of the housing price (Def. 3) is below 1 for all (p_j, t_j) on the PCF_j. If a household $\tilde{y}_j(\alpha)$ prefers the triple (p_j, g_j, t_j) on the PCF_j to all other triples on the PCF_j, then any richer (poorer) household opposes a reduction (increase) in taxes.

Proof: The proof refers to Figure 3. Consider the indifference curves of three voters with household income $y' < \tilde{y} < y''$ respectively, given the same taste parameter α. These indifference curves take into account the reaction of the housing prices to a change in the income tax rates. The straight line is the PCF in community j. One can verify in the figure that the pivotal voter $\tilde{y}$ prefers the pair (g_j, t_j) to all other combinations on the PCF. It is shown in Appendix 1 that the indifference curve is monotonically increasing in t and that its derivative w.r.t. t is decreasing in y if the equilibrium income elasticity of the housing price (Def. 3) is below 1. Therefore, all richer voters, e.g. y'', dislike all $(g, t) \in PCF_j$ combinations with taxes lower than t_j, while all poorer voters, e.g. y', dislike higher taxes. $\Box$

$\tilde{y}_j(\alpha)$ is called the locus of pivotal voters. It is a decreasing function in α, as the price reduction induced by higher taxes is more appreciated by households with a stronger taste for housing. Note that from the perspective of a naïve voter who ignores the housing market, Proposition 3 holds without the additional assumption of the housing market tightness.

Definition 5 (Majority rule voting equilibrium)
A triple (p_j, g_j, t_j) on community j's PCF is called a majority rule voting equilibrium when no other triple on the PCF is strictly preferred by a majority of the community's residents.

As an implication of Proposition 3, a majority rule voting equilibrium in community j is established when

$$\int_0^1 \int_{\underline{y}_j(\alpha)}^{Min(\tilde{y}_j(\alpha),\overline{y}_j(\alpha))} f(y,\alpha)\,dy\,d\alpha = \frac{1}{2}\int_0^1 \int_{\underline{y}_j(\alpha)}^{\overline{y}_j(\alpha)} f(y,\alpha)\,dy\,d\alpha \tag{4}$$

and if the equilibrium income elasticity of the housing price (Def. 3) is below 1.

2.5 Equilibrium

The overall equilibrium of the multi-community model is a situation in which the location choice and the political equilibrium are consistent, i.e. no household has an incentive to move, local taxes and public good provision is the outcome of a majority rule vote by the local residents and the local housing markets are in equilibrium.

Definition 6 (Equilibrium)
A set of community characteristics (p_j, g_j, t_j), $j = 1, ..., J$, *and an allocation of individual households across communities is an equilibrium if and only if*

- *all households choose their community to maximise their utility,*
- *the housing market clears in all communities,*
- *there is a majority rule voting equilibrium in all communities.*

Note that there is always a *symmetric* equilibrium in which all communities show identical characteristics and the local income and taste distribution of households is a replication of the universe when $c_0 = 0$ and housing supply is homogeneous of degree one in land area. However, symmetric equilibria may not be stable.[9] The focus of this paper is on the empirically interesting case of asymmetric equilibria where all communities exhibit distinct characteristics.

Existence of *asymmetric* equilibrium is proved by Goodspeed (1986) in a model with income taxes, taste homogeneity, naïve voters and a uniform income distribution. Epple, Filimon and Romer (1993) show existence in a model with property taxes and homogeneous tastes. Unfortunately, as in other models with taste heterogeneity (Epple and Platt, 1998), a proof of existence and uniqueness of this equilibrium can not be established. However, equations (1) to (4) provide the basis for a computational strategy to find equilibria numerically.

[9] The notion of 'stability' in an intrinsically static model is rather peculiar. Nevertheless, equilibria in static multi-community models are often judged by their 'dynamic' behavior. In this ad-hoc interpretation, an equilibrium is called 'stable' when the change of community characteristics induced by the migration of 'few' households gives these households an incentive to move back.

3 Numerical Equilibrium

In this section the qualitative and quantitative properties of the model are investigated in a fully specified and calibrated model.

3.1 Calibration

I calibrate the above outlined model to the metropolitan area of Zurich in Switzerland. The area around the city of Zurich forms the biggest Swiss metropolitan area. The city of Zurich has about 330,000 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 the introduction, each of these communities can set its own level of income taxes.

The analysis is restricted to a ring of the most integrated communities around the center. This ring is formed by all communities in the canton of Zurich with more than 1/3 of the working population commuting to the center.[10] The center community itself, the city of Zurich, is not included in the analysis, because the model cannot be expected to predict the center-periphery pattern well.[11]

The whole area of the peripheral communities has a physical size of $261km^2$ of which $87km^2$ are dedicated to development. In 1997, the fringe communities were populated by around 281,882 inhabitants (the city of Zurich had 335,943 inhabitants).[12] The system of 40 communities in the periphery is modelled as two distinct jurisdictions with equal land area. The two groups of communities consist of the 17 communities with the highest income tax rates and the 23 communities with the lowest tax rates, respectively. The community characteristics of this area are discussed in the introduction (see Figure 1). The calibrated parameters are summarized at the bottom of Table 1.

Income Distribution. The income distribution is calibrated with data from the Swiss Federal Tax Administration. We use a log-normal distribution to approximate this

[10]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 somewhat arbitrary definition of the urban area is chosen to justify the model's assumption that the 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. A wider area around the city of Zurich would include smaller but locally important job clusters such as the city of Wintherthur and the towns around the airport.

[11]I thank an anonymous referee for pointing to this fact and urging me to rethink the explanatory power of the model. Multi-community models are inherently spaceless and do therefore not include any attractiveness from being central. The center community does - not unexpectedly - not fit the general pattern predicted by the model: Despite its very high taxes, the city of Zurich, has relatively high housing prices. Furthermore, the center community provides many services which are in fact consumed to a large extent by households living in the suburbs. The per capita expenditures of the city of Zurich are therefore higher than in any community in the periphery. Note that Epple and Sieg (1999) also exclude the city of Boston in their empirical study of the Boston Metropolitan area.

[12]Source: Statistisches Amt des Kantons Zürich, Gemeindedaten per 31.12.1998.

right-skewed distribution. The estimation of the mean $E(\ln y) = 11.05$ and standard deviation $SD(\ln y) = 0.53$ from the observed income bins is described in Appendix 2. For numerical tractability, the model distribution is truncated at a minimum income of $y_{min} = 23,000$ and a maximum income $y_{max} = 500,000$.[13]

Taste Distribution. The taste distribution is calibrated with data from the Swiss labor force survey.[14] The 1995 cross-section contains monthly housing expenditure of renters for 1124 households in the above defined region.[15] Using the housing demand function in section 2.1, the taste parameter α of a household with disposable income y_d can be calculated as $(ph - ph_{min})/(y_d - y_{d,min})$, where ph is expenditure on housing and ph_{min} is the housing expenditure of the household with minimal disposable income $y_{d,min}$. The disposable income of a household y_d is calculated as reported household income minus federal, state and communal taxes. The average yearly housing expenditure of households around subsistence level is taken to approximate ph_{min}. This enables to approximate each household's taste parameter α. A beta distribution with mean $E(\alpha) = 0.17$ and standard deviation $SD(\alpha) = 0.11$ adequately describes the distribution of the so calculated taste parameter. Taste and income are assumed to be uncorrelated.[16]

Housing and Public Good Production. The price elasticity of housing supply is $\theta = 3$ as in Epple and Romer (1991) and Goodspeed (1989). The production of the public good exhibits constant per capita costs, i.e. $c_0 = 0$ and $c_1 = 1$.

Preference Parameters. The parameters $\beta_h = 700$, $\beta_b = 13000$ are chosen such that the consumption bundle of the minimal income household in equilibrium corresponds to the empirical findings. The benefit from additional units of the public good is taken from Goodspeed (1989) as $\gamma = 0.02$. The existential needs for the public good $\beta_g = 3100$ is set to produce equilibrium public expenditures close to the observed ones. Average local expenditures for education, security, traffic, culture, health, planning and administration were 4023 CHF per taxpayer in 1997.[17]

Additional data is used to assess the accuracy of the calibrated model (summarized in columns 6 to 8 in Table 1). In the year 1997, the communal tax rate for a married couple with taxable income of CHF 70,000 ranged from 4.7% to 5.7 % in the low-tax

[13] The minimum income is subsistence level for a one-person-household as defined by the Schweizerische Konferenz für Sozialhilfe (SKOS) and adjusted for inflation. The maximum income is chosen arbitrarily, but has no influence on the numerical simulation due to the low weight on high incomes.

[14] Swiss Federal Statistical Office, Schweizerische Arbeitskräfterhebung (SAKE) 1995.

[15] Of course, there is a selection bias by only considering renters. Because the proportion of renters is very high in Switzerland (65% in the data set used), this is not a first-order concern.

[16] Allowing for correlation between taste and income would introduce an (additional) systematic relationship between income and housing demand. This systematic relationship is, however, already modelled by the non-homothetic preferences that lead to decreasing relative housing expenditures with increasing incomes. The random component of preferences is therefore assumed to be orthogonal to income. This assumption is also used by Epple and Platt (1998).

[17] Source: Statistisches Amt des Kantons Zürich.

communities and from 5.8% to 7.2% in the high tax communities.[18] The rental price for housing was on average CHF 210 per m^2 in the high tax communites and 237 in the low tax communities.[19] The median income in the high-tax communities is CHF 63,900 opposed to CHF 80,600 in the low-tax communities. This large differences in the tax base allow the low tax communities to have about 10% higher public expenditures than the high-tax communities (CHF 4241 vs. 3810 per taxpayer). Figure 1 (in the introduction) visualizes the spatial distribution of tax rates, incomes, housing prices and public goods provision.

3.2 Simulated Equilibrium

The equilibrium values p_j, g_j and t_j, $i = 1, 2$, must satisfy equations (2), (3) and (4) and guarantee that the households reside in the community they prefer as expressed in equation (1). Unfortunately, there is no closed form solution to this nonlinear system of 6 equations, i.e. Equations 2, 3, and 4 in both communities, and 6 unknowns. The equation system is therefore numerically solved for the equilibrium values of the model.[20]

Table 1 shows the equilibrium values for the calibrated model in columns 4 and 5. As can be seen, the equilibrium values of the two communities differ substantially. The tax rate t_1 in the high-tax community is 39% higher than in the low-tax community, whereas the housing price is 12% and the public provision 19% lower. The average household income in the high-tax community is CHF 52,964 a year compared to CHF 87,401 in the low-tax community. Thus, the simulated model explains very well not only the observed sign but also the magnitude of the differences in tax rates, housing prices and public goods provision. The model also explain, but slightly overestimates, the observed difference in incomes. The equilibrium values for the case of taste homogeneity are given for comparison in columns 2 and 3. With taste homogeneity, the model predicts perfect income segregation and hence much higher than observed differences in tax rates, land rents and public goods provisions.

The segregation of the population in the two communities is shown in Figure 4. The left picture shows the locus of indifferent households, $\hat{y}_{12}$ which turns out to be an increasing function of income in the present equilibrium. This implies that, given a

[18] Source: Statistisches Amt des Kantons Zürich, Steuerfüsse 1997 and Tax scheme according to Steuergesetz vom 8. Juni 1997, Tarif a.

[19] Source: Wüest und Partner, Zurich. Offer prices for apartments in newspapers and online in 1997.

[20] Numerically solving the equation system is tedious and time-consuming. The aggregation of individual demand and voting behavior requires double integrals over the community population. These integrals cannot be calculated analytically. Gauss-Legendre Quadrature with 40 nodes in each dimension is used to approximate the various double integrals. Numerically minimizing the sum of squared deviations from the equilibrium conditions with the Gauss-Newton method solves for the equilibrium values. Appropriate scaling of the arguments and of the equilibrium conditions is important for the accuracy of the result. Convergence is only achieved with good starting values. Starting values are obtained from a grid search over the six-dimensional space of possible values.

Table 1: Equilibrium values of the simulation and data.

	model simulation					data 1997		
		homogeneous preferences		heterogeneous preferences		Zurich metropolitan area (excluding city of Zurich)		
	harmonized	high-tax	low-tax	high-tax	low-tax	whole[a]	high-tax[a]	low-tax[a]
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Community characteristics								
t: income tax rate	0.055	0.081	0.046	0.068	0.049	0.058[b]	0.065[b]	0.052[b]
p: rent	11.29	9.89	12.29	10.62	11.87	223[c]	210[c]	237[c]
p: rent, relative	1	1	1.24	1	1.12	1	1	1.13
g: public good prov.	4064	3531	4587	3624	4321	4023[d]	3810[d]	4241[d]
Size								
L: area	1	0.5	0.5	0.5	0.5	1	0.497[e]	0.503[e]
n: inhabitants	1	0.461	0.539	0.389	0.611	1	0.494[f]	0.506[f]
Income Distribution								
Ey: mean income	73,992	43,730	99,842	52,964	87,401	73,992[g]	63,864[g]	80,562[g]
Median income	64,156	44,205	88,239	43,332	76,900	64,156[g]	56,163[g]	68,803[g]
Welfare consequences (compensating variation w.r.t. tax harmonization)								
Average cv		18.0	234.2	56.6	34.8			
n (cv > 0)		52%	93%	61%	66%			

The calibrated model parameters: $\beta_h = 700$, $\beta_b = 13000$, $\beta_g = 3100$, $\gamma = 0.02$, $E(\alpha) = 0.17$, $SD(\alpha) = 0.11$ (heterogeneous tastes), $SD(\alpha) = 0$ (homogeneous tastes), $E(ln\, y) = 11.05$, $SD(ln\, y) = 0.53$, $y_{min} = 23,000$, $y_{max} = 500,000$, $\theta = 3$, $c_0 = 0$ and $c_1 = 1$.

[a] 23 communities with lowest taxes vs. 17 communities with highest taxes, excluding city of Zurich

[b] local income tax rate, married couple with taxable income of CHF 70,000, taxpayer weighted average

[c] CHF per m^2, population weighted mean

[d] expendit. for education, security, traffic, culture, health, planning and admin., CHF per taxpayer

[e] area dedicated for housing

[f] local resident population

[g] see Appendix 2, truncated as in simulation

subpopulation with equal tastes, richer households prefer the low-tax-high-price community.[21] However, this does not lead to perfect income segregation between the two communities since the households have different preferences. Although the average income in the high-tax community is much lower than in the low-tax community, households from almost all income groups can be found in both communities. The right picture in Figure 4 presents the resulting income distributions in the two communities. Figure 4 left also shows the loci of pivotal voters which split the communities' populations into half. Note that the distribution in the high-tax community as well as the one in the low-tax community are skewed to the right and the mean is considerably above the median. This replicates the observed pattern in the Zurich area. Households in rich low-tax communities vote for more public goods than households in poor high-tax communities,

[21]Note that the households with a very high taste for housing prefer to live in the low-price community. This, however, applies to only 5% of the population, as the weight on taste parameters above 0.33 is low.

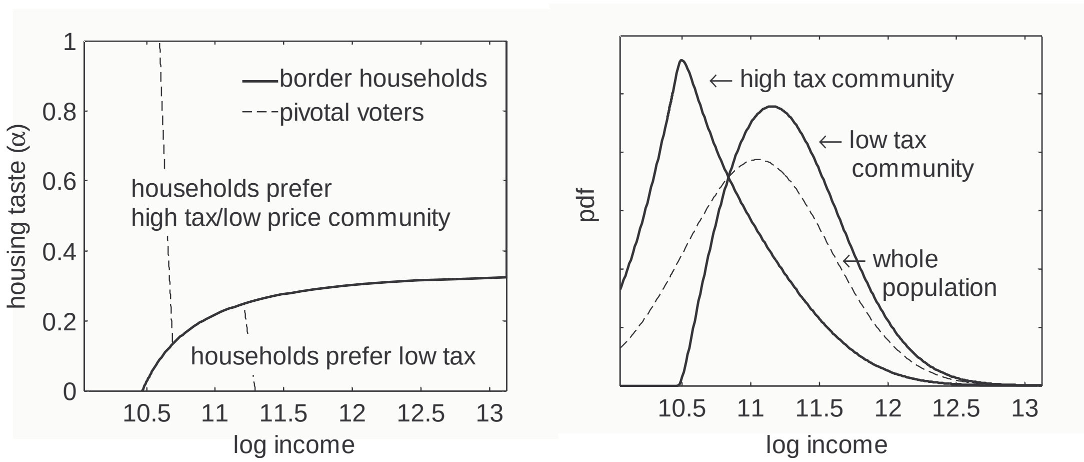

Figure 4: Income and taste segregation in equilibrium. The left figure shows the preferred community for all household types. The right figure shows the resulting income distributions in both communities.

yet this generous public good provision can be financed by a lower tax rate, due to the higher average income of the residents.

The prediction that rich households locate in low-tax communities is also established by Kessler and Lülfesmann (2005). However, in their model the attractiveness of low taxes is offset by low public good provision in the rich community. In my more general model, it is the high housing prices that deter the poor from moving to rich communities which offer both attractive taxes and public good provision. The high public good provision in the rich community corresponds with the empirical study by Epple and Sieg (1999) and the theoretical work by Epple, Filimon and Romer (1984, 1993) as well as with the specified version of Nechyba's (1997, p. 297) model.[22] All of these models predict higher after-property-tax housing prices in the rich community, but they do note make a statement about the relation of taxes. Note that Epple and Romer (1991) and Epple and Platt (1998) predict - somewhat against their own evidence - the exact opposite relation of public good provision (modelled as a lump-sum transfer) and housing prices across communities. Hansen and Kessler (2001a) predict lower taxes, lower public good provision (modelled as a lump-sum transfer) and higher housing prices in the rich community. Hence none of these models is able to explain the pattern of community characteristics as observed in Swiss metropolitan areas.

[22]The general model of Nechyba provides an powerful existence proof under very weak conditions for the case of property taxation but yields only limited statements on the possibly emerging segregation of the population.

3.3 Welfare Consequences of Decentralization

The above segregated equilibrium is now compared to the equilibrium when jurisdictions harmonize their income tax levels and households locate randomly.[23] The equilibrium values with tax harmonization are presented in column 1 in Table 1. One can immediately see that the housing price, tax level and public good provision lie between the corresponding values in the two-community model. The decentralization of tax authority to two communities does thus not lead to an overall reduction of taxes, but to relatively lower taxes in the rich community and higher taxes in the poor community.

The welfare effects from decentralized taxation and the associated segregated equilibrium depend on both the households' incomes and tastes. They are revealed by inspecting the compensating variation (cv), defined as the additional gross income that compensates a household for a shift from the equilibrium with harmonized taxes (t_h, p_h, g_h) to the segregated equilibrium:[24]

$$V(t_h, p_h, g_h; y, \alpha) = V(t_j, p_j, g_j; y + cv, \alpha).$$

The implied compensating variation cv depends on both the household's income y and taste α as well as the community j it chooses in the decentralized case:

$$cv_j(y, \alpha) = \frac{[y(1-t_h) - \beta_b - p_h\beta_h](\frac{p_h}{p_j})^{-\alpha}(\frac{g_h-\beta_g}{g_j-\beta_g})^{\gamma} - [y(1-t_j) - \beta_b - p_j\beta_h]}{1-t_j}.$$

Table 1 reports the average cv for each community. Households in the poor community have to be compensated by an average income allowance of CHF 57 compared to CHF 35 in the rich community. Note that this amount is only about one-tenth of a percent of the average gross income.

The reported average of the compensating variation hides the heterogeneity of welfare consequences across different household types. The left picture in Figure 5 shows contour lines of the cv for all household types. Households in the shaded band between the two zero contour lines exhibit positive values of the cv and thus prefer tax harmonization. Households further away from the border household prefer competing jurisdictions. The right picture in Figure 5 shows the cv across incomes for a household with housing taste $\alpha = \overline{\alpha} = 0.17$. Note that given these housing tastes, households with income below CHF 44,400, $log(y) < 10.7$, live in the poor community 1. The poorest in the poor community, $log(y) < 10.4$, prefer decentralication, the poorest by a cv of more than CHF -150. The relatively richer in the poor community, $10.4 > log(y) < 10.6$, prefer tax harmonization. The border household, which is indifferent between the two communities,

[23] Tax harmonization is equivalent to the equilibrium that would emerge if the two distinct jurisdictions merged. It is also equivalent to the (non-stable) symmetric equilibrium in the decentralized situation.

[24] The compensating variation defined above ignores the welfare implications for the (absentee) landlord.

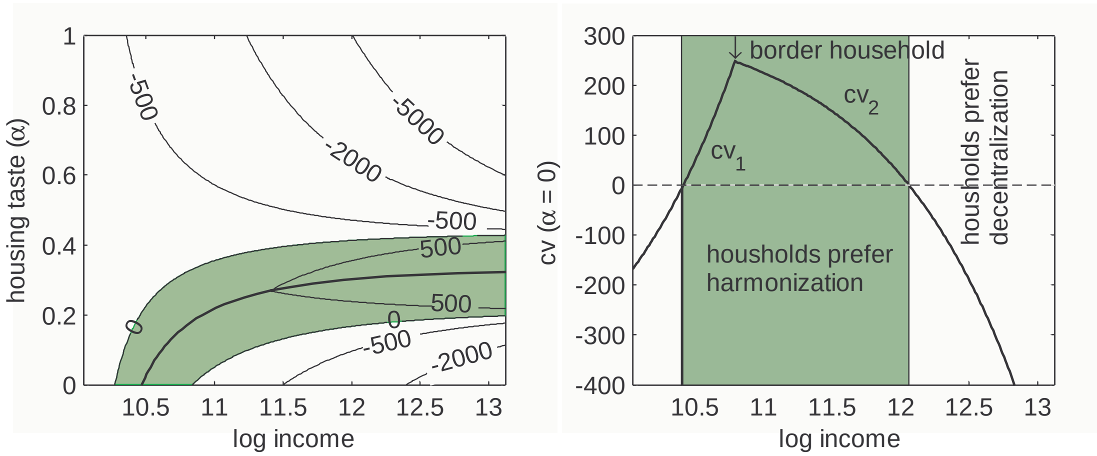

Figure 5: Welfare costs of decentralization. The left figure shows contour lines of the compensating variation (cv) for all household types. The right figure shows the compensating variation for households with housing taste $\alpha = E\alpha = 0.17$.

needs a compensation of about CHF 250 to prefer decentralization. The households in the rich community have opposed preferences towards fiscal decentralization: the relatively richer, $log(y) > 12.1$, prefer decentralization whereas the relatively poorer, $10.6 > log(y) < 12.1$, would prefer harmonized taxes. Summing up, the households with a clear preference for one of the two communities will also benefit the most from the decentralized arrangement. The number of households that prefer tax harmonization is also given for each community in Table 1: 61% of the population in the poor community and 67% of the population in the rich community prefer tax harmonization.[25]

Note that the above outlined welfare consequences are very sensitive to the assumed model parameters. The following three sections discuss the sensitivity of the simulation results to changes in the main parameter.

3.4 The Role of Preference Heterogeneity

How does the heterogeneity of tastes affect the properties of the equilibrium? I will answer this question by studying the calibrated model assuming different levels of taste variance while leaving the average level of tastes constant.

I start with the extreme case of homogeneous tastes. Columns 2 and 3 in Table 1 give the equilibrium values for this case. The population is now perfectly segregated

[25]It is the middle class which prefers tax harmonization and a coalition of the very rich and the very poor which prefer tax decentralization. This result reminds of Epple and Romano's (1996) "ends against the middle", where a coalition of rich and poor households vote in favor of privatization of schooling against the middle class which is in favor of public provision of schooling. A detailed politico-economic analysis of the choice for fiscal decentralization is, however, beyond the scope of this paper.

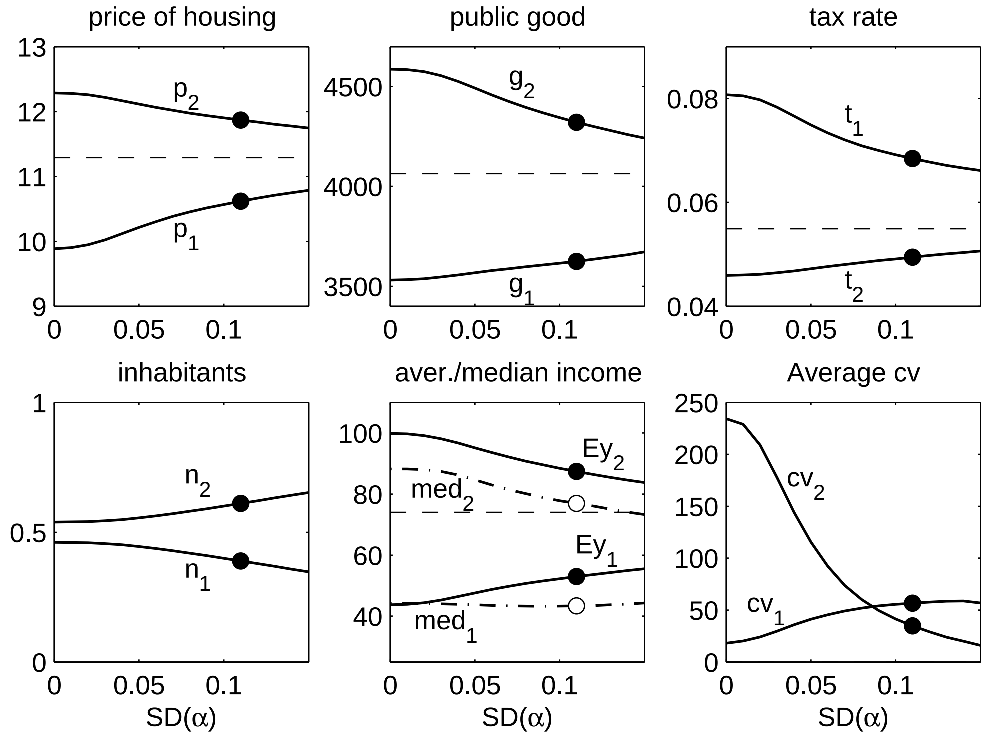

Figure 6: Equilibrium values for different variance of tastes. The dashed lines indicate the values with tax harmonization. The circles indicate the calibrated equilibrium.

by incomes. Consequently, the income difference between the two communities is much larger than with heterogeneous tastes. Also, the differences in prices, taxes and public good provision across the communities are much stronger. Separating the single peaked right-skewed distribution into two closed intervals leaves a right-skewed distribution (median smaller than average income) in the rich community and a left-skewed distribution (median greater than average income) in the poor community. The welfare effects under the assumption of taste homogeneity are substantially bigger than under heterogeneity. Note that taste heterogeneity has no impact on the equilibrium values with tax harmonization, i.e. column 1 shows the equilibrium values for all degrees of taste heterogeneity.[26]

Figure 6 shows a series of equilibrium values for different degrees of taste heterogeneity measured by the standard deviation $SD(\alpha)$, leaving the mean of tastes constant. The horizontal axes cover a range from $SD(\alpha) = 0$ (taste homogeneity) to a maximum

[26] The unified equilibrium in the case of taste homogeneity is theoretically different from the one in the case of taste heterogeneity as the pivotal voter varies with the taste parameter. However, this difference is numerically negligible.

$SD(\alpha) = 0.143$[27] including the calibrated case $SD(\alpha) = 0.11$. The picture reveals that the equilibrium approaches the values of the situation with harmonized taxes, indicated by the dotted lines. This is explained by the fact that with increasing taste variance, the population is more and more segregated by taste rather than income. This result suggests that taste heterogeneity is able to lower the negative distributional effects of a decentralized tax regime. Figure 6 (far right) shows the corresponding change of the average compensating variation. While the average cv in the poor community is almost unaffected by the amount of taste heterogeneity in the population, it falls sharply in the rich community. The fraction of households in the rich community which would prefer harmonized taxes (not reported) falls accordingly from 100% ($SD(\alpha) = 0$) to 63% ($SD(\alpha) = 0.143$). Partial taste segregation also generates right-skewed distributions in both communities.

3.5 The Role of Relative Land Area

How does the relative land area in the two jurisdictions affect the properties of the equilibrium? Up to now, the two communities were assumed to be of equal physical land area. However, one could also partition the 40 communities into a large and small group.

Figure 7 shows a series of equilibrium values for different relative size of the two jurisdictions. The horizontal axes specifies the land area L_1 of the high-tax community, indexed by 1, as fraction of total land area. The housing price and the public good provision in both communities increase with the physical size of the poor community. The poor community shows higher taxes, lower public good provision and lower housing prices than the rich community throughout all possible partitions of land between the two communities. The order of community characteristics is hence not affected by the relative land area. Not surprisingly, the equilibrium values of a community that virtually covers the whole area ($L_1 = 1$ or $L_1 = 0$, hence $L_2 = 1$) equal the values of the equilibrium with tax harmonization, marked by the dotted lines. The equilibrium values in the remaining very small community differ maximally from the values with harmonized taxes. The tax rate in the rich community, indexed by 2, declines with increasing relative land area in community 1. This shows that the rich community has more power to set low taxes when it is physically small. This echoes the findings by Hansen and Kessler (2001a), who show that segregated equilibria exist only when the rich community is small. However, my results demonstrate that their finding is very model specific and segregation can also emerge in situations with equally sized local jurisdictions.

The influence of the relative land area on welfare is particularly interesting. Recall

[27] Given the mean $E\alpha = 0.17$, $SD(\alpha) = 0.143$, is the maximal standard deviation that preserves the bell-shaped form of the beta distribution. Higher values lead to a u-shaped distribution.

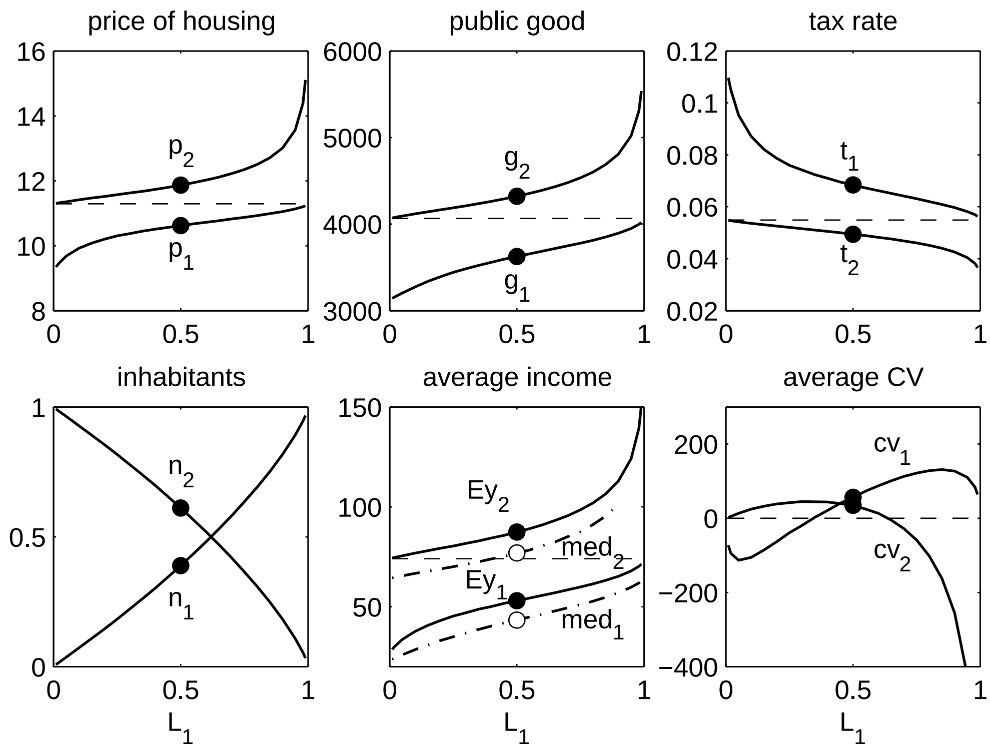

Figure 7: Equilibria for different relative area of the communities. The dashed lines indicate the values with tax harmonization. The circles indicate the calibrated equilibrium.

that in the calibrated situation ($L_1 = 0.5$), the average household prefers harmonized taxes: average compensating variation is CHF 57 in the high-tax and CHF 35 in the low-tax community. This result does strongly depend on the relative community size as can be seen in Figure 7 (far right). The average compensating variation in the poor high-tax community, cv_1, is negative if this community is small ($L_1 < 0.35$), meaning that the population does on average prefer (higher) local taxes to (lower) harmonized taxes, as they are associated with lower housing prices. Note that it is the poorer part of the population in the poor community that profits most from the local differences. The rich low-tax community shows a similar picture. Its increased ability to set low taxes when it is small ($L_1 > 65$ hence $L_2 < 0.35$) leads to a negative average compensating variation, i.e. an average preference for decentralized taxation.[28] Note that it is the richer part of the population that profits most from the decentralized tax setting.

[28]This result indicates - speculating beyond the limits of the model - that the rich would choose to locate in a small community. Anecdotic evidence from Switzerland and Europe suggests that tax havens are indeed usually small.

3.6 The Role of Preferences towards the Public Good

How do the assumed preferences for the local public good affect the properties of the equilibrium? This section explores the effect of the elasticity of substitution between the public and the private good.[29] The elasticity of substitution measures how easily a household can substitute the public good with the composite private goods. Unfortunately, there is no single parameter that sets the elasticity of substitution given Stone-Geary utility. The elasticity of substitution

$$\sigma_{g,b} \quad := \quad -\frac{\partial ln(\frac{g}{b})}{\partial ln(M_{g,b})} = \frac{\frac{b-\beta_b}{b(1-\alpha)} + \frac{g-\beta_g}{g\gamma}}{\frac{1}{1-\alpha} + \frac{1}{\gamma}}$$

is described by the parameters β_g and γ and depends on household income (via the private good consumption b) and the taste α for housing (see Appendix 1 for derivation). In the calibrated equilibrium, I set γ to 0.02 and $\beta_g = 3100$ to a value that produces average tax rates as observed. In this section I vary σ_g, b while holding the marginal rate of substitution between the public and the privat good

$$M_{g,b} \quad := \quad \frac{dg}{db} = -\frac{\partial U/\partial b}{\partial U/\partial g} = -\frac{(1-\alpha)\,(g-\beta_g)}{(b-\beta_b)\,\gamma}$$

for any household type constant. This is guaranteed by holding $(g - \beta_g)/\gamma$ constant. A constant marginal rate of substitution has several favorable properties. Firstly, this procedure changes the curvature of indifference curves while leaving them tangential at any given value of g and b, hence separating the effect from curvature from the overall esteem for the public good. Secondly, it leaves the values of the benchmark equilibrium of unified communities unchanged because the marginal rate of substitution between the local public good and local tax rates, $M_{g,t}$, and therefore the pivotal voter's perceived trade-off between public good provision and taxes $dg/dt\mid_{dV=0,HD=HS}$ remains also constant.

Figure 8 shows a series of the equilibrium values for γ ranging from 0 to 0.18 and β_g accordingly from 4064 to -4610. The implied elasticity of substitution between the public and the private good for a household with average income and taste is graphed on the horizontal axes and ranges from 0 to 1.94. The extreme case on the left with $\gamma = 0$, $\beta_g = 4064$, $\sigma_{b,g} = 0$ means that the public good and the private good are perfect complements. In this situation, the equilibrium public good provision is exogenously given as $g = \beta_g = 4064$. Despite the identical public good provision, there is income sorting across the two communities resulting in the typically observed pattern of community characteristics: the rich community 2 exhibits lower taxes and higher housing

[29]This exercise could as well be based on the elasticity of substitution between the public good and *housing*. See the formulas in Appendix 1.

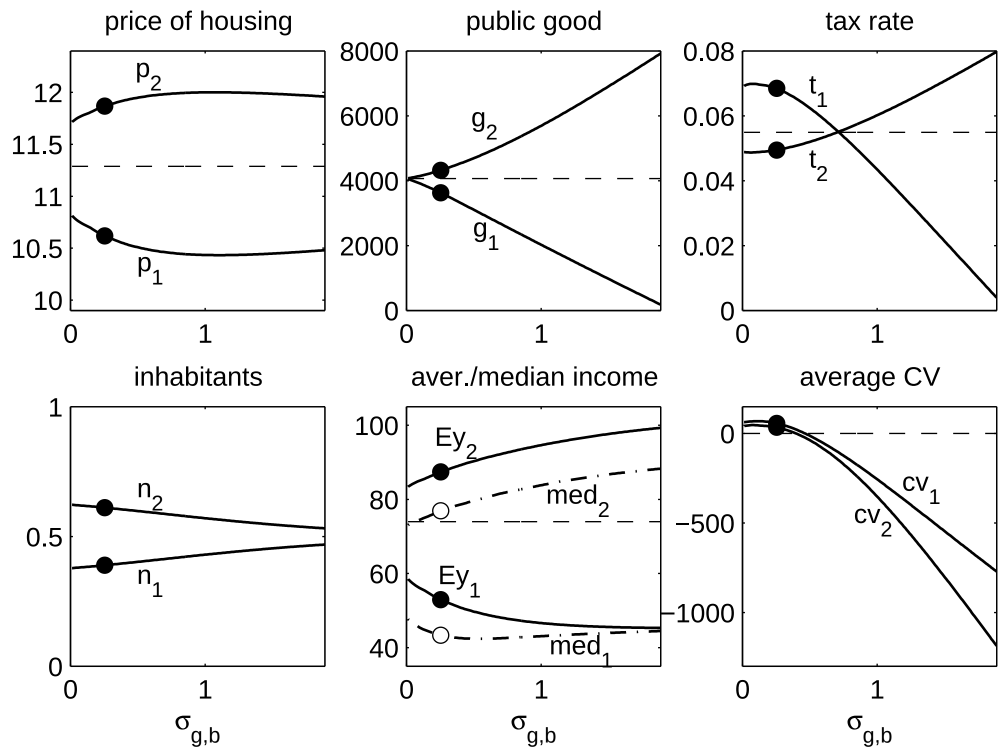

Figure 8: Equilibrium value for different levels of substitutability between public and private good. The dashed lines indicate the values with tax harmonization. The circles indicate the calibrated equilibrium.

prices than the poor community 1. The lower taxes in the rich community are a direct consequence of the larger tax base per capita. The high housing prices in the rich community scare the poor more than the rich and keep them from locating the low-tax community.

Increasing the elasticity of substitution between g and b fundamentally changes the properties of the equilibrium. The population n_2 in the rich community shrinks and becomes richer (Ey_2) with increasing $\sigma_{g,b}$. The public good provision g increases monotonically in the rich community and decreases monotonically in the poor community. Note that median household income is smaller than average income over the whole range of $\sigma_{g,b}$ in both communities. The pivotal voters therefore successfully use the public good provision as an instrument for redistribution. The pivotal voters in the rich community target a higher level of public good than the voters in the poor community because the marginal tax rate to finance an additional unit of public goods decreases with average income. The tax rates in the two communities show the most dramatic changes: the tax rate in the poor community is initially larger, increases slightly before it falls monotonically to zero. The tax rate in the rich community is initially smaller, falls slightly and then increases steadily. Housing prices in the rich community are higher than in the poor community over the whole range of $\sigma_{g,b}$. Interestingly, for a higher degree of substitutability of the public good the rich choose the community that offers more public good provision even though that community yields high taxes *and* high housing prices. This is a consequence of the assumed linear expenditure system which implies that the marginal rate of substitution between the public and the private good increases with income, i.e. the public good becomes a more and more scarce good. Note that there is an equilibrium in which the tax rates in the two communities equal. In this situation, the income sorting of the population is fully driven by the trade-off between local public good provision and local housing prices.

The numerical results in the exercise in this section are very model and parameter specific. However, there are two general lessons. Firstly, different form the literature on property taxation, there is no natural order of community characteristics in multi-community models with income taxation. Rich communities may be attractive to rich people because of either low taxes or a high level of public good provision. Secondly, the non-existence proposition of Hansen and Kessler (2001b) is very specific to their model and does not apply in general income tax models.

4 Conclusions

This paper presents a model of an urban area with local income taxes used to finance a local public good. The main assumptions of the model are: Households differ in incomes and tastes for housing. The demands for housing and non-housing consumption is a linear expenditure system. The share of housing in the budget of the households is on average declining with income. Non-housing consumption is only a partial substitute for the publicly provided good.

The existence of a segregated equilibrium is shown in a calibrated two-community model assuming realistic single-peaked distributions for income and taste in housing. The low-tax community exhibits both higher housing prices and higher public good provision than the high-tax community. The equilibrium features segregation of households by both incomes and tastes. The emerging segregation pattern is such that most rich households prefer the low-tax high-price community. As tastes differ across households, this does not lead to a perfect income segregation but to an income distribution in the rich low-tax community that stochastically dominates the income distribution in the poor high-tax community: while households from all income groups can be found in both communities, average income in the high-tax community is much lower than in the low-tax community. The model is able to explain the substantial differences in local income tax levels, average incomes, local housing prices and the local provision of public goods across communities in Swiss urban areas.

The numerical investigation shows that the order of community characteristics depends on the preferences for the local public good. The above ordering of community characteristics holds for low degrees of substitutability between public and private goods. When the public good is easily substituted by private goods the rich community exhibits higher housing prices and higher public goods provision as well as higher taxes.

The numerical investigation also suggests that taste heterogeneity reduces the distributional effects of local tax differences. The differences of characteristics across communities are maximal when tastes are equal for all households and when the population is accordingly perfectly segregated by income. These differences decrease with increasing taste heterogeneity as the income segregation of the population becomes more and more diffuse.

The numerical investigation furthermore suggests that the relative size of the individual jurisdictions has great impact on the equilibrium outcome. The characteristics of a relatively large community are close to the equilibrium characteristics of a single jurisdiction that covers the whole area. Conversely, the relatively small community differs substantially from the single jurisdiction. For example, rich communities are able to set lower taxes when they are small. However, contrary to the findings by Hansen and Kessler (2001a), a tax haven need not be small.

Multi-community models are especially well-suited to study metropolitan areas as they assume that the residence choice of a household is made after and independent of the decision of where its members work. Nevertheless the results presented in this paper may also shed light on fiscal decentralization at the level of states or countries.

Appendix 1

The household problem is

$$\max_{h,b} U(h,b,g,\alpha) = \alpha\, ln(h-\beta_h) + (1-\alpha)\, ln(b-\beta_b) + \gamma\, ln(g-\beta_g)$$
$$s.t. \quad ph + b \leq y(1-t)\,.$$

This leads to the housing demand

$$h^* = h(t,p,y,\alpha) = \frac{\alpha[y(1-t) - p\beta_h - \beta_b]}{p} + \beta_h\,,$$

the income elasticity of housing

$$\varepsilon = \frac{\partial h^*}{\partial y(1-t)} \frac{y(1-t)}{h^*} = \frac{\alpha y(1-t)}{\alpha[y(1-t) - p\beta_h - \beta_b] + p\beta_h}$$

and the indirect utility function

$$V = \alpha\, ln(\alpha) + (1-\alpha)\, ln(1-\alpha) - \alpha\, ln(p) + ln[y(1-t) - p\beta_h - \beta_b] + \gamma\, ln(g-\beta_g)\,.$$

The marginal rates of substitution in Property 1 are derived by totally differentiating the indirect utility function:

$$\begin{aligned} M_{g,t} &:= \frac{dg}{dt} = -\frac{\partial V/\partial t}{\partial V/\partial g} = \frac{y(g-\beta_g)}{\gamma[y(1-t) - p\beta_h - \beta_b]}\,, \\ M_{g,p} &:= \frac{dg}{dp} = -\frac{\partial V/\partial p}{\partial V/\partial g} = \frac{h^*(g-\beta_g)}{\gamma[y(1-t) - p\beta_h - \beta_b]}\,, \\ M_{t,p} &:= \frac{dt}{dp} = -\frac{\partial V/\partial p}{\partial V/\partial t} = -\frac{h^*}{y}\,. \end{aligned}$$

Differentiation of the MRS w.r.t. income and taste yields Property 2:

$$\begin{aligned} \frac{\partial M_{g,t}}{\partial y} &= -\frac{(g-\beta_g)(p\beta_h+\beta_b)}{\gamma[y(1-t) - p\beta_h - \beta_b]^2}\,, & \frac{\partial M_{g,t}}{\partial \alpha} &= 0\,, \\ \frac{\partial M_{g,p}}{\partial y} &= -\frac{(1-t)(g-\beta_g)\beta_h}{\gamma[y(1-t) - p\beta_h - \beta_b]^2}\,, & \frac{\partial M_{g,p}}{\partial \alpha} &= \frac{g-\beta_g}{p\gamma}\,, \\ \frac{\partial M_{t,p}}{\partial y} &= \frac{(1-\alpha)p\beta_h - \alpha\beta_b}{py^2}\,, & \frac{\partial M_{t,p}}{\partial \alpha} &= -\frac{y(1-t) - p\beta_h - \beta_b}{py}\,. \end{aligned}$$

The independence of the MRS ratio in Property 3 follows directly:

$$\frac{\partial M_{g,t}}{\partial y} \Big/ \frac{\partial M_{g,p}}{\partial y} = \frac{p\beta_h + \beta_b}{(1-t)\beta_h}\,,$$

$$\frac{\partial M_{t,g}}{\partial y} / \frac{\partial M_{t,p}}{\partial y} = \frac{\gamma p (p\beta_h + \beta_b)}{[(1-\alpha)p\beta_h - \alpha\beta_b](g - \beta_g)} ,$$

where $M_{t,g} = 1/M_{g,t}$.

The locus of indifferent households between community j and i

$$\hat{y}_{ji}(\alpha) = \frac{(p_j\,\beta_h + \beta_b)\,{p_i}^{\alpha}\,(g_j - \beta_g)^{\gamma} - (\beta_b + p_i\,\beta_h)\,{p_j}^{\alpha}\,(g_i - \beta_g)^{\gamma}}{(1-t_j)\,{p_i}^{\alpha}\,(g_j - \beta_g)^{\gamma} - (1-t_i)\,{p_j}^{\alpha}\,(g_i - \beta_g)^{\gamma}} .$$

solves $V(t_j, p_j, g_j, y, \alpha) = V(t_i, p_i, g_i, y, \alpha)$ for y. Alternatively, the locus solves for α:

$$\hat{\alpha}_{ji}(y) = \frac{ln[\frac{y(1-t_j) - p_j\beta_h - \beta_b}{y(1-t_i) - p_i\beta_h - \beta_b}] + \gamma\, ln[\frac{g_j - \beta_g}{g_i - \beta_g}]}{ln(p_j/p_i)} .$$

The locus $\hat{\alpha}_{ji}(y)$ is either strictly increasing and concave in y or strictly decreasing and convex, as can easily be verified by inspecting the first and second derivative

$$\frac{\partial \hat{\alpha}_{ji}}{\partial y} = -\frac{(1-t_j)[p_i\beta_h - \beta_b] - (1-t_i)[p_j\beta_h - \beta_b]}{[y(1-t_j) - p_j\beta_h - \beta_b][y(1-t_i) - p_i\beta_h - \beta_b]\, ln(p_j/p_i)}$$

$$\frac{\partial^2 \hat{\alpha}_{ji}}{\partial y^2} = -\frac{\partial \hat{\alpha}_{ji}}{\partial y} \cdot \frac{(1-t_j)[y(1-t_i) - p_i\beta_h - \beta_b] + (1-t_i)[y(1-t_j) - p_j\beta_h - \beta_b]}{[y(1-t_j) - p_j\beta_h - \beta_b][y(1-t_i) - p_i\beta_h - \beta_b]}$$

and provided that all household reach the subsistence level, i.e. $y(1-t) > p\beta_h + \beta_b > 0$, in both communities.

The utility difference between community j and i is

$$V_j(y,\alpha) - V_i(y,\alpha) = -\alpha\, ln(\frac{p_j}{p_i}) + ln[\frac{y(1-t_j) - p_j\beta_h - \beta_b}{y(1-t_i) - p_i\beta_h - \beta_b}] + \gamma\, ln(\frac{g_j - \beta_b}{g_i - \beta_b}).$$

Differentiation of the above expression w.r.t. y and α is used in the proof of Propositions 1 and 2:

$$\frac{\partial (V_j - V_i)}{\partial y} = \frac{1}{y - \frac{p_j\beta_h + \beta_b}{1-t_j}} - \frac{1}{y - \frac{p_i\beta_h + \beta_b}{1-t_i}} , \quad \frac{\partial (V_j - V_i)}{\partial \alpha} = ln(p_i) - ln(p_j).$$

The rate of substitution between tax rate and public good provision a voter faces is derived from totally differentiating the indirect utility function considering the housing market reaction, $dp/dt|_{HD=HS}$ (community subscripts omitted):

$$\begin{aligned} \left.\frac{dg}{dt}\right|_{dV=0,HD=HS} &= \frac{-\frac{\partial V}{\partial t} - \frac{\partial V}{\partial p} \cdot \left.\frac{dp}{dt}\right|_{HD=HS}}{\frac{\partial V}{\partial g}} = M_{g,t} + M_{g,p} \cdot \left.\frac{dp}{dt}\right|_{HD=HS} \\ &= \frac{g - \beta_g}{\gamma[y(1-t) - p\beta_h - \beta_b]} \left[y + h^* \left.\frac{dp}{dt}\right|_{HD=HS} \right] . \end{aligned}$$

The voter's rate of substitution is positive when the price effect on the housing market is not too large:

$$\left.\frac{dg}{dt}\right|_{dV=0,HD=HS} > 0 \quad \text{iff} \quad \left.\frac{dp/p}{d(1-t)/(1-t)}\right|_{HD=HS} < \frac{y(1-t)}{ph^*} \text{ for all } \alpha.$$

The voter's rate of substitution decreases with income

$$\begin{aligned}\frac{\partial \left.\frac{dg}{dt}\right|_{dV=0,HD=HS}}{\partial y} &= \frac{\partial M_{g,t}}{\partial y} + \frac{\partial M_{g,p}}{\partial y}\left.\frac{dp}{dt}\right|_{HD=HS}. \\ &= -\frac{g-\beta_g}{\gamma[y(1-t)-p\beta_h-\beta_b]^2}\left[p\beta_h + \beta_b + (1-t)\beta_h\left.\frac{dp}{dt}\right|_{HD=HS}\right]\end{aligned}$$

if the price effect on the housing market is not too large:

$$\frac{\partial \left.\frac{dg}{dt}\right|_{dV=0,HD=HS}}{\partial y} < 0 \quad \text{iff} \quad \left.\frac{dp/p}{d(1-t)/(1-t)}\right|_{HD=HS} < \frac{p\beta_h+\beta_b}{p\beta_h}.$$

Both the condition on the sign of the voter's marginal rate of substitution and the sign of its derivative w.r.t. y are fulfilled if $\left.\frac{dp/p}{d(1-t)/(1-t)}\right|_{HD=HS} < 1$ and all households reach the subsistence level.

The compensating variation cv_j is the additional gross income that a household in e.g. community j needs in order to be compensated for a shift from the symmetric (tax harmonization) equilibrium, (t_h, p_h, g_h), to the asymmetric (segregated) equilibrium, (t_j, p_j, g_j). Solving

$$V(t_h, p_h, g_h; y, \alpha) = V(t_j, p_j, g_j; y + cv, \alpha)$$

for cv yields the compensating variation for a household with income y and taste α in community j:

$$cv_j(y,\alpha) = \frac{[y(1-t_h)-\beta_b-p_h\beta_h](\frac{p_h}{p_j})^{-\alpha}(\frac{g_h-\beta_g}{g_j-\beta_g})^{\gamma} - [y(1-t_j)-\beta_b-p_j\beta_h]}{1-t_j}.$$

The average compensating variation in community j is then computed as

$$\overline{cv}_j = \frac{1}{n_j}\int_0^1\int_{\underline{y}_j(\alpha)}^{\overline{y}_j(\alpha)} cv_j(y,\alpha)\, f(y,\alpha)\, dy\, d\alpha\,.$$

The marginal rates of substitution between the public and the private goods used in section 3.6 are derived by totally differentiating the tility function:

$$\begin{aligned} M_{g,b} &:= \frac{dg}{db} = -\frac{\partial U/\partial b}{\partial U/\partial g} = -\frac{(1-\alpha)\,(g-\beta_g)}{(b-\beta_b)\,\gamma}, \\ M_{g,h} &:= \frac{dg}{db} = -\frac{\partial U/\partial h}{\partial U/\partial g} = -\frac{\alpha\,(g-\beta_g)}{(h-\beta_h)\,\gamma}. \end{aligned}$$

The elasticity of substitution between the public and the private goods is:

$$\begin{aligned}
\sigma_{g,b} &:= -\frac{\partial ln(\frac{g}{b})}{\partial ln(\frac{\partial u/\partial g}{\partial u/\partial b})} = -\frac{\frac{1}{b\cdot\partial u/\partial b}+\frac{1}{g\cdot\partial u/\partial g}}{\frac{\partial^2 u/\partial b^2}{(\partial u/\partial b)^2}-2\frac{\partial^2 u/\partial b\partial g}{\partial u/\partial b\cdot\partial u/\partial g}+\frac{\partial^2 u/\partial g^2}{(\partial u/\partial g)^2}} = \frac{\frac{b-\beta_b}{b(1-\alpha)}+\frac{g-\beta_g}{g\gamma}}{\frac{1}{1-\alpha}+\frac{1}{\gamma}}\,, \\
\sigma_{g,h} &:= -\frac{\partial ln(\frac{g}{h})}{\partial ln(\frac{\partial u/\partial g}{\partial u/\partial h})} = -\frac{\frac{1}{h\,\partial u/\partial h}+\frac{1}{g\,\partial u/\partial g}}{\frac{\partial^2 u/\partial h^2}{(\partial u/\partial h)^2}-2\frac{\partial^2 u/\partial h\partial g}{\partial u/\partial b\cdot\partial u/\partial g}+\frac{\partial^2 u/\partial g^2}{(\partial u/\partial g)^2}} = \frac{\frac{h-\beta_h}{h\alpha}+\frac{g-\beta_g}{g\gamma}}{\frac{1}{\alpha}+\frac{1}{\gamma}}\,.
\end{aligned}$$

Both elasticities of substitution are increasing in γ holding $(g-\beta_g)/\gamma$ constant (thus also holding $M_{g,b}$, $M_{g,h}$ and $M_{g,t}$ constant):

$$\begin{aligned}
\left.\frac{\partial\sigma_{g,b}}{\partial\gamma}\right|_{\frac{g-\beta_g}{\gamma}} &= \frac{\frac{b-\beta_b}{b(1-\alpha)}+\frac{g-\beta_g}{g\gamma}}{(\frac{1}{1-\alpha}+\frac{1}{\gamma})^2}\frac{1}{\gamma^2} > 0\,, \\
\left.\frac{\partial\sigma_{g,h}}{\partial\gamma}\right|_{\frac{g-\beta_g}{\gamma}} &= \frac{\frac{h-\beta_h}{h\alpha}+\frac{g-\beta_g}{g\gamma}}{(\frac{1}{\alpha}+\frac{1}{\gamma})^2}\frac{1}{\gamma^2} > 0\,.
\end{aligned}$$

Appendix 2

I estimate the local income density $\hat{f}(y|j)$ from publicly available local income distribution data. The federal tax administration publishes the number of households with taxable income in seven different income classes.[30] I assume that incomes are log-normally distributed and estimate the mean μ_j and the variance σ_j^2 of this distribution using maximum likelihood.[31] I estimate a truncated log-normal distribution as the first reported income interval is empty for data collecting reasons. The log likelihood function for any community i is

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

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 $c_k \in \{log(15000), log(20000),$ $log(30000), log(40000), log(50000), log(75000), \infty\}$. $\Phi(.)$ is the cdf of the standard normal distribution.

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

[31] Note that this maximum likelihood estimator corresponds to an ordered probit with known thresholds.

References

Bayer, Patrick, Robert McMillan, and Kim Rueben, 2004, An Equilibrium Model of Sorting in an Urban Housing Market, NBER working paper 10865.

Calabrese, Stephen M., 2001, Local Redistribution Financed by Income Tax, Public Finance Review 29, 259-303.

Ellickson, 1971, Jurisdictional Fragmentation and Residential Choice, American Economic Review 61, 334-339.

Epple, Dennis, Radu Filimon and Thomas Romer, 1984, Equilibrium among Local Jurisdictions: toward an Integrated Treatment of Voting and Residential Choice, Journal of Public Economics 24, 281-308.

Epple, Dennis, Radu Filimon and Thomas Romer, 1993, Existence of Voting and Housing Equilibrium in a System of Communities with Property Taxes, Regional Science and Urban Economics 23, 585-610.

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.

Epple, Dennis and Thomas Nechyba, 2004, Fiscal Decentralization. In: Vernon Henderson and J. Thisse, eds., Handbook of Regional and Urban Economics Vol. 4, Amsterdam: North-Holland.

Epple, Dennis and Richard E. Romano, 1996, Ends against the middle: Determining public service provision when there are private alternatives, Journal of Public Economics 62, 297-325.

Epple, Dennis and Holger Sieg, 1999, Estimating Equilibrium Models of Local Jurisdictions, Journal of Political Economy 107, 645-681.

Epple, Dennis and Thomas Romer, 1991, Mobility and Redistribution, Journal of Political Economy 99,828-858.

Fernandez, Raquel and Richard Rogerson, 1996, Income Distribution, communities, and the Quality of Public Education, The Quarterly Journal of Economics 111, 135-64.

Goodspeed, Timothy J., 1986, Local Income Taxation: Equilibrium, Social Optimality and Welfare Loss in a General Equlibrium Metropolitan Model, Ph.D. Dissertation, University of Maryland.

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.

Hansen, Nico A. and Anke S. Kessler, 2001a, The Political Geography of Tax H(e)avens and Tax Hells, American Economic Review 91, 1103-1115.

Hansen, Nico A. and Anke S. Kessler, 2001b, (Non-)Existence of Equilibria in Multi-Community Models, Journal of Urban Economics 50, 418-435.

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.

Hindriks, Jean, 2001, Mobility and Redistributive Politics, Journal of Public Economic Theory 3, 95-120.

Ioannides, Yannis M, 2004, Neighborhood income distributions, Journal of Urban Economics 56, 435-457.

Keely, Louise C., 2004, A Theory of Community Formation, Mimeo, University of Wisconsin-Madison.

Kessler Anke S. and Christoph Lülfesmann, 2005, Tiebout and redistribution in a model of residential and political choice, Journal of Public Economics 89, 501-528.

Konishi, H., 1996, Voting with Ballots and Feet: Existence of Equilibrium in a Local Public Good Economy, Journal of Economic Theory 68, 480-509.

Nechyba, Thomas J., 1997, Existence of equilibrium and stratification in local hierarchical Tiebout economies with property taxes and voting, Economic Theory 10, 277-304.

Oates, Wallace E., 1972, Fiscal Federalism, New York: Harcourt Brace Jovanovich.

Rhode, Paul W. and Koleman S. Strumpf, 2003, Assessing the Importance of Tiebout Sorting: Local Heterogeneity form 1850 to 1990, American Economic Review 93, 1648-1677.

Rose-Ackerman, Susan, 1979, Market Models of Local Government: Exit, Voting, and the land market, Journal of Urban Economics 6, 319-337.

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.

Schmidheiny, Kurt, 2002, Income Stratification in Multi-Community Models, VWI Discussion Paper 15-02, University of Bern.

Schmidheiny, Kurt, forthcoming, Income Segregation and Local Progressive Taxation: Empirical Evidence from Switzerland, Journal of Public Economics.

Stone, J., 1954, Linear Expenditure Systems and Demand Analysis. An Application to the Pattern of British, Demand, Economic Journal 64, 511-27.

Tiebout, Charles M., 1956, A Pure Theory of Local Expenditures, Journal of Political Economy 64, 416-424.

Westhoff, F., 1977, Existence of Equilibria in Economies with a Local Public Good, Journal of Economic Theory 14, 84-112.