Submission Number: PET11-11-00309

Head taxes vs income taxes in a Tiebout model with crowding types

Francisco Martinez mora
University of Leicester

Abstract

This paper studies a two-district Tiebout model with housing markets and crowding types to re-evaluate the optimality properties of equilibrium district compositions induced by head and income taxes. When local governments can adjust taxes to crowding types, it is shown that both head taxes and income taxes can lead to the optimal distribution of households across districts. When governments cannot tax-discriminate across crowding types, both head and income taxes distort the distribution of households across districts. Interestingly, in some cases, income taxes lead to smaller welfare losses than the alternative, as they (partially) internalise the emerging location externalities. The results cast doubts over previous conclusions in the literature such as the superiority on efficiency grounds of head over income taxes, the incompatibility between income redistribution and location optimality, the benefit view of head taxes, or the efficiency enhancing properties of zoning regulations.

I am very grateful for many useful comments to Subir Bose, Phillipe De Donder, Gianni De Fraja, Íñigo Iturbe, Miltos Makris, Clara Ponsatí, Javier Rivas, to seminar attendants at the Institut d   Economia de Barcelona (IEB) and to the Department of Economics at Universitat Autonoma de Barcelona (UAB) where part of this research was completed. All remaining errors are my own responsibility.
Submitted: March 16, 2011.

Head taxes vs income taxes in a Tiebout model with crowding types*

Francisco Martínez-Mora
University of Leicester and FEDEA

February 18, 2011

Abstract

This paper studies a two-district Tiebout model with housing markets and crowding types to re-evaluate the optimality properties of equilibrium district compositions induced by head and income taxes. When local governments can adjust taxes to crowding types, it is shown that both head taxes and income taxes can lead to the optimal distribution of households across districts. When governments cannot tax-discriminate across crowding types, both head and income taxes distort the distribution of households across districts. Interestingly, in some cases, income taxes lead to smaller welfare losses than the alternative, as they (partially) internalise the location externalities. The results cast doubts over previous conclusions in the literature such as the superiority on efficiency grounds of head over income taxes, the incompatibility between income redistribution and location optimality, the benefit view of head taxes, or the efficiency enhancing properties of zoning regulations.

Key-words: Tiebout, crowding types, head tax, income tax, segregation.
JEL classification numbers:

*I am very grateful for many useful comments to Subir Bose, Phillipe De Donder, Gianni De Fraja, Íñigo Iturbe, Miltos Makris, Clara Ponsatí, Javier Rivas, to seminar attendants at the Institut d'Economia de Barcelona (IEB) and to the Department of Economics at Universitat Autonoma de Barcelona (UAB) where part of this research was completed. All remaining errors are my own responsibility.

1 Introduction

One of the most active areas of research in urban public finance concerns the normative properties of alternative local tax systems.[1] The main distinguishing feature of that literature with respect to the general analysis of optimal taxation is the choice tax-payers have among a number of distinct taxing jurisdictions. Two efficiency questions emerge in that context: whether the equilibrium allocation of households to districts is efficient and whether the levels of local public good provision and taxation locally selected are efficient.

In his seminal contribution, Tiebout (1956) suggested that head taxes would lead to an efficient sorting of households into homogenous jurisdictions, which would therefore select the efficient level of local public good unanimously. Tiebout did not formalise his argument, which thus became the Tiebout hypothesis. Bewley (1981) among others formalised the result, providing a set of conditions under which the Tiebout hypothesis holds. These included at least as many districts as household types, a technology of production of the local public good exhibiting constant and anonymous marginal costs and the use of local head taxes.[2] In a recent contibution, Calabrese et al. (2010) extended the result to a more realistic framework with housing markets and a smaller number of exogenous districts than of household types.

Whereas head taxes are optimal in that setting, local income taxes are not (Wildasin, 1986; Goodspeed, 1989, 1995): poor households face lower tax prices (that is, they receive greater amounts of spending per unit of tax paid) which allows them to live in richer districts than they would otherwise do and to benefit from the induced redistribution. That distorts residential choices and leads to aggregate welfare losses.[3]

[1] Different aspects of that literature have been surveyed by Zodrow and Mieszkowski (1989), Ross and Yinger (1999), Nechyba (2002), Scotchmer (2002).and Epple and Nechyba (2004).

[2] The Tiebout hypothesis has also been formalised in the slightly different framework of club theory. See Wooders (1978) for the case with anonymous crowding and footnote 9 below.

[3] On a related matter, two results in the literature suggest that local income taxes could be *less distortionary* than property taxes. On the one hand, Calabrese et al. (2010) compare aggregate welfare in two cases: under centralised school finance, where taxes and spending are identical in all jurisdictions, and under local (decentralised) finance. They show that the welfare losses caused by housing market distortions under local property taxation (with respect to the efficient head-tax equilibrium) are *large* and can easily offset the welfare gains stemming from Tiebout-matching (with respect to the equilibrium

A crucial assumption behind all these results is that households are of a single crowding type, that is to say, that households impose identical congestion costs in the production of the local public good. Such assumption is sometimes difficult to justify, especially if the local public good considered is schooling.[4] It is therefore important to investigate the sensitivity of these results to the presence of non-anonymous crowding.

This paper studies a two-district Tiebout model with housing markets and crowding types to re-evaluate the optimality properties of the equilibrium community compositions induced by head and income taxes. The model represents an urban area divided into two jurisdictions (school-districts) with exogenously given boundaries. Districts provide tax-funded public education to residents and households choose where to live, a decision that subsumes the choice of school.[5] Furthermore, local governments provide the optimal level of school quality which allows to focus the analysis on locational optimality.[6] In the normative analysis, the model adopts a utilitarian approach, defining the Social Welfare Function (SWF) as an unweighted sum of utility in the economy.[7]

The main distinguishing feature with respect to previous analyses is the joint consideration of housing markets and non-anonymous crowding. Households differ along three dimensions: income, tastes (i.e. the offspring's ability to benefit from school quality) and crowding costs (i.e. the cost of providing her with one unit of school quality).[8] These characteristics are exogenously

---

emerging under centralised school finance). On the other hand, Goodspeed (1989) compares aggregate welfare when local governments use head and income taxes. His main result is that the welfare losses income taxes cause are *small* if the share of income spent in the local public good is small with respect to that spent in housing.

[4]Notice for example that while some households have no children at school age, others have different number of children, or that the costs of education vary across children of different ages, ability and behaviour.

[5]The case of local public education is used in the exposition. Of course, the analysis applies to other local public goods for which the technological assumptions presented below are reasonable.

[6]Most of the literature, like this paper, focuses on one of the two efficiency questions. A recent exception to this norm is Calabrese et al. (2010). In that paper, welfare losses stemming from voting distortions are small. Goodspeed (1989) notes that the main critic to the local use of income taxation concerns the distortions it introduces in the distribution of the population across districts.

[7]That is the reason to speak about optimality instead of about efficiency.

[8]Most previous analyses in urban public finance consider models in which households differ along a single-dimension (e.g. Epple et al., 1984, 1993; Bénabou, 1996; Nechyba

given. Likewise, districts have an exogenous and fixed supply of homogenous housing units, which avoids well-known existence problems illustrated, for example, by Rose-Ackerman (1979) or Epple et al. (1984).[9]

The analysis considers two situations: one where local governments observe households crowding characteristics and are allowed to levy personalised taxes; another in which, either for information constraints or for political reasons, they are not able to do so. In the first case, the following results emerge. On the one hand, uniform head taxes are inefficient, making personalised ones necessary to achieve optimality. This result is in line with the previous literature (e.g. de Bartolome, 1990; Schwabb and Oates, 1991; Brueckner and Lee, 1989, Conley and Wooders, 1998).[10] Interestingly, however, in a model with housing markets, these need neither cover the marginal cost of entry into a district, nor punish households with lower ability children, as previous results suggested. On the contrary, the extra taxes households of different types pay in the urban area with respect to the suburbs must provide them with the correct *relative location incentives*. Hence, a menu of many different head-tax combinations can lead the economy to an optimal outcome, which opens the door for personalised head taxes to effect some redistribution across crowding types.

On the other hand, a system of local income taxes can always implement the optimal equilibrium. There exist multiple combinations of income tax rates simultaneously satisfying the public sector budget constraints and providing households with the optimal location incentives. This result holds under a mild condition but is not sufficient, for the tax rates profiles may not fulfill two necessary single-crossing conditions. Nevertheless, it is always possible to decentralise the optimal equilibrium by combining uniform income taxes with lump-sum transfers or taxes. At most, the necessary

1999). There are however exceptions in which households differ, additionally, along a taste parameter (Epple and Platt, 1998; Kessler and Lülfessmann, 2005; Schmidheiny, 2006a, 2006b), a productivity one such as the offspring ability (Epple and Romano, 2003), or both (Wildasin, 1986).

[9] Club models incorporate more general demographic frameworks but not a explicit representation of land markets. Also, they are not constrained by an exogenous number of districts whose boundaries stem from some historical process, as models in urban public finance, but allow clubs to form endogenously. A joint review of the two literatures can be found in Scotchmer (2002). A succint review of the club literature on the Tiebout model is in Wooders (1999).

[10] Notice that these personalised taxes depend on crowding types (which may be publicly observable) and not on tastes (which are not). See Conley and Wooders (1998).

single-crossing conditions may require the tax rates to be equated across districts, thereby limiting the amount of redistribution that is compatible with location optimality.[11]

A counter-intuitive result emerges in the second case, that is, when local governments cannot tailor taxes to household types. In those cases, it is not possible to unambiguously rank head and income taxes according to the distortions they introduce in the location equilibrium. This result clashes with the benefit view of local head taxes: in the more general framework this paper studies, uniform head taxes not only induce an inefficient distribution of households across districts but may also be more distortionary than an ability-to-pay tax such as a proportional income tax. The intuition is the following: whereas anonymous head taxes do not affect households relative location incentives, income taxes do generate differences in entry prices across crowding types. The reason is that, in general, marginal (or cut-off) households of each type living in the rich area (i.e. the lowest income ones of their type residing there) have different incomes in equilibrium and, hence, pay different entry prices. Under some circumstances, such differences will imperfectly internalise the location externalities created by the marginal residents of the rich district.[12]

Overall, these results cast doubts over important previous conclusions in the urban public finance literature, such as the superiority on efficiency grounds of head over income taxes, the incompatibility between income redistribution and location optimality, the benefit view of head taxes, or the efficiency enhancing properties of zoning regulations (Hamilton, 1975; Calabrese et al., 2007). Remarkably, they are derived in a model that purposedly favours head taxes: because the utility specification is quasi-linear in private consumption, the planner has no preference for redistribution; on the other hand, the analysis focuses on solutions to the Social Planner Problem where

[11] Wildasin (1986) showed that a set of personalised income tax rates adjusted so that the tax payment of every household is equal to its marginal congestion cost in every district would achieve the normative objective. That solution, which is valid when all types are present in all districts in the optimal outcome, does not necessarily extend to the current setting where only two indifferent types will exist in equilibrium. In general, the proposed taxes must also provide the optimal location incentives for types that concentrate in a subset of locations. This further restricts the set of income taxes that implement the optimal solution. In the current context in particular, the proposed taxes would need to satisfy the relevant single-crossing conditions.

[12] This possibility was conjectured by Goodspeed (1989, footnote 12).

income segregation is efficient for households of the same type.[13]

The rest of the paper is organised as follows. The next section presents the model. Section 3 derives the housing markets equilibrium condition, which restricts the set of allocations of households to districts that may be sustained as a segregated equilibrium. The following section obtains the optimal allocation. The analysis then turns to the comparison between head and income taxes. Section 5 studies the case where governments can use personalised taxes, whereas section 6 does the same for the case in which taxes are anonymous. Finally, section 7 offers some concluding remarks and suggests questions for future research.

2 The model

A metropolitan area is divided into two school districts (or communities) with fixed boundaries, labeled as the urban area (or central district) and the suburbs, and indexed with $j = u, s$. Districts provide tax-funded, tuition-free public education of homogeneous quality (e_j) to all their school-aged residents[14]. With no loss of generality, for the sake of definiteness and following the typical pattern of many European cities, the urban school corresponds to that of higher quality.[15]

A population of households with mass normalised to 1 lives in the city. Every household has a school-aged child. Households differ continuously according to income and discretely along two additional dimensions: the cost of providing them with a given level of school quality and the benefit they receive from it. To present results in the simplest possible manner, I consider that two cost-benefit types of households coexist at a time. Household types are indexed with $i = 1, 2$, with households of type 1 corresponding to the low cost type;[16] $\gamma \in (0, 1)$ measures the proportion of type 1 households in the

[13] This restricts preferences and the weights of the social welfare function in ways that will be made explicit in the analysis.

[14] Attendance to school is assumed compulsory. The probability of admission to a local school is equal to one if the household resides in the district and equal to zero otherwise.

[15] For simplicity sake, private schools are excluded from the analysis. Note, however, that the choice of private schools over the public educational sector by some households automatically tranforms them into a different crowding and benefit type from the perspective of the local government. Intuitively, the presence of private education would only reinforce the results of this paper.

[16] These definitions will be stated formally once some additional notation is introduced.

population.

Household income is denoted with $y \in D \equiv \left[\underline{y}, \overline{y}\right]$, and is distributed in the population according to $\Phi_i(y) \in [0,1]$; $i = 1,2$. Income distribution functions are continuous, strictly increasing in all their support D and have densities $\phi_i(y) = \Phi_i'(y)$; $i = 1,2$. The total (and average) income is:

$$Y = \gamma \int_{\underline{y}}^{\overline{y}} y\phi_1(y)\,dy + (1-\gamma)\int_{\underline{y}}^{\overline{y}} y\phi_2(y)\,dy, \tag{1}$$

whilst district j's income distribution functions are given by $\Phi_j(y) \in [0,1]$, $i = 1,2$.

Districts have a fixed (and identical) supply of homogeneous houses, denoted H_j, the total supply being equal to the mass of households that resides in the city: $H_u = H_s = 1/2$.[17] Absentee landlords lend these houses out to households in exchange of a rent. To avoid a source of indeterminacy, I normalise the suburbs' rent away to zero.[18] Because housing is supplied inelastically, school quality and tax differentials will capitalise into housing prices in equilibrium. Equilibrium in the housing markets will therefore entail the existence of a rent premium in the good school district, r_u.[19]

Preferences are defined over a private composite good (the numeraire), x, and the offspring's future income, h.[20] The latter, in turn, depends on the quality of education received, e, and the availability of home inputs, y. A twice continuously differentiable utility function $U_i(x, e; y)$ represents these preferences. Following de Bartolome and Ross (2004), I adopt a quasi-linear and separable specification of utility:

$$U_i(x, e; y) = x + h_i(e; y); \quad i = 1,2, \tag{2}$$

where h is monotonically increasing in its arguments and strictly concave in e. The choice of a quasi-linear utility function not only simplifies the analysis but also completely separates efficiency and equity considerations: because

[17] This simplifying assumption does not affect any of the results presented below but avoids the need to determine whether it is optimal to have the larger district with the better school or viceversa (see Calabrese et al., 2010).

[18] This implies that the opportunity cost of the residential use of land –the value of industrial or agricultural alternative uses– is normalised to zero.

[19] The rent premium could be negative in equilibrium which would indicate the negative capitalisation of higher taxes.

[20] Because houses are homogeneous they can be excluded from the preference relation.

the marginal utility of private consumption is constant and equal to one (i.e. households are risk neutral), the level of aggregate welfare attained by a particular allocation of households to districts is invariant to the distribution of private consumption and taxes in the population. Preferences also satisfy:

Assumption 1 *Education is a normal good.*

Assumption 1 implies a positive income elasticity of the demand for school quality, which agrees with the empirical evidence (see for example Ross and Yinger, 1999). This assumption restricts quasi-linear preferences. In particular, it requires home and school inputs to be complements in the production of future income (or human capital): $\partial^2 h(e, y) / \partial e \partial y > 0$.[21]

The technology of production of school quality is identical across districts and linear. It is described by the cost function:

$$C\left(n_j^1, n_j^2, e_j\right) = \left(n_j^1 c_1 + n_j^2 c_2\right) e_j; \qquad c_1 \leq c_2$$

where n_j^i stands for the mass of households of type i living in the district, and $c_1 \leq c_2$ means that households of type 2 are (weakly) more costly to educate, i.e. impose greater congestion costs. Let Δ_c denote the crowding cost differential $\Delta_c = c_2 - c_1 \geq 0$. The analysis considers cases where crowding and benefit types are correlated as follows:

Assumption 2 *Type 1 households impose weakly smaller congestion costs on the production of education: $c_1 \leq c_2$, and derive greater benefits from school quality: $\partial h_1(e, y) / \partial e \geq h_2(e, y) / \partial e$, $\forall e$.*

Type 1 households may thus be interpreted as having higher ability offsprings and the cost difference can be seen as the result of differential peer effects operating at the school (or district) level.[22]

In order to focus the analysis on locational efficiency, I assume local governments behave as follows:

[21] This could be due, for example, to better labour market networking of better-off parents.

[22] I will also explain how results change when the low cost type derives smaller benefits from education ($\partial y_1'(e, y) / \partial e < \partial y_2'(e, y) / \partial e$). In that case, type 1 households may be interpreted as childless ones who do not impose a cost on public education but still derive benefits from better school quality as they care for the behaviour of their young neighbours.

Assumption 3 *Local governments provide the (locally) efficient level of school quality, given the local population of pupils, and balance their budget.*

That is to say, levels of local spending satisfy the Samuelsonian condition for publicly provided public goods, i.e. are such that the Average Marginal Rate of Substitution between school quality and numeraire consumption ($AMRS_{ex}$) equals the marginal cost of school quality in the district ($n_1 c_1 + n_2 c_2$).[23]

Local governments fund their spending with some form of taxation captured by the tax-bill function $\tau_i \left(e_j, n_j^1, n_j^2, y\right)$. The tax-bill function must not only meet the local budget constraint but also satisfy the feasibility constraint $\tau_i \left(e_j, n_j^1, n_j^2, y\right) \leq y$ for all $y \in \left[\underline{y}, \bar{y}\right]$, i.e. household tax payments must be smaller than household income.

The indirect utility function of a household of type i that has income y and lives in district j can now be formally stated:

$$v_j^i \left(e_j, \tau \left(e_j, n_j^1, n_j^2, y\right), r_j; y\right) = y - \tau_i \left(e_j, n_j^1, n_j^2, y\right) - r_j + h_i \left(e, y\right)$$

A useful tool in the analysis that follows are the so-called bid-rent functions, which I define next[24]:

Definition 1 *The* bid-rent function r_i^τ *($i = u, s$; $\tau = H, I$) provides the maximum amount of the numeraire a household of income y and type $i = 1, 2$ is willing to pay as rent premium in the urban area, given the tax system τ and the vector of school qualities and tax bills $(y, e_u, e_s, \tau_s, \tau_u)$. Bid rent functions are obtained by setting $r_s = 0$ in the indifference condition:*

$$v_u^i \left(e_u, \tau_u, r_u, y\right) = v_s^i \left(e_s, \tau_s, r_s, y\right). \tag{3}$$

Bid-rent functions provide the value of the urban area rent premium that makes households of type i and income y indifferent among the two areas. They thus depend on the tax system considered.[25]

[23] This condition is derived formally in the next section when the Social Planner Problem is solved.

[24] For a review of the literature on urban public finance that extensively uses bid-rent functions see Ross and Yinger (1999).

[25] For that reason, I do not give a generic expression for them at this point but delay its formal presentation to the analysis of market equilibrium under the two alternative tax systems considered.

Although the model is static, the sequence of events unfolds in three stages. First, households choose which district to live in and rent a house there. The choice of district is the single strategic decision households make, which they take anticipating (correctly in equilibrium) the vector of local levels of spending and taxation. At this stage, housing markets clear and the rent premium is determined. At the second stage, given the distribution of households across districts, local public spending and taxes are determined by the Samuelsonian efficiency condition and the local governments budget constraints. At the third stage, households send their children to school and consume all available income.[26]

Given the tax system, an equilibrium in this model is an allocation of households to districts, a vector of local tax rates and school qualities, and a value of the urban rent premium satisfying the following conditions:

(E1) *Rational choices*: No household can increase utility by moving into the other school district.

(E2) *Housing markets clearance*: $H_j = n_j^1 + n_j^2$; $j = u, s$.

(E3) *Local governments budget balance*: $\tau\left(e_j, n_j^1, n_j^2, y\right)$ satisfies the local budget constraint of district j.

(E4) *Locally efficient school qualities*: e_j satisfy the Samuelsonian condition given the local population of households.

3 The housing market constraint

Given the objectives of the paper, the analysis focuses on cases where schools are of different quality. In the utilitarian normative framework considered (explained in the next section), the assumption that education is a normal good implies that, in an optimal allocation, households of the same type will be segregated across districts according to income, and that higher income ones will be allocated to the better school district. I call this property *within-types income segregation.*

Definition 2 *An allocation of households to districts satisfies* within-types income segregation (WTS) *if, for any pair of households of the same crowding type but of different income, and who are assigned to different districts, the higher income one lives in the (good) urban area.*

[26] Notice that stages 2 and 3 are degenerate.

Within-types income segregation simplifies the problem by restricting the set of allocations that may potentially be an equilibrium or an optimal solution. In an allocation of households to districts satisfying the WTS property, households of the same type living in the same district belong to a single income interval, and the income intervals containing households of the same type living in each district do not overlap. Therefore, there exists a unique pair of cut-off incomes y_i^τ, $i = 1, 2$, such that households of type i and income above (below) y_i^τ reside in the urban area (the suburbs) in that allocation. Clearly, n_j^i can be rewritten in that case as $n_u^i\left(y_i^\tau\right) = \gamma_i\left(1 - \Phi_i\left(y_i^\tau\right)\right)$ and $n_s^i\left(y_i^\tau\right) = \gamma_i \Phi_i\left(y_i^\tau\right)$ (where $\gamma_1 = \gamma$ and $\gamma_2 = 1 - \gamma$). The urban area land market constraint (which, if satisfied, implies the suburbs land market constraint) can thus be expressed in terms of the cut-off incomes:

$$H_u = \gamma\left(1 - \Phi_1\left(\widetilde{y}_1\right)\right) + (1 - \gamma)\left(1 - \Phi_2\left(\widetilde{y}_2\right)\right) \tag{4}$$

Lemma 1 *Consider an allocation that satisfies WTS. Then, cut-off incomes are linked through a continuously decreasing function z defined on a compact set: $\widetilde{y}_1 = z\left(\widetilde{y}_2\right)$, $\widetilde{y}_2 \in \left[\underline{y}^2, \overline{y}^2\right]$; z is implicitly defined by:*[27]

$$H_u = \gamma\left(1 - \Phi_1\left(z\left(\widetilde{y}_2\right)\right)\right) + (1 - \gamma)\left(1 - \Phi_2\left(\widetilde{y}_2\right)\right) \tag{5}$$

Proof. Establishing that the urban area housing market constraint links $\widetilde{y}_1$ and $\widetilde{y}_2$ through a functional relationship requires showing that, for any value of $\widetilde{y}_2$, there exists a unique value of $\widetilde{y}_1$ satisfying (4). Suppose $\gamma \in [1/2, 1)$; then, $0 \le (1 - \gamma)(1 - \Phi_2(y)) \le 1/2 \; \forall y \in \left[\underline{y}, \overline{y}\right]$. Given that $\gamma\left(1 - \Phi_1\left(\overline{y}\right)\right) = 0$ and $1/2 \le \gamma\left(1 - \Phi_1\left(\underline{y}\right)\right) < 1$, continuity and strict monotonicity of the income distribution functions and the intermediate value theorem imply that, for any $\widetilde{y}_2 \in \left[\underline{y}, \overline{y}\right]$, there exists a unique $\widetilde{y}_1 \in \left[\underline{y}, \overline{y}\right]$ such that the land market constraint (4) holds. Suppose instead that $\gamma \in (0, 1/2]$; then $0 \le (1 - \gamma)(1 - \Phi_2(y)) \le 1 \; \forall y \in \left[\underline{y}, \overline{y}\right]$. Hence, by continuity and strict monotonicity of Φ_2 there exists a unique income $\underline{y}^2 = y$ such that $(1 - \gamma)\left(1 - \Phi_2\left(\overline{y}\right)\right) = 1/2$, and another $\overline{y}^2 = y$ such that $(1 - \gamma)\left(1 - \Phi_2\left(\overline{y}\right)\right) = 1/2 - \gamma$. Moreover, given that $0 \le \gamma(1 - \Phi_1(y)) \le \gamma \; \forall y \in \left[\underline{y}, \overline{y}\right]$, again the continuity and strict monotonicity of Φ_1 and Φ_2 and the intermediate value theorem imply that, for any $\widetilde{y}_2 \in \left[\underline{y}^2, \overline{y}^2\right]$, there exists a unique $\widetilde{y}_1 \in \left[\underline{y}, \overline{y}\right]$ such that the land market constraint (4) holds. Furthermore, these two

[27] For uneven district sizes, the domain of $z(y)$ depends on the size of the urban district relative to the proportion of each type of households in the population.

properties guarantee that z is continuous and decreasing. Finally, apply the implicit function theorem to (4) in order to find the derivative:

$$\frac{d\widetilde{y}_1}{d\widetilde{y}_2} = -\frac{(1-\gamma)\,\phi_2(\widetilde{y}_2)}{\gamma\phi_1(z(\widetilde{y}_2))} < 0. \tag{6}$$

■

Function z is continuous and decreasing, which implies that cut-off incomes move continuously in opposite directions. This implies, for example, that as $\widetilde{y}_2$ goes up above the level inducing perfect income segregation, $\widetilde{y}$, relatively low income households of type 1 replace higher income type 2 households in the urban area.[28] Hereafter, I will characterise community compositions satisfying WTS simply with the type 2 cut-off income $\widetilde{y}_2$.

Every allocation satisfying WTS can be linked to a unique vector of local spending and tax policies. By assumption 3, local governments' provide the efficient level of school quality,[29] given the demographic composition of each district. Because the human capital production function displays decreasing marginal returns to school quality, the optimal levels of school quality are unique and hence can be written as a function of $\widetilde{y}_2$: $e_j(\widetilde{y}_2)$. In turn, given school qualities, the local governments' budget constraints allow to express local tax payments as a function of $\widetilde{y}_2$ as well. Therefore, to every allocation of households to districts satisfying WTS $\widetilde{y}_2$ corresponds a unique vector of local policy variables: $\mathbf{\Pi}^{\tau}(\widetilde{y}_2) = [e_u(\widetilde{y}_2)\ \ e_s(\widetilde{y}_2)\ \ \tau_u(\widetilde{y}_2)\ \ \tau_s(\widetilde{y}_2)]$, $\tau = H, I$, where τ_j represents the tax-bill of district j under head taxation, T_j, or its tax rate under income taxation, t_j.

Restricting attention to household allocations exhibiting WTS, define:

Definition 3 Cut-off income bid rent functions, *denoted* $\rho_i^{\tau}(\widetilde{y}_2)$, $i = 1, 2$, $\tau = T$, t, *provide for any type 2 cut-off income* $\widetilde{y}_2$, *the maximum rent premium households of each type with the corresponding cut-off income are willing to pay for a house in the urban area, when local policy variables are set at* $\mathbf{\Pi}^{\tau}(\widetilde{y}_2)$. *Cut-off income bid rent functions are implicitly defined by setting* $r_s = 0$ *in the indifference condition:*

$$v_u^i(e_u, \tau_u, \rho_i^{\tau}; y_i^{\tau}) = v_s^i(e_s, \tau_s, r_s; y_i^{\tau})\,; \ \ i = 1, 2. \tag{7}$$

where local policies are those in $\mathbf{\Pi}^{\tau}(\widetilde{y}_2)$.

[28] Defining the "amount" of income mixing as the mass of households with incomes such that households of the other type with the same income live in the other district, it is also clear that income mixing rises as $\widetilde{y}_2$ gets away from $\widetilde{y}$.

[29] These conditions are stated formally in the next section.

Importantly, note that because $z(\widetilde{y}_2)$, $e_j(\widetilde{y}_2)$ and $\tau_j(\widetilde{y}_2)$ are continuous, cut-off bid-rent functions are continuous too.

4 The optimal allocation

This section characterises the solution to the Social Planner Problem (SPP). I adopt a utilitarian approach and define the Social Welfare Function (SWF) as the unweighted sum of utility in the economy. I consequently speak of optimality instead of efficiency.[30] The SWF includes the utility of the absentee landowners, which is assumed linear in the private composite good. The planner is allowed to use head taxes, which may differ across household types, and can also make transfers to households and landlords, denoted $R_i(y)$ and R_L, respectively.[31] The indirect utility function is thus:

$$V_j^i\left(e_j, T_j^i, r_j, y, R_i(y)\right) = y + R_i(y) - T_j^i - r_j + h_i(e_j, y), \tag{8}$$

while the SWF is:

$$\begin{aligned}
&\int_{\underline{y}}^{\widetilde{y}_1} y + R_1(y) - T_s^1 - r_s + h_1(e_s, y)\,\gamma\phi_1(y)\,dy + \\
&\int_{\widetilde{y}_1}^{\overline{y}} y + R_1(y) - T_u^1 - r_u + h_1(e_u, y)\,\gamma\phi_1(y)\,dy + \\
&\int_{\underline{y}}^{\widetilde{y}_2} y + R_2(y) - T_s^2 - r_s + h_2(e_s, y)\,(1-\gamma)\,\phi_2(y)\,dy + \\
&\int_{\widetilde{y}_2}^{\overline{y}} y + R_2(y) - T_u^2 - r_u + h_2(e_u, y)\,(1-\gamma)\,\phi_2(y)\,dy + \\
&+\left[R_L + r_s H_s + r_u H_u\right].
\end{aligned} \tag{9}$$

The SWF is maximised with respect to $e_j, y_i, T_j^i, r_j, R_i(y), R_L$ subject to ten constraints, which include six nonnegativity constraints ($e_j \geq 0$, $T_j^i \geq 0$), the

[30] The results in the paper emerge when it is optimal to have differentiated and segregated districts, as it is the case with the unweighted utilitarian SWF. Instead, one could assume a weighted SWF and consider sets of weights for which maximising the SWF requires districts providing different levels of school quality and households segregating in the form described by WTS.

[31] This is only for completeness; quasi-linearity of the utility function implies that such transfers do not affect aggregate welfare if all the weights in the utility function are equal to one, as it is assumed below.

housing market constraint and two local governments budget constraints.[32] The housing market constraint has λ_h as its multiplier and, by Lemma 1, can be written as:

$$\widetilde{y}_1 - z\left(\widetilde{y}_2\right) = 0. \tag{10}$$

The two local budget constraints have associated multipliers λ_j and are given by:

$$e_j\left[c_1 n_j^1\left(\widetilde{y}_1\right) + c_2 n_j^2\left(\widetilde{y}_2\right)\right] - T_j^1 n_j^1\left(\widetilde{y}_1\right) + T_j^2 n_j^2\left(\widetilde{y}_2\right) = 0;\ j = u, s. \tag{11}$$

The optimal demographic composition of districts is determined by the FOCs on the cut-off incomes y_i. These yield the following *Marginal Social Value* functions:

$$MSV_i\left(y\right) = h_i\left(e_u, y\right) - h_i\left(e_s, y\right) + c_i\left[e_s - e_u\right];\ i = 1, 2. \tag{12}$$

Marginal Social Value functions provide the marginal impact on social welfare of moving a household of type i with cut-off income y_i from the suburbs to the urban area.[33] The following proposition characterises the solution to the SPP.[34]

Proposition 1 *A solution to the Social Planner's Problem with $e_u^* > e_s^*$ exhibits WTS. Furthermore:*
i) In an interior solution, cut-off incomes $\widetilde{y}_1^*$, $\widetilde{y}_2^*$ satisfy: $MSV_1\left(y_1^*\right) = MSV_2\left(\widetilde{y}_2^*\right)$, or

$$\left[h_1\left(e_u^*, y_1^*\right) - h_1\left(e_s^*, y_1^*\right)\right] - \left[h_2\left(e_u^*, y_2^*\right) - h_2\left(e_s^*, y_2^*\right)\right] = \Delta_c\left[e_s^* - e_u^*\right], \tag{13}$$

the housing market constraint, $\widetilde{y}_1^* = z\left(\widetilde{y}_2^*\right)$ and, if assumption 2 holds, $\widetilde{y}_1^* <$

[32] The tenth constraint requires the transfers' budget to balance :

$$\gamma\int_{\underline{y}}^{\overline{y}} R_1\left(y\right)\phi_1(y)dy + (1-\gamma)\int_{\underline{y}}^{\overline{y}} R_2\left(y\right)\phi_2(y)dy + R_L = 0$$

and has multiplier λ_R.

[33] I express MSV functions as the social value of moving a household from the suburbs to the urban area (i.e. the value of *decreasing* the cut-off income of a particular type) to facilitate comparability with the bid-rent functions and only for expositional purposes.

[34] In an interior solution the two types of households are present in the two districts. In a corner solution, one of them concentrates in one of the districts. If $\gamma \geq 1/2$, then all type 2 households live in the suburbs, $y_1^* \geq \underline{y}$, $y_2^* = \overline{y}$; if $\gamma < 1/2$, then all type 1 households live in the urban district, $y_1^* = \underline{y}$, $y_2^* < \overline{y}$.

$\widetilde{y}_2^*$.
ii) In a corner solution, $\widetilde{y}_1^*$, $\widetilde{y}_2^*$ *satisfy:* $MSV_1(y_1^*) \geq MSV_2(\widetilde{y}_2^*)$, *the housing market constraint,* $\widetilde{y}_1^* = z(\widetilde{y}_2^*)$ *and, if assumption 2 holds,* $\widetilde{y}_1^* < \widetilde{y}_2^*$.
iii) School qualities satisfy the Samuelsonian conditions:

$$\begin{aligned} c_1 n_u^1(\widetilde{y}_1) + c_2 n_u^2(\widetilde{y}_2) &= \gamma \int_{\widetilde{y}_1}^{\overline{y}} \frac{\partial h_1(e_u^*, y)}{\partial e_u} \phi_1(y) dy \\ &\quad + (1-\gamma) \int_{\widetilde{y}_2}^{\overline{y}} \frac{\partial h_2(e_u^*, y)}{\partial e_u} \phi_2(y) dy \end{aligned} \tag{14}$$

$$\begin{aligned} c_1 n_s^1(\widetilde{y}_1) + c_2 n_s^2(\widetilde{y}_2) &= \gamma \int_{\underline{y}}^{\widetilde{y}_1} \frac{\partial h_1(e_s^*, y)}{\partial e_s} \phi_1(y) dy \\ &\quad + (1-\gamma) \int_{\underline{y}}^{\widetilde{y}_2} \frac{\partial h_2(e_s^*, y)}{\partial e_s} \phi_2(y) dy. \end{aligned} \tag{15}$$

Proof. It is straightforward to check that normality of education implies that MSV functions are increasing in income, which confirms that the solution satisfies WTS. The FOCs on T_j^i are:

$$-n_j^i(y) + \lambda_j n_j^i(y) = 0;\ j = u, s,\ i = 1, 2, \tag{16}$$

which yield $\lambda_j = 1$. Using that initial result and equation (6), the ones corresponding to the cut-off incomes simplify to:

$$[h_1(e_u, y) - h_1(e_s, y) - c_1(e_u - e_s)]\, \gamma \phi_1(\widetilde{y}_1) = -\lambda_h \tag{17}$$

$$[h_2(e_u, y) - h_2(e_s, y) - c_2(e_u - e_s)]\, (1-\gamma) \phi_2(\widetilde{y}_2) = -\frac{\lambda_h (1-\gamma) \phi_2(\widetilde{y}_2)}{\gamma \phi_1(\widetilde{y}_1)} \tag{18}$$

which together imply that (13) must hold in an interior solution. Next, note that, given $e_u > e_s$, Assumption 2 entails $MSV_1(\widetilde{y}) > MSV_2(\widetilde{y})$, where recall $\widetilde{y}$ satisfies $\widetilde{y} = z(\widetilde{y})$. The marginal social value functions are increasing so that the optimal cut-off incomes must satisfy $\widetilde{y}_1 < \widetilde{y}$ and $\widetilde{y} < \widetilde{y}_2$. Finally, the FOCs corresponding to the school quality variables e_u and e_s yield the usual Samuelsonian conditions (14) and (15). ■

Income mixing is optimal in a model with differential crowding costs: households of type 1 with incomes between $\widetilde{y}_1^*$ and $\widetilde{y}_2^*$ should live in the urban area instead of households of type 2 and identical income. The reasons are that they impose smaller congestion costs on and derive greater benefits from the district offering higher quality schooling.[35]

5 Decentralising the optimal allocation

In order to decentralise the optimal allocation as a market equilibrium, local tax variables must satisfy the local governments' budget constraints and provide households with the adequate location incentives. That is, they must ensure that households derive (weakly) higher utility in the district they are assigned to in the optimal allocation. To ensure the induced equilibrium allocation of households to districts exhibits WTS as required by optimality, the bid-rent functions need to be increasing in income and the optimal cut-off types must be made indifferent between the two residential alternatives; the latter will generate a *location-incentive contraint*.

5.1 Personalised head taxes

With differentiated head taxes, the local budget constraints are

$$E_j = T_j^1 n_j^1 + T_j^2 n_j^2,\ j = u, s, \tag{19}$$

where $E_j = e_j \left(n_j^1 c_1 + n_j^2 c_2 \right)$, whilst the indirect utility functions are:

$$v_j^i \left(e_j, T_j^i, r_j, y \right) = y - T_j^i - r_j + h_i \left(e_j, y \right) \tag{20}$$

Normalising the rent of the suburbs to $r_s = 0$, one can derive the head-tax bid-rent functions (denoted r_i^H) from the indifference condition (3):

$$r_i^H \left(y, e_u, e_s, T_s^i, T_u^i \right) = h_i \left(e_u, y \right) - h_i \left(e_s, y \right) + T_s^i - T_u^i,\ i = 1, 2. \tag{21}$$

Lemma 2 proves that the head-tax bid-rent functions are increasing in income and that the relevant single-crossing conditions hold in this case.

[35] If households preferences satisfy $\partial y_1' \left(e, y \right) / \partial e < \partial y_2' \left(e, y \right) / \partial e$, while some income mixing will be optimal (except in special cases) which of the cut-off incomes should be smaller is ambiguous. In that case, the two forces work in different directions: while the lower costs of educating type 1 children tends to make optimal that $\widetilde{y}_1 < \widetilde{y}_2$, the greater benefit type 2 households derive from school quality has the opposite effect.

Lemma 2 *Suppose that $e_u > e_s$, then, the head-tax bid-rent functions are increasing in income, which implies that households induced preferences satisfy the following single-crossing conditions:*

$$\begin{aligned} v_i^s(e_s, T_s^i, y_i) &= v_i^u\left(e_u, T_u^i, r_u; y_i\right) \Rightarrow \\ v_i^s\left(e_s, T_s^i, y\right) &< v_i^u\left(e_u, T_u^i, r_u; y\right); \forall y > y_i \\ v_i^s\left(e_s, T_s^i, y\right) &> v_i^u\left(e_u, T_u^i, r_u; y\right); \forall y < y_i,\ i = 1, 2. \end{aligned} \tag{22}$$

A household's tax burden does not depend on income; the normality of education thus ensures that head-tax bid-rent functions are increasing in income. That property, in turn, implies that if households of type i and income y_i^H are indifferent between the two districts, then households of the same type and higher (lower) income strictly prefer the urban area (the suburbs).

Consider an allocation of households to districts characterised by WTS and with type 2 cut-off income y and let

$$\Delta_1^h(y) = h_1\left(e_u(y), z(y)\right) - h_1\left(e_s(y), z(y)\right)$$

$$\Delta_2^h(y) = h_1\left(e_u(y), y\right) - h_1\left(e_s(y), y\right)$$

denote the gap in human capital each type's cut-off household obtains from attending the urban school instead of the suburban one, when school qualities are determined optimally. Cut-off income bid-rent functions, derived from equations (5) and (7), can then be written as:

$$\rho_1^H(y_2) = \Delta_1^h(y_2) + T_s^1 - T_u^1 \tag{23}$$

$$\rho_2^H(y_2) = \Delta_2^h(y_2) + T_s^2 - T_u^2 \tag{24}$$

Proposition 2 below proves that in an *interior market equilibrium* satisfying WTS, the cut-off income bid-rent functions must be equal to each other at y_2. That is, $\rho_1^H(y_2) = \rho_2^H(y_2)$, or:

$$\Delta_1^h(y_2) - \Delta_2^h(y_2) = \left(T_s^1 - T_u^1\right) - \left(T_s^2 - T_u^2\right) \tag{25}$$

The *location-incentives constraint* is obtained by equating the RHS of the market equilibrium condition (25) to the RHS of the efficiency condition (13), which recall characterises interior solutions to the Social Planner Problem:

$$\left(T_u^1 - T_s^1\right) - \left(T_u^2 - T_s^2\right) = \Delta_c\left[e_s(y_2^*) - e_u(y_2^*)\right]. \tag{26}$$

Definition 4 *Let $\Omega^H(y_2^*)$ be the set of all combinations of personalised head taxes T_j^i satisfying the location-incentives constraint (26) and the local budget constraints (19) at y_2^*.*

Clearly, the combination of head taxes covering the marginal cost of admitting a household of a given type in a particular district, $T_j^i = c_i e_j(y_2^*)$, belongs to $\Omega^H(y_2^*)$. In that case:

$$\left(T_s^i - T_u^i\right) = c_i \left[e_s(y_2^*) - e_u(y_2^*)\right],$$

so that the budget and the location-incentives constraints are satisfied. Interestingly, however, many other combinations of differentiated head taxes lead to the optimal equilibrium as well.

Proposition 2 *Consider a solution to the Social Planner Problem with optimal cut-off incomes (y_1^*, y_2^*) and satisfying $e_u^* > e_s^*$. For every combination of head-tax bills in $\Omega^H(y_2^*)$ there exists an equilibrium exhibiting WTS with cut-off incomes $y_1^H = y_1^*$ and $y_2^H = y_2^*$ and rent premium $r^H = \rho_1^H(y_2^*) = \rho_2^H(y_2^*)$. There are infinitely many such combinations.*[36]

Proof. First, notice that the elements of $\Omega^H(y_2^*)$ satisfy a system of three linearly independent equations in four unknowns, so that the system has one degree of freedom and infinitely many solutions. Proving existence requires checking that the four equilibrium conditions (E1)-(E4) hold: (E3) is satisfied by the definition of $\Omega^H(y_2^*)$; (E4) is fulfilled as well by the assumption on the behaviour of local governments (assumption 3), while proposition 1 implies that the housing market constraint (E2) is satisfied for (y_1^*, y_2^*). The rational choices condition (E1), in turn, requires, first, the single-crossing conditions embedded in (22) to hold and, second, the cut-off bid-rent functions to be equal to each other and to the equilibrium housing rent premium r_H at y_2^*: i.e. $r_H = \rho_1^H(y_2^*) = \rho_2^H(y_2^*)$. The former was proved in Lemma 2, whereas the latter is again ensured by the definition of $\Omega^H(y_2^*)$, as its elements satisfy the location incentives constraint (26). ■

The result shows that, in a model with housing markets, while differentiated head taxes are necessary to attain optimality, these need neither cover the marginal cost of entry of a household nor be greater for households

[36] The result applies to a corner solution of the SPP as well because in a corner solution $MSV_1(y_1^*) \geq MSV_2(y_2^*)$ and $\Omega^H(y_2^*)$ imposes that border incomes in the associated market equilibrium satisfy $MSV_1\left(y_1^H\right) = MSV_2\left(y_2^H\right)$.

with lower ability children. Instead, they must provide households with the correct *relative* incentives. The result opens the door for differentiated head taxes to effect some redistribution across types.[37]

5.2 Income taxes

The analysis now turns to the implementation of the optimal outcome under proportional income taxation. Under differentiated income taxation, the local budget constraint of district j is:

$$e_j \left(n_j^1 c_1 + n_j^2 c_2\right) = t_j^1 n_j^1 Y_j^1 + t_j^2 n_j^2 Y_j^2,\ j = u, s. \tag{27}$$

where t_j^i stands for the local income tax rate district j imposes on households of type i. Indirect utility functions are now:

$$v_j^i \left(e_j, t_j^i, r_j; y\right) = y \left(1 - t_j^i\right) - r_j + h_i \left(e_j, y\right), \tag{28}$$

while the income-tax bid rent functions, obtained as before from the indifference condition $v_u^i \left(e_u, t_u^i, r_u; y\right) = v_s^i \left(e_s, t_s^i, r_s; y\right)$, are:

$$r_i^I \left(y, e_u, e_s, t_s^i, t_u^i\right) = h_i \left(e_u, y\right) - h_i \left(e_s, y\right) + y \left(t_s^i - t_u^i\right);\ i = 1, 2. \tag{29}$$

There are two district-level variables whose impact on utility varies with income: school quality and the income tax rate. The former makes richer households willing to pay more than lower income ones for a house in the district offering higher quality of education. The latter makes them willing to pay more for a house located where their group income tax rate is lower. Therefore, the single-crossing conditions will be satisfied if the richer district is able to fund its education spending with lower income tax rates than the poor district. More generally:

Lemma 3 *Suppose $e_u > e_s$. The income-tax bid-rent functions, r_i^I, are increasing in income if and only if*

$$\left[\frac{\partial h_i(e_u, y)}{\partial y} - \frac{\partial h_i(e_s, y)}{\partial y}\right] > \left(t_u^i - t_s^i\right) \quad \forall y \in S,\ i = 1, 2. \tag{30}$$

[37] Notice that the optimal demographic composition of districts could also be implemented by transfer schemes that create the optimal location incentives.

If inequality (30) holds, households induced preferences satisfy:

$$\begin{aligned} v_i^s(e_s, t_s^i, y_i) &= v_i^u\left(e_u, t_u^i, r_u; y_i\right) \Rightarrow \\ v_i^s\left(e_s, t_s^i, y\right) &< v_i^u\left(e_u, t_u^i, r_u; y\right); \forall y > y_i \\ v_i^s\left(e_s, t_s^i, y\right) &> v_i^u\left(e_u, t_u^i, r_u; y\right); \forall y < y_i,\ i = 1, 2. \end{aligned} \tag{31}$$

A sufficient but not necessary condition for these single-crossing conditions to hold is thus $t_s^i \geq t_u^i$. If instead $t_s^i < t_u^i$, they require the income elasticity of the demand for school quality to be strong enough relative to the tax rate differential.

The cut-off bid-rent functions (denoted ρ_i^I) are now deduced from equations (5) and (7), yielding:

$$\rho_1^I(y_2) = \Delta_1^h(y_2) + z(y_2)\left[t_s^1 - t_u^1\right] \tag{32}$$

$$\rho_2^I(y_2) = \Delta_2^h(y_2) + y_2\left[t_s^2 - t_u^2\right] \tag{33}$$

If an interior income-tax market equilibrium exists at y_2, then the equilibrum rent premium necessarily satisfies $r^I = \rho_1^I(y_2) = \rho_2^I(y_2)$, which in turn requires:

$$\Delta_1^h(y_2) - \Delta_2^h(y_2) = y_2\left[t_s^2 - t_u^2\right] - z(y_2)\left[t_s^1 - t_u^1\right]. \tag{34}$$

Hence, the *location-incentives constraint* in this case is:

$$z(y_2)\left[t_u^1 - t_s^1\right] - y_2\left[t_u^2 - t_s^2\right] = \Delta_c\left[e_s^* - e_u^*\right]. \tag{35}$$

Again, if the latter condition holds, the optimal cut-off income households are indifferent between the two districts.

Definition 5 *Let $\Omega^I(y_2^*)$ be the set of feasible income tax rate combinations $t_j^i \in [0, 1]$ satisfying the location-incentives constraint (35) and the local budget constraints (27) at y_2^*.*

Lemma 4 *$\Omega^I(y_2^*)$ is non-empty if and only if:*

$$\Delta_c\left[e_u^* - e_s^*\right] \leq y_2^*\bar{t}_u^2 + y_1\bar{t}_s^1 \tag{36}$$

Proof. It is straightforward to check that (35) and the two local budget constraints (27) conform a system of three linearly independent equations in four unknowns (t_j^i) which, therefore, has infinitely many solutions. That set must be restricted by eliminating the combinations of taxes that are not feasible. Let $B_j^i = n_j^i Y_j^i$ denote group i's tax base in district j. If $B_j^i \geq E_j$, $i = 1, 2$, feasibility requires: $t_j^i \in \left[0, \bar{t}_j^i\right]$, with $\bar{t}_j^i = \frac{E_j}{B_j^i} \leq 1$; if $B_j^i \leq E_j$ then $t_j^i \in [0, 1]$ and $t_j^{-i} \in \left[\underline{t}_j^i, \bar{t}_j^i\right]$ where $\underline{t}_j^i = \frac{B_j^{-i}}{B_j^i}\left(\frac{E_i}{B_j^{-i}} - 1\right)$. Next, let $t_u^2 = f\left(t_u^1, y_2^*\right)$ and $t_s^2 = g\left(t_s^1, y_2^*\right)$ denote the local government budget constraints and substitute them into the location-incentives constraint:

$$y_1 t_u^1 - y_2 f\left(t_u^1, y_2^*\right) = \Delta_c\left[e_s^* - e_u^*\right] + y_1 t_s^1 - y_2 g\left(t_s^1, y_2^*\right). \tag{37}$$

Express each side of the equation as a function of t_j^1, $\psi_j\left(t_j^1\right)$, $j = u, s$. These functions are both strictly increasing in t_u^1 and t_s^1, respectively, as $df/dt_u^1 = -B_u^1/B_u^2$ and $dg/dt_s^1 = -B_s^1/B_s^2$. They reach a minimum at $t_j^1 = 0$ where $\psi_u(0) = -y_2^* \bar{t}_u^2 < 0$ and $\psi_s(0) = \Delta_c\left[e_s^* - e_u^*\right] - y_2 \bar{t}_s^2$, and a maximum at $t_j^1 = \bar{t}_j^1$, where $\psi_u\left(\bar{t}_u^1\right) = y_1^* \bar{t}_u^1$ and $\psi_s\left(\bar{t}_s^1\right) = \Delta_c\left[e_s^* - e_u^*\right] + y_1 \bar{t}_s^1$. Let Υ_j be the set of tax rates imposed by district j on group 1 households, t_j^1, such that $\psi_j\left(t_j^1\right) \in \left[\psi_{-j}(0), \psi_{-j}\left(\bar{t}_{-j}^1\right)\right]$. Then, for any element of Υ_u, there exists a unique element in Υ_s such that (37) holds. Finally, note that the sets Υ_j, and thereby also $\Omega^I\left(y_2^*\right)$, are empty sets if the images of ψ_u and ψ_s do not overlap. Otherwise, i.e. if (36) is satisfied, Υ_j and $\Omega^I\left(y_2^*\right)$ are non-empty sets. ■

Condition (36) is mild; it demands the sum of the maximum amount of taxes a type 1 cut-off household would pay in district s (if type 2 residents did not pay any) and a type 2 cut-off household would pay in district u (if type 1 neighbours did not pay any) to be greater than the additional cost of educating a type 2 household (instead of a type 1) in the urban area (rather than in the suburbs).

Proposition 3 *Consider a solution to the Social Planner Problem with optimal cut-off incomes (y_1^*, y_2^*) and satisfying $e_u^* > e_s^*$. For every combination of local income tax rates in $\Omega^I\left(y_2^*\right)$ such that lemma 3 holds, there exists an equilibrium with cut-off incomes $y_1^I = y_1^*$, $y_2^I = y_2^*$ and rent premium $r^I = \rho_1^I\left(y_2^*\right) = \rho_2^I\left(y_2^*\right)$.*

Proof. The proof is analogous to the one of proposition 2 and is omitted for the sake of brevity. ■

The result is not general because $\Omega^I(y_2^*)$ may be an empty set and, even if it is not, there may not be an element satisfying the relevant single-crossing conditions. Hence, the optimal allocation may not be sustainable as a market equilibrium with personalised income taxes. The remaining of this section shows that it is however possible to combine other fiscal tools with income taxes to correct for their distortionary location effects. The next proposition assumes that local governments use anonymous income taxation to fund education. If the single-crossing conditions are satisfied, then the location externalities can be internalised with a self-funded lump-sum transfer scheme among the residents of the urban district.[38] If not, it must be the case that $t_u\left(y_2^I\right) > t_s\left(y_2^I\right)$. Then, the proposal involves applying the suburbs' tax-rate to the urban area and imposing personalised head taxes to correct for the location externalities and to fund the resulting deficit, D_u.

Proposition 4 *Consider a solution to the Social Planner Problem with optimal cut-off incomes (y_1^*, y_2^*) and satisfying $e_u^* > e_s^*$. Suppose that local governments use anonymous income taxes to fund education.*
1) If the vector of local policies $\Pi^I(y_2^*)$ and household preferences satisfy lemma 3, then, the unique self-funded lump-sum transfers scheme from type 2 to type 1 urban residents (L_1, L_2) satisfying

$$L_1 n_u^1 - L_2 n_u^2 = 0 \tag{38}$$

$$-L_1 - L_2 = \Delta_c\left[e_s^* - e_u^*\right] - (y_1^* - y_2^*)(t_u - t_s) \tag{39}$$

sustains the optimal allocation as an income tax equilibrium.
2) In other case, setting the urban area tax rate at $t_u = t_s(y_2^*)$, the unique personalised head-tax scheme imposed on urban residents $\left(\widehat{T}_1, \widehat{T}_2\right)$ satisfying

$$\widehat{T}_1 n_u^1 + \widehat{T}_2 n_u^2 = D_u \tag{40}$$

$$\widehat{T}_1 - \widehat{T}_2 = \Delta_c\left[e_s^* - e_u^*\right] \tag{41}$$

sustains the optimal allocation as an income tax equilibrium.

[38] Correcting for location externalities could be achieved in other ways as well. For example, with a transfer between households of a given type living in different districts.

Proof. Clearly, both systems of equations have a unique solution. Equations (38) and (40) guarantee, in each case, that the scheme is either self-funded or covers the budget deficit arising in the urban district D_u. Hence, both proposals ensure that the two local budget constraints are satisfied. Under anonymous income taxation, the location incentives constraint (35) reduces to

$$\Delta_c [e_s^* - e_u^*] = (y_1^* - y_2^*) (t_u (y_2^*) - t_s (y_2^*)) \tag{42}$$

Without the proposed schemes, the optimal allocation cannot be sustained as an equilibrium. 1) In this case, the proposed scheme reduces the willingness to pay for a house in the urban area of every type 2 household by an amount equal to L_2 and increases that of type 1 households by L_1. Equation (39) requires the sum of both to cover the difference between cut-off households' "optimal" *relative* willingness to pay $\Delta_c [e_s^* - e_u^*]$ and the one induced by income taxes $(y_1^* - y_2^*) (t_u (y_2^*) - t_s (y_2^*))$. Because lemma 3 holds by assumption, the scheme is thus able to sustain the optimal allocation as an equilibrium. 2) Here, $t_u = t_s (y_2^*)$, so that lemma 3 is satisfied. Moreover, income taxes do not affect the location incentives of households: $(y_1^* - y_2^*) (t_u - t_s (y_2^*)) = 0$. Therefore, the proposed scheme needs to increase type 1 households' willingness to pay for living in the urban area with respect to that of type 2 ones by $\Delta_c [e_s^* - e_u^*]$, which is precisely what equation (41) imposes. ■

The lump-sum transfers implied by the first proposal could be from type 2 to type 1 households or viceversa. The reason is that the latter have lower income so that the tax-price of entry to the urban area induced by anonymous income taxes may be too low for them relative to that required from type 2 cut-off ones. On the contrary, the personalised head taxes that complement the anonymous and uniform income tax scheme in the second proposal need to be greater for type 2 households. This result demonstrates that, while single-crossing conditions mark the limits to the compatibility between optimality and redistribution, these two sometimes conflicting normative objectives are not incompatible with each other in the current setting. Actually, the second proposal sets a lower bound for the amount of tax redistribution effected in the rich district, as the single-crossing conditions will be satisfied for some $t_u > t_s (y_2^*)$.

6 The ambiguous comparison between anonymous head and income taxes

Suppose now that local governments observe the marginal cost of providing an additional unit of school quality to the district $(n_1c_1 + n_2c_2)$ but cannot identify individual marginal congestion costs c_i or cannot use that information to tax-discriminate across household types. This section proves that anonymous head and income taxes cannot be unambiguously ranked according to the distortions they generate.

6.1 Head taxes

In this case, the budget constraints, indirect utility, bid-rent and cut-off bid-rent functions are obtained by setting $T_j^1 = T_j^2$ in (19), (20), (21), (23) and (24), respectively. The next proposition reveals that, while corner head-tax equilibria are optimal, interior ones are suboptimal.

Proposition 5 *1) A head-tax equilibrium exists.*
2) If the SPP has an interior solution, then there exists an interior head-tax equilibrium. Every interior head-tax equilibrium is suboptimal.
3) If the SPP has a corner solution and the sign of $\Delta_1^h(y_2^*) - \Delta_2^h(y_2^*)$ is positive (negative), then there exists a corner (interior) head-tax equilibrium. Every corner head-tax equilibrium is optimal.

Proof. 1) If $e_u(y) > e_s(y)\ \forall y \in [\widetilde{y}, \overline{y}_2]$, then, by lemma 2, the single-crossing conditions ensuring WTS hold in that interval. Assumption 2 implies that $\rho_1^H(\widetilde{y}) - \rho_2^H(\widetilde{y}) > 0$. Then, if $\rho_1^H(\overline{y}_2) - \rho_2^H(\overline{y}_2) < 0$, the intermediate value theorem and the continuity of the cut-off bid-rent functions ensure that there is a level of income $y_2^H \in [\widetilde{y}, \overline{y}_2]$ such that $\rho_1^H\left(y_2^H\right) - \rho_2^H\left(y_2^H\right) = 0$. Hence, such allocation satisfies the rationality condition of equilibrium E1 for $r_H = \rho_1^H\left(y_2^H\right) = \rho_2^H\left(y_2^H\right)$. Because cut-off bid-rent functions embed the equilibrium conditions E2 to E4, y_2^H and the associated vector of local policies $\Pi_H\left(y_2^H\right)$ constitute an interior head-tax equilibrium with rent premium r_H. If instead $\rho_1^H(\overline{y}_2) - \rho_2^H(\overline{y}_2) \geq 0$, the corner allocation of households to districts $\overline{y}_2$ also satisfies E1. In the case where $\gamma > 0.5$, cut-off incomes satisfy $\overline{y}_2 = \overline{y}$ and $z(\overline{y}_2) > \underline{y}$ and, for the same reasons aforementioned, for every $r_H \in \left[\rho_2^H(\overline{y}_2), \rho_1^H(\overline{y}_2)\right]$ there exists a corner equilibrium with all type 2 households living in the suburbs. If $\gamma \leq 0.5$, then $\overline{y}_2 \leq \overline{y}$, $z(\overline{y}_2) \geq \underline{y}$ and there is a corner equilibrium with every type 1 household living in the urban area with rent premium $r_H = \rho_2^H(\overline{y}_2)$. Finally, note that if there exists $y \in [\widetilde{y}, \overline{y}_2]$ such that $e_u(y) = e_s(y)$, it can be readily checked that such allocation y, $\Pi_H(y)$ and the rent premium $r_H = T_s(y) - T_u(y)$ leave every household indifferent between the two districts so that they constitute an equilibrium.

2) By proposition 1, in an interior solution to the SPP:

$$\Delta_1^h(y_2^*) - \Delta_2^h(y_2^*) = -\Delta_c(e_u^* - e_s^*). \tag{43}$$

In an interior head-tax equilibrium, in turn, $\rho_1^H(y_2^H) - \rho_2^H(y_2^H) = 0$ implies:

$$\Delta_1^h(y_2^H) - \Delta_2^h(y_2^H) = 0. \tag{44}$$

Because the RHS of (43) is negative: $\rho_1^H(y_1^*) < \rho_2^H(y_2^*)$ and $MSV_1(y_2^H) > MSV_2(y_2^H)$. The latter implies that any interior head-tax equilibrium is suboptimal. The former, along with the fact that $\rho_1^H(\widetilde{y}) > \rho_2^H(\widetilde{y})$, the continuity of the cut-off bid rent functions and the intermediate value theorem, entail the existence of a level of income $y_2^H \in (\widetilde{y}, y_2^*)$ such that $\rho_1^H(y_2^H) = \rho_2^H(y_2^H)$ for which an interior head-tax equilibrium exists.

3) In a corner solution to the SPP:

$$\Delta_1^h(\overline{y}_2) - \Delta_2^h(\overline{y}_2) \geq -\Delta_c(e_u^* - e_s^*). \tag{45}$$

Because the RHS of (45) is negative, its LHS may be negative or positive. In the former case, $\rho_1^H(\overline{y}_2) - \rho_2^H(\overline{y}_2) < 0$ and, as proved above, there is a level of income $y_2^H \in (\widetilde{y}, y_2^*)$ such that $\rho_1^H(y_2^H) = \rho_2^H(y_2^H)$ for which an interior head-tax equilibrium exists. In the latter case, note that

$$\rho_1^H(\overline{y}_2) - \rho_2^H(\overline{y}_2) > 0 \Leftrightarrow \Delta_1^h(\overline{y}_2) - \Delta_2^h(\overline{y}_2) > 0$$

so that existence of a corner head-tax equilibrium implies that (45) holds and so the existence of a corner solution to the SPP, and viceversa. ■

Interior equilibria emerging with anonymous local head taxes generate a suboptimal distribution of households across districts. Residential choices generate negative externalities because the homogenous tax-bill levied on every resident of the good area does not cover the marginal costs of admitting households into the district, being too low for high-cost households and too high for low-cost ones. Hence, too many high-cost households live in the good (urban) school district in equilibrium.

6.2 Income taxes

This subsection compares the distortions emerging in market equilibrium when local governments use non-differentiated head and income taxes. In the latter case, the local budget constraints and the indirect utility, bid rent and cut-off bid-rent functions are obtained by setting $t_j^1 = t_j^2$ in (27), (28), (29), (32) and (33), respectively.

As in the case with differentiated taxes, the possibility that the single-crossing conditions might not be satisfied implies that existence of an income-tax equilibrium satisfying WTS is not guaranteed. Nevertheless, following a similar argument as in the proof of the previous proposition, it can be shown that, if $e_u(y) > e_s(y)\ \forall y \in [\widetilde{y}, \overline{y}_2]$, then either there is a level of income $y \in (\widetilde{y}, \overline{y}_2)$ for which $\rho_1^I(y) = \rho_2^I(y)$, or $\rho_1^I(\overline{y}_2) \geq \rho_2^I(\overline{y}_2)$. In both cases, equilibrium requirements (E2)-(E4) hold. The implied allocation will be an equilibrium if the single-crossing conditions are also met. If an interior equilibrium exists, its cut-off incomes are derived from the equation of the two types' cut-off bid-rent functions, which yields:

$$\Delta_1^h\left(y_2^I\right) - \Delta_2^h\left(y_2^I\right) = \left(z\left(y_2^I\right) - y_2^I\right)\left(t_u\left(y_2^I\right) - t_s\left(y_2^I\right)\right). \tag{46}$$

The comparison between (46) and the optimality requirement (13) confirms that anonymous income taxes distort the allocation of households across districts. The next result shows that head taxes may induce greater welfare losses than income taxes.[39]

Proposition 6 *Suppose an interior income-tax equilibrium exists with cut-off incomes $y_1^I < y_2^I < y_2^*$ and tax rates $t_u\left(y_2^I\right) > t_s\left(y_2^I\right)$. Then, another head-tax equilibrium inducing larger locational distortions exists.*

Proof. In an interior income-tax equilibrium, cut-off incomes satisfy (46). In turn, in an interior head-tax equilibrium, they fulfill $\rho_1^H\left(y_2^H\right) = \rho_2^H\left(y_2^H\right)$. Because $t_u\left(y_2^I\right) > t_s\left(y_2^I\right)$ by assumption in the case considered, the RHS of (46) is negative, implying $\rho_1^H\left(y_2^I\right) < \rho_2^H\left(y_2^I\right)$. Using again the fact that

[39] It can also be shown that, if t_u is set equal to $t_s(y)$ and the resulting urban budget deficit is financed through a uniform head-tax levied on urban residents, income taxes do as well as head taxes. The reason is that in that case:

$$\rho_1^I(y) - \rho_2^I(y) = \rho_1^H(y) - \rho_2^H(y) = \Delta_1^h(y) - \Delta_2^h(y)$$

$\rho_1^H(\widetilde{y}) > \rho_2^H(\widetilde{y})$, continuity of the cut-off bid rent functions and the intermediate value theorem imply the existence of a level of income $y_2^H \in \left(\widetilde{y}, y_2^I\right)$ such that $\rho_1^H(y_2^H) = \rho_2^H(y_2^H)$ and for which a head-tax equilibrium exists. Therefore, $y_2^* > y_2^I > y_2^H$, $y_1^* < y_2^I < y_1^H$. ■

Results in this section clash with the view of local head taxes as efficiency-enhancing benefit taxes. When governments cannot observe the cost-parameters or use that information to tax-discriminate across households of different types, head taxation may be more distortionary than an ability-to-pay tax such as the proportional income tax. The reason is that, if an equilibrium has higher tax rates in the urban area, $t_u > t_s$, the cut-off households of the high-cost type face a greater tax-price of entry into the urban area than the lower income cut-off households of the low-cost type. Thereby, the negative cost-externalities they impose on the rest are (partially) internalised.

7 Concluding remarks

The analysis in this paper offers new insights on the relative normative merits of local head and local income taxes. The main message is that, in the presence of non-anonymous crowding, head taxes are not necessarily superior to income taxes. In cases where local governments can observe the crowding costs of different types and tax-discriminate across them, both head and income taxes are able to implement the optimal outcome. Because what matters for the outcome of the location game is the *relative* willingness to pay for entering the good district, "optimal" personalised head taxes need not cover marginal coongestion costs and allow for redistribution across crowding types. In turn, income taxes pemit redistribution of income and, possibly, across types too. Because optimality requires segregation in the utilitarian normative framework considered, the necessary single-cossing conditions for a segregated equilibrium to exist mark the limits of the compatibility between redistribution and location optimality but do not rule it out. When differentiated taxes are not available, the two tax systems lead to suboptimal equilibria and they cannot be unambiguougsly ranked according to the location distortions they generate.

It is important to check the robustness of these conclusions to alternative specifications of the technology and the housing markets. That is to say, to relax the assumptions of linear crowding costs and of inelastic housing supplies. On the other hand, the results were obtained in a quasi-linear utility framework, where the social planner has no taste for redistribution, eliminating an efficiency advantage of income taxes over head taxes.

Two important questions for further research emerge. The first one concerns the comparison of income taxes to property taxes with and without zoning regulations, extending Calabrese et al. (2007) to include differential crowding effects and income taxes. The second concerns the relative performance of alternative tax systems when local tax and spending policies are selected through an electoral process. It is worth mentioning to conclude that these results suggest as well that, in the presence of asymmetric and unobservable crowding costs, the public sector could use exams and condition access to schools on the results to derive welfare gains.

References

[1] Brueckner, J. K., Lee, K., 1989. Club theory with a peer group effect. Regional Science and Urban Economics, 19: 399-420.

[2] Bénabou, R., 1996. Equity and efficiency in human capital investment: the local connection. Review of Economic Studies 63: 237–264.

[3] Bewley, T., 1981. A critique of Tiebout's theory of local public expenditures. Econometrica, 49: 713-740.

[4] Calabrese, S., Epple D. and Romano, R. E., 2007. On The Political Economy of Zoning. Journal of Public Economics, 91: 25-49.

[5] Calabrese, S., Epple D. and Romano, R. E., 2010. Inefficiencies from Metropolitan Political and Fiscal Decentralization: Failures of Tiebout Competition. Working Paper.

[6] Conley, J. and Wooders, M., 1998. Equivalence of the core and competitive equilibrium in a Tiebout economy with cowding types. Journal of Urban Economics, 41: 421-440.

[7] de Bartolome, C. A. M., 1990. Equilibrium and Inefficiency in a Community Model with Peer Group Effects. Journal of Political Economy, 98(1): 110-133.

[8] de Bartolome, C. A. M. and Ross, S. L., 2004. Who's in charge of the central city? The conflict between efficiency and equity in the design of a metropolitan area. Journal of Urban Economics, 56(3): 458-483.

[9] Epple, D, Filimon, R. and Romer, T., 1984. Equilibrium among local jurisdictions: Toward an integrated treatment of voting and residential choice. Journal of Public Economics 24, 281-308.

[10] Epple, D, Filimon, R. and Romer, T., 1993. Existence of Voting and Housing Equilibrium in a System of Communities with Property Taxes", Regional Science and Urban Economics 23: 585-610.

[11] Epple, D. and Romano, R., 2003. Neighbourhood schools, choice and the distribution of educational benefits. In: Hoxby, C.M. (Ed.), The Economics of School Choice, National Bureau of Economic Research, University of Chicago Press, pp. 227-286.

[12] Epple, D. and Platt, G. J., 1998. Equilibrium and local redistribution in an urban economy when households differ in both preferences and income. Journal of Urban Economics 43, 23– 51.

[13] Goodspeed, T. J., 1989. A reexamination of the use of ability to pay taxes by local governments. Journal of Public Economics, 38: 319-342.

[14] Goodspeed, T. J., 1995. Local income taxation: an externality, pigouvian solution and public policies. Regional Science and Urban Economics, 25: 279-296.

[15] Hamilton, Bruce W., 1975. Zoning and Property Taxation in a System of Local Governments, Urban Studies, 12: 205-211.

[16] Kessler, A. and Lülfessmann, C., 2005. Tiebout and redistribution in a model of residential and political choice. Journal of Public Economics, 89: 501-528.

[17] Nechyba, 1999. School-finance induced migration patterns: The case of private school vouchers. Journal of Public Economic Theory, 1: 5-50.

[18] Rose-Ackerman, S., 1979. Market models of local governments: exit, voting and the land market. Journal of Urban Economics, 6: 319-337.

[19] Ross, S. and Yinger, J., 1999. Sorting and voting: A rreview of the literature on urban public finance. In Ceshire and Mills (eds.): Handbook of Regional and Urban Economics, vol. 3., chapter 47. North-Holland, Amsterdam.

[20] Schmidheiny, K., 2006a. Income segregation and local progressive taxation: Empirical evidence from Switzerland. Journal of Public Economics, 90: 429-458.

[21] Schmidheiny, K., 2006b. Income segregation from local income taxation when households differ in both preferences and incomes. Regional Science and Urban Economics, 36: 270-299.

[22] Scotchmer, S., 2002. Local public goods and clubs. In Auerbach and Feldstein (eds): Handbook of Public Economics. vol. 4, chapter 29, 1997-2042. North-Holland, Amsterdam.

[23] Schwabb, R. M. and Oates, W. E., 1991. Community conmposition and the provision of local public goods: A nomative analysis. Journal of Public Economics, 44: 217-237.

[24] Tiebout, C. M., 1956. A pure theory of local expenditures. Journal of Political Economy, 64(5): 416-424.

[25] Wildasin, D., 1986. Urban public finance. Hardwood Academic Publishers, New York.

[26] Wooders, M., 1978. Equilibria, the core, and jurisdiction structures in economies with a local public good. Journal of Economic Theory, 18: 328-348.

[27] Wooders, M., 1999. Multijurisdictional economies, the Tiebout hypothesis and sorting. Proceedings of the National Academy of Science, 96: 10585-10587.