



**University of
Leicester**

DEPARTMENT OF ECONOMICS

The peer group effect and the optimality properties of head and income taxes

Francisco Martínez-Mora, University of Leicester, UK

**Working Paper No. 11/26
April 2011**

The peer group effect and the optimality properties of head and income taxes

Francisco Martínez-Mora

University of Leicester and FEDEA

Abstract

This paper studies a Tiebout model with two school districts, housing markets and peer effects to re-evaluate the optimality properties of the allocation of households to districts induced by head and income taxes. The main novel results reveal that head taxes are not superior to income taxes and that the indirect redistribution implied by income taxation is not necessarily at odds with location optimality or associated to welfare losses. Many combinations of head taxes differentiated by household type can sustain the optimal outcome as an equilibrium. While this may not be possible using differentiated income taxes, a combination of non-differentiated ones and differentiated head taxes levied on the residents of the rich district can lead to the optimal outcome and effect significant local redistribution. In turn, non-differentiated head taxes are suboptimal (unless optimality requires one of the districts to be type-homogeneous) and a combination of uniform income taxes and head taxes levied on the rich district's population can do as well as them. Moreover, non-differentiated income taxes may generate smaller welfare losses than their lump-sum counterpart, a result which clashes with the benefit view of head taxes.

Key words: Tiebout, peer effects, head tax, income tax, optimality.

1 Introduction

In his seminal contribution, Tiebout (1956) suggested that head taxes would lead to an optimal sorting of households into homogenous jurisdictions, which would therefore select the optimal level of local public good provision 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 jurisdictions as household types, the use of local head taxes and a technology of production of the local public good exhibiting constant and identical marginal costs with respect to the population. In

a recent contribution, Calabrese et al. (2010) extended the result to a less stylised framework with housing markets and a smaller number of exogenous jurisdictions than of household types.¹

Income taxes, in turn, generate welfare losses (Wildasin, 1986; Goodspeed, 1989, 1995). Since households with below average income face lower tax prices (that is, they receive greater amounts of spending per unit of tax paid), they do not internalise the true costs they impose on local finances, and end up living in richer districts than they would otherwise do and benefitting from the implied redistribution. Moreover, to the extent that income taxes distort tax prices they also distort the political process.

The assumption that the costs of provision of the local public good are independent of the characteristics of the population is crucial to attain these conclusions.² Yet, such assumption is difficult to justify, especially if the local public good considered is schooling. In that context, as parents, teachers, policy-makers and academics know well, the diversity of the student population and the importance of peer influences make the peer group effect an essential element of the problem.³

The objective of this paper is to re-evaluate the optimality properties of the equilibrium district compositions induced by head and income taxes in the presence of peer effects (or, more generally, of non-anonymous crowding). The model represents a city divided into two jurisdictions (school districts) with exogenously given boundaries and housing supplies. Each district provides tax-funded public education to its residents and households choose where to live, a decision that subsumes the choice of school.⁴ Households differ continuously in income, and are of one of two types that depend on the child's ability to benefit

¹ The Tiebout hypothesis has also been formalised in the slightly different framework of club theory. Club models do not incorporate an explicit representation of land markets and allow clubs to form endogenously. A joint review of the two literatures can be found in Scotchmer (2002). A succinct review of the club literature on the Tiebout model is in Wooders (1999).

² For example, de Bartolome (1990) revealed that the peer group effect could generate inefficiencies even if local governments relied on head taxes to fund schools. In similar contexts, other authors (including Schwabb and Oates, 1991, Brueckner and Lee, 1989, and Conley and Wooders, 1998) have shown that differentiated taxes are required to achieve optimality.

³ Recent evidence documenting the impact of peer effects on student outcomes includes the work on Kenyan primary schools of Duflo et al. (forthcoming) and that of Lavy et al. (2009) on English secondary schools.

⁴ As the most prominent example of a local public good, the case of local schooling motivates the analysis. Results apply more generally to any local public good whose technology of production depends on the characteristics of the population that receives it.

from education and on the cost of providing her with a unit of school quality. Peer effects are simply modeled as the outcome of these differential crowding costs: a school with a greater proportion of low cost pupils can provide the same level of school quality using less resources or provide greater quality with equal spending per pupil. In order to focus the analysis on locational optimality, the model assumes that local governments provide the (locally) optimal level of school quality.⁵

The paper adopts a utilitarian normative framework, with the Social Welfare Function (SWF) being the unweighted sum of utility in the economy. In that setting, optimality requires households of the same type to be perfectly segregated by income across districts, with higher income ones living in the better school district. That requirement generates two necessary single-crossing conditions, one per household type. Optimality also demands some income mixing: relatively low income households of the low cost and high benefit type should reside in the good school district instead of higher income ones of the other type.

The results of the analysis critically depend on whether household cost-types are observable and taxes can be differentiated across them or not. In case they are, on the one hand, multiple head taxes that differentiate across household types can achieve optimality.⁶ In a model with housing markets, therefore, these taxes need not cover the marginal cost of admitting a household into a district, as in a model without them (e.g. Schwabb and Oates, 1991). Instead, the extra taxes households of different types pay in the rich district with respect to the poor one must provide them with the optimal *relative location incentives* (that is, they must generate differences in the cost of entry each type of household faces to live in the rich district that internalise the externalities emerging from the peer effect). On the other hand, although it is often possible to find combinations of differentiated income tax rates simultaneously satisfying the public sector budget constraints and providing cut-off households with the optimal location incentives, it may well be the case that none of them fulfills the two necessary single-crossing conditions optimality requires. Notwithstanding, it is always possible to reach optimality by combining anonymous income taxes with differentiated head taxes levied on the

⁵ Most of the literature focuses on one of two efficiency questions: whether the emerging allocation of households to districts is efficient and whether the political process selects the efficient level of provision and taxation. 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.

⁶ Notice that these differentiated taxes depend on crowding types (which may be publicly observable) and not on tastes (which are not). See Conley and Wooders (1998).

residents of the rich district. At most, the necessary single-crossing conditions will require the income tax rates to be equated across districts. Therefore, whilst the single-crossing conditions mark the limits to the compatibility between income taxes and optimality, they do not rule it out.

Additional counter-intuitive results arise when local governments cannot tailor taxes to household types. In that case, it is not possible to unambiguously rank head and income taxes. Anonymous head taxes not only induce a suboptimal distribution of households across districts (unless optimality requires one of the districts to be type-homogeneous) but may also be more distortionary than an ability-to-pay tax. The intuition is the following: whereas anonymous head taxes do not affect households relative location incentives, income taxes generate differences in entry prices between the marginal (or cut-off) household of each type living in the rich district (i.e. the lowest income residents of their type in that district). The reason is that, in general, these households have different incomes in equilibrium. Under some circumstances, such differences will imperfectly internalise the location externalities emerging from the peer effect.⁷ Furthermore, non-differentiated income taxes, combined with head taxes that are levied on (or with lump-sum transfers to) the rich district's population, can match the outcome attained by head taxes. These results clash with the benefit view of local head taxes and imply that income taxes will not necessarily generate welfare losses with respect to them.

The rest of the paper is organised as follows. The next section presents the model. Section 3 derives the housing markets equilibrium condition when households of the same type are income-segregated across districts. The following section obtains the optimal allocation of households to districts. The analysis then turns to the comparison between head and income taxes. Section 5 studies the case where governments can use differentiated 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 communities (school districts) with fixed boundaries, labeled as the urban area and the suburbs, and indexed with $j = u, s$. Districts provide tax-funded, tuition-free public education of homogeneous quality ($e_j > 0$) to all their school-aged residents⁸. With no loss

⁷ This possibility was conjectured by Goodspeed (1989, footnote 12).

⁸ 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

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.⁹

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 the child with a given level of school quality and the benefit she receives from it.¹⁰ To present results in the simplest possible setting, there are two cost-benefit types of households. Household types are indexed with $i = 1, 2$; $\gamma \in (0, 1)$ measures the proportion of type 1 households in the population. The analysis presented below considers cases where crowding and benefit types are correlated as follows:¹¹

Assumption 1 *Type 1 pupils impose smaller congestion costs onto schools and derive greater benefits from school quality than type 2 ones.*

The interpretation is that type 1 households have higher ability offsprings and that the cost difference is the result of differential peer effects operating at the school level. Peer effects are thus modeled *a-la-Lazear* (2001) as the result of differential crowding effects: schools with a larger proportion of low cost pupils are able to provide the same level of school quality using less resources (e.g. with larger classes) or greater quality with equal spending per pupil. The technology of production of school quality is identical across districts and linear. It is described by the cost function:

$$C(n_j^1, n_j^2, e_j) = (n_j^1 c_1 + n_j^2 c_2) e_j; \quad c_1 < c_2$$

where n_j^i stands for the mass of households of type i living in district j , and $c_1 < c_2$. Let Δ_c denote the crowding cost differential $\Delta_c = c_2 - c_1 > 0$.

Household income is denoted with $y \in D \equiv [\underline{y}, \bar{y}]$, 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

otherwise.

⁹ For simplicity sake, private schools are excluded from the analysis.

¹⁰ 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 1999). There are however exceptions in which households differ, additionally, along a taste parameter (Epple and Platt, 1998; Kessler and Lülffessmann, 2005; Schmidheiny, 2006a, 2006b), a productivity one such as the offspring ability (Epple and Romano, 2003), or both (Wildasin, 1986).

¹¹ It is important to note that the results on the comparison between head and income taxes do not depend on the direction of correlation between cost and benefit types assumed. Only the circumstances under which income taxes may generate smaller welfare losses than head taxes change.

$\phi_i(y) = \Phi'_i(y)$; $i = 1, 2$. The total (and average) income is:

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

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

The model of the housing market avoids well-known existence problems illustrated, for example, by Rose-Ackerman (1979) or Epple et al. (1984): districts have a fixed (and identical) supply of homogeneous houses, denoted H_j , and the total supply is equal to the mass of households that resides in the city: $H_u = H_s = 1/2$.¹² 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.¹³ 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 .¹⁴

Preferences are defined over a private composite good (the numeraire), x , and the offspring's future human capital (or income), h .¹⁵ 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 preferences over pairs (x, e) . Following de Bartolome and Ross (2004), I adopt a quasi-linear specification of utility:

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

where h is monotonically increasing in both arguments and strictly concave in e . Assumption 1 implies:

$$\partial h_1(e, y) / \partial e > \partial h_2(e, y) / \partial e, \forall (e, y).$$

The choice of a quasi-linear utility function not only simplifies the analysis but also completely separates efficiency and equity considerations: because the marginal utility of private consumption is constant and equal to one (i.e.

¹² The assumption of equal district sizes 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).

¹³ This implies that the opportunity cost of the residential use of land –the value of industrial or agricultural alternative uses– is normalised to zero.

¹⁴ The rent premium may be negative in equilibrium which would indicate the capitalisation of higher taxes in the urban area.

¹⁵ Because houses are homogeneous they are excluded from the preference relation.

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 satisfy:

Assumption 2 *Education is a normal good.*

Assumption 2 implies a positive income elasticity of demand for school quality, which agrees with the empirical evidence (Ross and Yinger, 1999). This assumption restricts quasi-linear preferences, requiring home and school inputs to be complements in the human capital production function.¹⁶ The next assumption establishes the behaviour of local governments, focusing the analysis on locational optimality.

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

Local governments fund their spending with head or proportional income taxes (indexed with $k = H, I$) that may be differentiated across types or not. In the latter case, which I refer to as anonymous taxation, a tax-bill function, denoted $\tau_j^k(e_j, n_j^1, n_j^2, y)$ generically, derives from the corresponding local budget constraint. The tax-bill function must also meet the feasibility constraint that household tax payments must be smaller than household income. In the former case, the differentiated head tax bills or income tax rates will be required to fulfill these two constraints too. The indirect utility function of a household of type i that has income y and lives in district j is:

$$v_j^i(e_j, \tau_j^{ik}(y), r_j; y) = y - \tau_j^{ik}(y) - r_j + h_i(e, y)$$

where $\tau_j^{ki}(y)$ represents their tax bill under the tax system k . A useful tool in the analysis that follows are the so-called bid-rent functions, which I define next¹⁷:

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

$$v_u^i(e_u, \tau_u^{ik}(y), r_i^k(y), y) = v_s^i(e_s, \tau_s^{ik}(y), r_s, y). \quad (3)$$

The model is static. An equilibrium is an allocation of households to districts,

¹⁶ This could be due, for example, to better labour market networking of better-off parents.

¹⁷ For a review of the literature on urban public finance that extensively uses bid-rent functions see Ross and Yinger (1999).

a vector of local head tax bills or income 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 clear: $H_j = n_j^1 + n_j^2$; $j = u, s$.*

E3 *Local governments balance their budget.*

E4 *School qualities are optimal given the local population of households.*

3 The housing market constraint

In the utilitarian normative framework considered (explained in detail 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 type but of different income who reside in different districts, the higher income one lives in the urban area.*

In an allocation satisfying the WTS property, households of the same type living in the same district belong to a single income interval and the intervals corresponding to each district do not overlap. The monotonicity of the income distribution functions then implies that, when households of a certain type i are present in the two districts, a unique cut-off income, denoted y_i , exists such that households of that type with income $y > (<)y_i$ reside in the urban (suburban) district. A solution to the Social Planner Problem or an equilibrium may however have all households of a given type concentrated in one district. In particular, type 1 households may all concentrate in the urban district, where the good school is located, or type 2 ones may all concentrate in the suburbs. Clearly, the former case may only happen when $\gamma \leq H_u$, that is, if the proportion of type 1 households in the population is no larger than the size of the urban district. In that case, cut-off incomes are $y_1 = \underline{y}$ and $y_2 = \bar{y}_2 \leq \bar{y}$, where \bar{y}_2 is defined by:

$$1/2 - \gamma = (1 - \gamma) [1 - \Phi_2(\bar{y}_2)].$$

The latter case may arise if $\gamma \geq H_u$ and will have cut-off incomes $y_1 \geq \underline{y}$ and $y_2 = \bar{y}$. Figure 1 represents examples of these three possibilities: panel (a) depicts an allocation with mixed districts, while panels (b) and (c) represent cases where the urban or the suburban district is type-homogenous. It is thus

possible to express the mass of households of each type living in each district as a function of the corresponding cut-off income: $n_u^i(y_i) = \gamma_i [1 - \Phi_i(y_i)]$ and $n_s^i(y_i) = \gamma_i \Phi_i(y_i)$ (where $\gamma_1 = \gamma$ and $\gamma_2 = 1 - \gamma$). Moreover, the urban district housing market constraint (which implies the suburbs housing market constraint) can also be expressed in terms of the cut-off incomes:

$$H_u = 1/2 = \gamma [1 - \Phi_1(y_1)] + (1 - \gamma) [1 - \Phi_2(y_2)] \quad (4)$$

Lemma 1 *Consider allocations satisfying WTS. Cut-off incomes are linked through a continuously decreasing function z defined on a compact set $[\underline{y}^2, \bar{y}^2]$, where $\underline{y}^2 = \underline{y}$ and $\bar{y}^2 = \bar{y}$ if $\gamma \geq 1/2$, and $\underline{y}^2 > \underline{y}$ and $\bar{y}^2 < \bar{y}$ if $\gamma < 1/2$; z is implicitly defined by:*¹⁸

$$1/2 - (1 - \gamma) [1 - \Phi_2(y_2)] = \gamma [1 - \Phi_1(z(y_2))] \quad (5)$$

Proof. The proof needs to establish that, for any value of y_2 , there exists a unique value of y_1 satisfying (4). Suppose $\gamma > 1/2$; then,

$$0 \leq (1 - \gamma) (1 - \Phi_2(y)) \leq 1/2; \quad \forall y \in [\underline{y}, \bar{y}]$$

and the LHS of (5) belongs to the interval $(0, 1/2]$. Given that $\gamma [1 - \Phi_1(\bar{y})] = 0$ and $1/2 \leq \gamma [1 - \Phi_1(\underline{y})] < 1$, continuity and strict monotonicity of the income distribution functions and the intermediate value theorem imply that, for any $y_2 \in [\underline{y}, \bar{y}]$, there exists a unique $y_1 \in [\underline{y}, \bar{y}]$ such that the housing market constraint (4) holds. Suppose instead that $\gamma \leq 1/2$; then

$$0 \leq (1 - \gamma) [1 - \Phi_2(y)] \leq (1 - \gamma); \quad \forall y \in [\underline{y}, \bar{y}].$$

Hence, by continuity and strict monotonicity of Φ_2 there exists a unique income $\underline{y}^2 = y$ such that $(1 - \gamma) [1 - \Phi_2(y)] = 1/2$, and another $\bar{y}^2 = y$ such that $(1 - \gamma) [1 - \Phi_2(y)] = 1/2 - \gamma$. Moreover, given that $0 \leq \gamma [1 - \Phi_1(y)] \leq \gamma \forall y \in [\underline{y}, \bar{y}]$, again the continuity and strict monotonicity of Φ_1 and Φ_2 and the intermediate value theorem imply that, for any $y_2 \in [\underline{y}^2, \bar{y}^2]$, there exists a unique $y_1 \in [\underline{y}, \bar{y}]$ such that the housing market constraint (4) holds. Furthermore, these two properties guarantee that z is continuous and decreasing. Finally, apply the implicit function theorem to (4) in order to find the derivative:

$$\frac{dy_1}{dy_2} = -\frac{(1 - \gamma) \phi_2(y_2)}{\gamma \phi_1(z(y_2))} < 0. \quad (6)$$

■

The analysis focuses on allocations satisfying WTS. These will be characterised simply with the type 2 cut-off income y_2 . Since the human capital production

¹⁸ For uneven district sizes, the domain of z depends on the size of the urban district relative to the proportion of each type of households in the population.

function displays decreasing marginal returns to school quality, Assumption 3 implies that every allocation of households to districts is linked to a *unique* pair of school qualities. For allocations satisfying WTS, then, optimal school qualities can be written as a function of y_2 : $e_j(y_2)$. In the case of anonymous taxation, that is also true of the tax variables: given school qualities and the distribution of households across districts, and with some abuse of notation, it is possible to write:

$$\tau_j^k(y_2) = \tau_j^k(e_j(y_2), n_j^1(y_2), n_j^2(y_2), y_2).$$

Therefore, under anonymous taxation, to every allocation satisfying WTS, y_2 , corresponds a vector of local policy variables:

$$\mathbf{\Pi}^k(y_2) = [e_u(y_2) \ e_s(y_2) \ \tau_u^k(y_2) \ \tau_s^k(y_2)], k = H, I$$

Definition 3 Cut-off income bid rent functions, denoted $\rho_i^k(y_2)$, $i = 1, 2$, $k = H, I$, provide for any type 2 cut-off income 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 those in $\mathbf{\Pi}^k(y_2)$ if local governments rely on anonymous taxation, or those chosen by the planner otherwise.

For instance, with anonymous taxation, cut-off income bid rent functions are implicitly defined by setting $r_s = 0$ in the indifference conditions:

$$v_u^1(e_u, \tau_u^k, \rho_1^k(y_2); z(y_2)) = v_s^1(e_s, \tau_s^k, r_s; z(y_2)) \quad (7)$$

$$v_u^1(e_u, \tau_u^k, \rho_2^k(y_2); y_2) = v_s^2(e_s, \tau_s^k, r_s; y_2). \quad (8)$$

Importantly, note that because $z(y)$, $e_j(y)$ and $\tau_j^k(y)$ 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.¹⁹ The SWF includes the utility of the absentee

¹⁹ 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.

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.²⁰ The indirect utility function is thus:

$$V_j^i(e_j, T_j^i, r_j, y, R_i(y)) = y + R_i(y) - T_j^i - r_j + h_i(e_j, y), \quad (9)$$

while the SWF is:

$$\begin{aligned} & \int_{\underline{y}}^{y_1} y + R_1(y) - T_s^1 - r_s + h_1(e_s, y) \gamma \phi_1(y) dy + \\ & \int_{y_1}^{\bar{y}} y + R_1(y) - T_u^1 - r_u + h_1(e_u, y) \gamma \phi_1(y) dy + \\ & \int_{\underline{y}}^{y_2} y + R_2(y) - T_s^2 - r_s + h_2(e_s, y) (1 - \gamma) \phi_2(y) dy + \\ & \int_{y_2}^{\bar{y}} y + R_2(y) - T_u^2 - r_u + h_2(e_u, y) (1 - \gamma) \phi_2(y) dy + \\ & + [R_L + r_s H_s + r_u H_u]. \end{aligned} \quad (10)$$

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 housing market constraint and two local governments budget constraints.²¹ The housing market constraint has λ_h as its multiplier and, by Lemma 1, can be written as:

$$y_1 - z(y_2) = 0. \quad (11)$$

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

$$e_j [c_1 n_j^1(y_1) + c_2 n_j^2(y_2)] - T_j^1 n_j^1(y_1) + T_j^2 n_j^2(y_2) = 0; j = u, s. \quad (12)$$

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:

$$MSV_i(y) = h_i(e_u, y) - h_i(e_s, y) + c_i [e_s - e_u]; i = 1, 2.$$

²⁰ This is only for completeness; quasi-linearity of the utility function implies that such transfers do not affect aggregate welfare if this is given by the unweighted SWF considered.

²¹ The tenth constraint requires the transfers' budget to balance :

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

and has multiplier λ_R .

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.²² The following proposition characterises the solution to the SPP.

Proposition 1 *A solution to the Social Planner's Problem with school qualities $e_u^* > e_s^*$ exhibits WTS. Furthermore:*

i) In an interior solution, cut-off incomes y_1^, y_2^* satisfy: $MSV_1(y_1^*) = MSV_2(y_2^*)$, the housing market constraint and $y_1^* < y_2^*$.*

ii) In a corner solution, y_1^, y_2^* satisfy: $MSV_1(y_1^*) \geq MSV_2(y_2^*)$, the housing market constraint and $y_1^* < y_2^*$.*

iii) School qualities satisfy the Samuelsonian conditions:

$$c_1 n_u^1(y_1^*) + c_2 n_u^2(y_2^*) = \gamma \int_{y_1^*}^{\bar{y}} \frac{\partial h_1(e_u^*, y)}{\partial e_u} \phi_1(y) dy + (1 - \gamma) \int_{y_2^*}^{\bar{y}} \frac{\partial h_2(e_u^*, y)}{\partial e_u} \phi_2(y) dy \quad (13)$$

$$c_1 n_s^1(y_1^*) + c_2 n_s^2(y_2^*) = \gamma \int_{\underline{y}}^{y_1^*} \frac{\partial h_1(e_s^*, y)}{\partial e_s} \phi_1(y) dy + (1 - \gamma) \int_{\underline{y}}^{y_2^*} \frac{\partial h_2(e_s^*, y)}{\partial e_s} \phi_2(y) dy. \quad (14)$$

Proof. Deriving the MSV functions with respect to y , it is straightforward to check that the normality of education implies that MSV functions are increasing in income, which entails that the solution satisfies WTS. The FOCs on T_j^i are:

$$-n_j^i(y_i^*) + \lambda_j^* n_j^i(y_i^*) = 0; \quad j = u, s, \quad i = 1, 2, \quad (15)$$

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

$$\begin{aligned} [h_1(e_u^*, y_1^*) - h_1(e_s^*, y_1^*) - c_1(e_u^* - e_s^*)] \gamma \phi_1(y_1^*) &= -\lambda_h^* \\ [h_2(e_u^*, y_2^*) - h_2(e_s^*, y_2^*) - c_2(e_u^* - e_s^*)] (1 - \gamma) \phi_2(y_2^*) &= -\frac{\lambda_h^* (1 - \gamma) \phi_2(y_2^*)}{\gamma \phi_1(y_1^*)} \end{aligned}$$

which together imply that $MSV_1(y_1^*) = MSV_2(y_2^*)$ must hold in an interior solution. Next, note that, since $e_u^* > e_s^*$, Assumption 1 implies $MSV_1(\tilde{y}) >$

²² Following Calabrese et al. (2010), 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.

$MSV_2(\tilde{y})$, where recall \tilde{y} satisfies $\tilde{y} = z(\tilde{y})$. Since the MSV functions are rising in income, then, the optimal cut-off incomes must satisfy $y_1^* < \tilde{y} < y_2^*$. If there is no $y \in (\tilde{y}, \bar{y}_2)$ such that $MSV_1(z(y)) = MSV_2(y)$, a corner solution emerges. In that case, since $MSV_1(\tilde{y}) > MSV_2(\tilde{y})$ and MSV functions are continuous, $MSV_1(z(\bar{y}_2)) \geq MSV_2(\bar{y}_2)$ and thus it is optimal to have the urban district populated exclusively by type 1 households if $\gamma \geq H_u$, or the suburbs exclusively by type 2 households if $\gamma < H_u$. Finally, the FOCs corresponding to the school quality variables e_u and e_s yield the usual Samuelsonian conditions (13) and (14). ■

An interior solution has households of the two types allocated to both districts, as in panel (a) of Figure 1; panels (b), where $\gamma > H_u$, and (c), where $\gamma < H_u$, of the same figure correspond to corner solutions. The equality of the MSV functions in an interior optimum implies the following optimality condition:

$$[h_1(e_u^*, y_1^*) - h_1(e_s^*, y_1^*)] - [h_2(e_u^*, y_2^*) - h_2(e_s^*, y_2^*)] = \Delta_c [e_s^* - e_u^*], \quad (16)$$

On the other hand, the Samuelsonian conditions that determine the optimal school qualities equate districts' average marginal rates of substitution between school quality and numeraire consumption to their respective marginal cost of school quality $(n_j^1 c_1 + n_j^2 c_2)$.

Remark 1 *Income mixing is optimal. Households of type 1 with incomes between y_1^* and y_2^* are allocated to the urban area, while households of type 2 and identical income are assigned to the suburbs in the optimal allocation. The explanation is doublefold as type 1 pupils impose smaller congestion costs onto and derive greater benefits from the better quality school.*²³

5 Decentralising the optimal allocation

In order to decentralise the optimal allocation, the planner must choose local tax variables that satisfy the local governments' budget constraints and ensure that, in the emerging equilibrium, households derive (weakly) higher utility in their socially optimal location. Bearing in mind that the optimal allocation satisfies the WTS property, optimal tax combinations will need to make households with cut-off incomes indifferent between the two residential alternatives. That requirement will generate a *location-incentives constraint*.

²³ Suppose assumption 1 does not hold but that, instead, preferences satisfy $\partial h_1(e, y) / \partial e < \partial h_2(e, y) / \partial e$. In that case, though some income mixing will be optimal (except in special circumstances) which of the cut-off incomes should be smaller is ambiguous. The reason is that, while the lower costs of educating type 1 children tends to make optimal that $y_1 < y_2$, the greater benefit type 2 children derive from school quality has the opposite effect.

Two-single crossing conditions will then guarantee that the remaining households strictly prefer their socially optimal location. The analysis in this section restricts attention to interior solutions to the SPP for, as proposition 5 below demonstrates, corner solutions will be sustained as an equilibrium with anonymous head taxes.

5.1 Differentiated head taxes

With differentiated head taxes, the local budget constraints are:

$$E_j = T_j^1 n_j^1 + T_j^2 n_j^2, \quad j = u, s, \quad (17)$$

where $E_j = e_j (n_j^1 c_1 + n_j^2 c_2)$ is district's j total spending, whilst the indirect utility functions are:

$$v_j^i(e_j, T_j^i, r_j, y) = y - T_j^i - r_j + h_i(e_j, y) \quad (18)$$

Substituting (18) into the indifference condition (3) and normalising the rent of the suburbs to $r_s = 0$, one can derive the head-tax bid-rent functions (denoted r_i^H):

$$r_i^H(y, e_u, e_s, T_s^i, T_u^i) = h_i(e_u, y) - h_i(e_s, y) + T_s^i - T_u^i, \quad i = 1, 2. \quad (19)$$

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:*

$$\begin{aligned} v_i^s(e_s, T_s^i, y_i) &= v_i^u(e_u, T_u^i, r_u; y_i) \Rightarrow \\ v_i^s(e_s, T_s^i, y) &< v_i^u(e_u, T_u^i, r_u; y); \quad \forall y > y_i \\ v_i^s(e_s, T_s^i, y) &> v_i^u(e_u, T_u^i, r_u; y); \quad \forall y < y_i, \quad i = 1, 2. \end{aligned} \quad (20)$$

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

Next, in order to derive the *location-incentives constraint*, let

$$\Delta_1^h(y_2) = h_1(e_u(y_2), z(y_2)) - h_1(e_s(y_2), z(y_2))$$

and

$$\Delta_2^h(y_2) = h_2(e_u(y_2), y_2) - h_2(e_s(y_2), y_2)$$

denote the gap in human capital each type's cut-off households obtain from attending the urban school instead of the suburban one, when school qualities are determined optimally. The cut-off income bid-rent functions, derived from equations (7) and (8), can then be written as:

$$\rho_1^H(y_2) = \Delta_1^h(y_2) + T_s^1 - T_u^1 \quad (21)$$

$$\rho_2^H(y_2) = \Delta_2^h(y_2) + T_s^2 - T_u^2 \quad (22)$$

If an allocation $y_2^H \in (\tilde{y}, \bar{y}_2)$ is to be sustained as an *interior head-tax equilibrium*, households with cut-off income must be indifferent between the two districts. That condition requires the equilibrium urban rent premium to satisfy $r_u^H = \rho_1^H(y_2^H) = \rho_2^H(y_2^H)$, implying:

$$\Delta_1^h(y_2^H) - \Delta_2^h(y_2^H) = (T_s^1 - T_u^1) - (T_s^2 - T_u^2) \quad (23)$$

The *location-incentives constraint* is obtained by setting $y_2^H = y_2^*$ in the head-tax equilibrium condition (23) and subtracting it from the optimality condition (16), which yields:

$$\Delta_c[e_s(y_2^*) - e_u(y_2^*)] = (T_u^1 - T_s^1) - (T_u^2 - T_s^2). \quad (24)$$

Combinations of head tax bills that fulfill the previous equation ensure that, when school qualities are given by $e_j(y_2^*)$, the maximum bids households of each type with the optimal cut-off income are willing to offer for a house in the urban area coincide; in other words, that there exists a level of the rent premium that makes them simultaneously indifferent between the two districts.

Definition 4 Let $\Omega^H(y_2^*)$ be the set of all combinations of differentiated head taxes T_j^i satisfying the location-incentives constraint (24) and the local budget constraints (17) at the optimal allocation.

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 the set $\Omega^H(y_2^*)$. In that case:

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

and it is straightforward to check that the budget and the location-incentives constraints are satisfied. Interestingly, infinitely many other combinations of differentiated head taxes lead to the optimal equilibrium as well.

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

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 by 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 (20) 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 (24). ■

Remark 2 *In a model with housing markets, many combinations of differentiated head taxes can attain optimality. These need neither cover the marginal cost of entry of a household nor be greater for households with lower ability children. Instead, they must provide households with the correct relative location incentives. In other words, they must generate differences in the cost of entry each type of household faces to live in the rich district that internalise the externalities emerging from the peer effect. The result opens the door for differentiated head taxes to effect some redistribution across household types.*

5.2 Income taxes

Under differentiated income taxation, the local budget constraint of district j is:

$$e_j (n_j^1 c_1 + n_j^2 c_2) = t_j^1 n_j^1 Y_j^1 + t_j^2 n_j^2 Y_j^2, \quad j = u, s. \quad (25)$$

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

$$v_j^i(e_j, t_j^i, r_j; y) = y(1 - t_j^i) - r_j + h_i(e_j, y), \quad (26)$$

while the income-tax bid rent functions, obtained as before, are:

$$r_i^I(y, e_u, e_s, t_s^i, t_u^i) = h_i(e_u, y) - h_i(e_s, y) + y(t_s^i - t_u^i); \quad i = 1, 2. \quad (27)$$

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*

$$\left[\frac{\partial h_i(e_u, y)}{\partial y} - \frac{\partial h_i(e_s, y)}{\partial y} \right] > (t_u^i - t_s^i) \quad \forall y \in S, \quad i = 1, 2. \quad (28)$$

If inequality (28) holds, households induced preferences satisfy:

$$\begin{aligned} v_i^s(e_s, t_s^i, y_i) &= v_i^u(e_u, t_u^i, r_u; y_i) \Rightarrow \\ v_i^s(e_s, t_s^i, y) &< v_i^u(e_u, t_u^i, r_u; y); \quad \forall y > y_i \\ v_i^s(e_s, t_s^i, y) &> v_i^u(e_u, t_u^i, r_u; y); \quad \forall y < y_i, \quad i = 1, 2. \end{aligned} \quad (29)$$

A sufficient 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 large enough relative to the tax rate differential.

The cut-off bid-rent functions (denoted ρ_i^I) are again deduced from equations (7) and (8), yielding:

$$\rho_1^I(y_2) = \Delta_1^h(y_2) + z(y_2) [t_s^1 - t_u^1] \quad (30)$$

$$\rho_2^I(y_2) = \Delta_2^h(y_2) + y_2 [t_s^2 - t_u^2] \quad (31)$$

If an allocation $y_2^I \in (\tilde{y}, \bar{y}_2)$ can be sustained as an interior income-tax equilibrium, the rent premium must satisfy $r_u^I = \rho_1^I(y_2^I) = \rho_2^I(y_2^I)$, which requires:

$$\Delta_1^h(y_2^I) - \Delta_2^h(y_2^I) = y_2^I [t_s^2 - t_u^2] - z(y_2^I) [t_s^1 - t_u^1]. \quad (32)$$

Therefore, the *location-incentives constraint*, obtained as before, is:

$$\Delta_c [e_s(y_2^*) - e_u(y_2^*)] = z(y_2^*) [t_u^1 - t_s^1] - y_2^* [t_u^2 - t_s^2]. \quad (33)$$

Combinations of income tax rates that satisfy this equation make households of each type with the optimal cut-off income indifferent between the two districts at the equilibrium rent premium when school qualities are given by $e_j(y_2^*)$.

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 (33) and the local budget constraints (25) at y_2^* .*

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

$$\Delta_c [e_u(y_2^*) - e_s(y_2^*)] \leq y_2^* \bar{t}_u^2 + y_1 \bar{t}_s^1 \quad (34)$$

Proof. It is straightforward to check that (33) and the two local budget constraints (25) 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 and the budget constraints require: $t_j^i \in [0, \bar{t}_j^i]$, with $\bar{t}_j^i = \frac{E_j}{B_j^i} \leq 1$; if $B_j^i < E_j$ then $t_j^i \in [0, 1]$ and $t_j^{-i} \in [\underline{t}_j^i, \bar{t}_j^i]$ where $\underline{t}_j^i = \frac{B_j^{-i}}{B_j^i} \left(\frac{E_i}{B_j^{-i}} - 1 \right) > 0$. Also, let $t_u^2 = f(t_u^1, y_2^*)$ and $t_s^2 = g(t_s^1, y_2^*)$ denote the local government budget constraints and substitute them into the location-incentives constraint:

$$y_1^* t_u^1 - y_2^* f(t_u^1, y_2^*) = \Delta_c [e_s^* - e_u^*] + y_1^* t_s^1 - y_2^* g(t_s^1, y_2^*). \quad (35)$$

Express each side of the equation as a function of t_j^1 , $\psi_j(t_j^1)$, $j = u, s$. These functions are 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 $\psi_s(0) = \Delta_c [e_s^* - e_u^*] - y_2^* \bar{t}_s^2$, and a maximum at $t_j^1 = \bar{t}_j^1$, where $\psi_u(\bar{t}_u^1) = y_1^* \bar{t}_u^1$ and $\psi_s(\bar{t}_s^1) = \Delta_c [e_s^* - e_u^*] + y_1^* \bar{t}_s^1$. Consequently, letting Υ_j be the set of tax rates imposed by district j on type 1 households, t_j^1 , that fulfills equation (35), i.e. such that $\psi_j(t_j^1) \in [\psi_{-j}(0), \psi_{-j}(\bar{t}_{-j}^1)]$, for any element of Υ_u , there exists a unique element in Υ_s such that (35) holds. Finally, note that the sets Υ_j , and thereby also $\Omega^I(y_2^*)$, are empty sets if the images of ψ_u and ψ_s do not overlap. Otherwise, that is if (34) is satisfied, Υ_j and $\Omega^I(y_2^*)$ are non-empty. ■

Lemma 4 does not impose a restrictive condition: the RHS of (35) is the sum of the taxes a type 1 cut-off household would pay in district s if type 2 residents did not pay any and those a type 2 cut-off household would pay in district u if their type 1 neighbours did not pay any. That sum must 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). Nevertheless, the elements of $\Omega^I(y_2^*)$ may not sustain an optimal equilibrium.

Proposition 3 Consider an interior solution to the Social Planner Problem with optimal cut-off incomes (y_1^*, y_2^*) and satisfying $e_u^* > e_s^*$. A combination of local income tax rates in $\Omega^I(y_2^*)$ sustains the optimal allocation as an equilibrium with rent premium $r^I = \rho_1^I(y_2^*) = \rho_2^I(y_2^*)$ and cut-off incomes $y_1^I = y_1^*$, $y_2^I = y_2^*$ if and only if the two pairs of tax rates (t_u^i, t_s^i) , $i = 1, 2$, satisfy Lemma 3.

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

Remark 3 *The existence problem emerges because $\Omega^I(y_2^*)$ may be an empty set and, even if it is not, it may be that no element satisfies the necessary single-crossing conditions.²⁴ Therefore, the optimal allocation may not be sustainable as a market equilibrium with differentiated income taxes.*

Notwithstanding, the next proposition proves that it is possible to combine anonymous income taxes with another fiscal tool to correct for their distortionary location effects. If the anonymous income tax rates that balance the local governments' budgets satisfy the single-crossing conditions, then the location externalities can be internalised with a self-funded lump-sum transfer scheme among households of different types that live in the urban district. If they do not satisfy them, it must be the case that $t_u(y_2^I) > t_s(y_2^I)$. Then, the proposal involves applying the suburbs' tax-rate to the urban area and imposing differentiated head taxes to correct for the location externalities and to fund the resulting deficit, D_u .²⁵

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*

$$L_1 n_u^1 - L_2 n_u^2 = 0 \quad (36)$$

$$-L_1 - L_2 = \Delta_c [e_s^* - e_u^*] - (y_1^* - y_2^*) (t_u - t_s) \quad (37)$$

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*

²⁴ 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 also need to satisfy the relevant single-crossing conditions.

²⁵ 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.

differentiated head-tax scheme imposed on urban residents $(\widehat{T}_1, \widehat{T}_2)$ satisfying

$$\widehat{T}_1 n_u^1 + \widehat{T}_2 n_u^2 = D_u \quad (38)$$

$$\widehat{T}_1 - \widehat{T}_2 = \Delta_c [e_s^* - e_u^*] \quad (39)$$

sustains the optimal allocation as an income tax equilibrium.

Proof. Clearly, both systems of equations have a unique solution. Equations (36) and (38) 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 (33) reduces to

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

Therefore, 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 (37) 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 (39) imposes. ■

Remark 4 *This result demonstrates that, while single-crossing conditions mark the limits to the compatibility between optimality, on one side, and income taxes and the implied redistribution, on the other, these are not incompatible with each other. As a matter of fact, the second proposal sets a lower bound for the amount of tax redistribution effected in the rich district, as the single-crossing conditions could be satisfied for some $t_u > t_s(y_2^*)$.*

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 differentiated head taxes that complement the uniform income tax scheme in the second proposal need to be greater for type 2 households.

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 $(n_j^1 c_1 + n_j^2 c_2)$ 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 and that income taxes can easily match the outcome of head taxes.

6.1 Head taxes

In this case, the budget constraints, indirect utility, bid-rent and cut-off bid-rent functions are obtained by setting $T_j^1 = T_j^2 = T_j$ in (17), (18), (19), and $T_j^1 = T_j^2 = T_j(y_2)$ in (21) and (22). The next proposition reveals that an interior head tax equilibrium is necessarily suboptimal, while a corner one is optimal.

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

such that $e_u(y) = e_s(y)$, it is straightforward to check that such allocation of households to districts, y , the policy vector $\Pi_H(y)$ and the rent premium $r_H = T_s(y) - T_u(y)$ leave every household indifferent between the two districts and, hence, constitute an equilibrium.

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

$$\Delta_1^h(y_2^*) - \Delta_2^h(y_2^*) = -\Delta_c(e_u^* - e_s^*). \quad (40)$$

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

$$\Delta_1^h(y_2^H) - \Delta_2^h(y_2^H) = 0. \quad (41)$$

Because the RHS of (40) is negative: $\rho_1^H(y_2^*) < \rho_2^H(y_2^*)$ and $MSV_1(y_2^H) > MSV_2(y_2^H)$. The latter implies that any interior head-tax equilibrium is sub-optimal (i.e. that $y_2^H \neq y_2^*$). The former, along with the fact that $\rho_1^H(\tilde{y}) > \rho_2^H(\tilde{y})$, the continuity of the cut-off bid rent functions and the intermediate value theorem, entail the existence of a level of income $y_2^H \in (\tilde{y}, y_2^*)$ 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:

$$\Delta_1^h(\bar{y}_2) - \Delta_2^h(\bar{y}_2) \geq -\Delta_c(e_u^* - e_s^*). \quad (42)$$

Because the RHS of (42) is negative, its LHS may be negative or positive. In the former case, $\rho_1^H(\bar{y}_2) - \rho_2^H(\bar{y}_2) < 0$ and, as proved above, there is a level of income $y_2^H \in (\tilde{y}, y_2^*)$ 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

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

so that existence of a corner head-tax equilibrium implies that (42) holds and so the existence of a corner solution to the SPP, and viceversa. ■

Remark 5 *Interior equilibria emerging with anonymous head taxes generate a suboptimal distribution of households across districts. Residential choices generate negative externalities because the homogenous tax-bill levied on the urban residents does not cover the marginal costs of admitting them into the district, being too low for high-cost (type 2) households and too high for low-cost (type 1) 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 $t_j^1 = t_j^2 = t_j$ in (25), (26), (27), and $t_j^1 = t_j^2 = t_j(y^2)$ in (30) and (31).

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. Following a similar argument as in the proof of the previous proposition, it can be shown that, if $e_u(y) > e_s(y) \forall y \in [\tilde{y}, \bar{y}_2]$, then either there is a level of income $y \in (\tilde{y}, \bar{y}_2)$ for which either $\rho_1^I(y) = \rho_2^I(y)$, or $\rho_1^I(\bar{y}_2) \geq \rho_2^I(\bar{y}_2)$. In both cases, equilibrium requirements E2-E4 hold but the implied allocation will only 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:

$$\Delta_1^h(y_2^I) - \Delta_2^h(y_2^I) = (z(y_2^I) - y_2^I)(t_u(y_2^I) - t_s(y_2^I)). \quad (43)$$

The comparison between (43) and the optimality requirement (16) confirms that anonymous income taxes lead to a suboptimal allocation of households across districts. The next result shows that head taxes may induce greater welfare losses than income taxes. Figure 2 illustrates this possibility.

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

Proof. In an interior income-tax equilibrium, cut-off incomes satisfy (43). In turn, in an interior head-tax equilibrium, they fulfill $\rho_1^H(y_2^H) = \rho_2^H(y_2^H)$. Because $t_u(y_2^I) > t_s(y_2^I)$ by assumption, the RHS of (43) is negative, implying $\rho_1^H(y_2^I) < \rho_2^H(y_2^I)$. Using again the fact that $\rho_1^H(\tilde{y}) > \rho_2^H(\tilde{y})$, continuity of the cut-off bid rent functions and the intermediate value theorem imply the existence of a level of income $y_2^H \in (\tilde{y}, y_2^I)$ such that $\rho_1^H(y_2^H) = \rho_2^H(y_2^H)$ and for which a head-tax equilibrium exists. Therefore, $y_2^* > y_2^I > y_2^H$, $y_1^* < y_2^I < y_1^H$. ■

Furthermore, anonymous income taxes can easily match the outcome achieved with anonymous head taxes.

Proposition 7 *The results in proposition 5 apply to an income tax scheme that sets t_u equal to $t_s(y)$ and funds (gives back) the resulting urban budget*

deficit (surplus) through a uniform head tax levied on (with a uniform transfer paid back to) urban residents.

Proof. Because $t_u = t_s(y)$, the single-crossing conditions are satisfied and the RHS of (43) is equal to zero so that:

$$\rho_1^I(y) - \rho_2^I(y) = \rho_1^H(y) - \rho_2^H(y).$$

■

Remark 6 *Results in this section clash with the view of local head taxes as welfare-enhancing benefit taxes. When governments cannot observe the cost-parameters or use that information to tax-discriminate across households of different types, head taxes may be more distortionary than an ability-to-pay tax such as the proportional income tax. The reason is that, if an income tax equilibrium has higher tax rates in the urban area 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 the former impose on the rest are (partially) internalised.*²⁶

Remark 7 *Moreover, income taxation can be combined with a uniform head tax levied on urban residents, or with a lump-sum transfer to them to match the outcome of head taxes. Therefore, the implied income redistribution does not generate more welfare losses than those resulting from the use of anonymous head taxes.*

7 Concluding remarks

The analysis in this paper offers new insights into the relative normative merits of local head and income taxes. The main novel results reveal that head taxes are not superior to income taxes and that the indirect redistribution implied by income taxation is not necessarily at odds with location optimality or associated to welfare losses. In cases where local governments can observe the cost parameters of different household types and tax-discriminate across them, both head and income taxes are able to sustain the optimal allocation in equilibrium. Remarkably, optimal differentiated head taxes need not cover marginal congestion costs, as the outcome of the location game depends on the relative willingness to pay for entering the good district. Because optimality

²⁶ It is important to stress that the essence of this result does not depend on the assumptions made over household types. In particular, an analogous result can be derived if households imposing smaller costs onto schools also derive smaller benefits from school quality.

requires the income segregation of households of the same type in the utilitarian normative framework considered, the necessary single-crossing conditions for a segregated equilibrium to exist mark the limits of the compatibility between income taxes and location optimality but do not rule it out. At most, it may be necessary to set anonymous and identical tax rates in the two districts and cover the budget deficit that results in the rich district by imposing differentiated head taxes on its residents. When differentiated taxes are not available, the two tax systems cannot be unambiguously ranked according to the welfare losses they generate as compared to the optimal outcome. Moreover, supplemented with head taxes levied on, or lump-sum transfers to, the residents of the rich district, income taxes could always match the outcome attained with head taxes.

It is important to check the robustness of these conclusions to alternative specifications of the peer effect and of housing markets, that is to say, to relax the assumptions of linear crowding costs and of inelastic housing supplies. 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 peer 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 peer effects, the public sector could use exams and condition access to schools on the results to derive welfare gains.

Acknowledgements

I am very grateful for many useful comments to Subir Bose, Phillippe 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 Universitat Rovira i Virgili and to the Department of Economics at Universitat Autònoma de Barcelona (UAB) where part of this research was completed. All remaining errors are my own.

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 crowding 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] Duflo, E., Dupas, P and Kremer, M., forthcoming. Peer Effects and the Impacts of Tracking: Evidence from a Randomized Evaluation in Kenya. *American Economic Review*.
- [10] 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.
- [11] 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.
- [12] 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.
- [13] 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.
- [14] Goodspeed, T. J., 1989. A reexamination of the use of ability to pay taxes by local governments. *Journal of Public Economics*, 38: 319-342.
- [15] Goodspeed, T. J., 1995. Local income taxation: an externality, pigouvian solution and public policies. *Regional Science and Urban Economics*, 25: 279-296.
- [16] Hamilton, Bruce W., 1975. Zoning and Property Taxation in a System of Local Governments, *Urban Studies*, 12: 205-211.

- [17] Kessler, A. and Lülfeßmann, C., 2005. Tiebout and redistribution in a model of residential and political choice. *Journal of Public Economics*, 89: 501-528.
- [18] Lavy, V., Silva, O. and Weinhardt, F. 2009. The Good, The Bad and The Average: Evidence on the Scale and Nature of Ability Peer Effects in School. NBER Working Paper 15600.
- [19] Lazear, E. P., 2001. Educational production, *Quarterly Journal of Economics*, 116 (3): 777-803.
- [20] Nechyba, T., 1999. School-finance induced migration patterns: The case of private school vouchers. *Journal of Public Economic Theory*, 1: 5-50.
- [21] Rose-Ackerman, S., 1979. Market models of local governments: exit, voting and the land market. *Journal of Urban Economics*, 6: 319-337.
- [22] Ross, S. and Yinger, J., 1999. Sorting and voting: A review 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.
- [23] Schmidheiny, K., 2006a. Income segregation and local progressive taxation: Empirical evidence from Switzerland. *Journal of Public Economics*, 90: 429-458.
- [24] 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.
- [25] 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.
- [26] Schwabb, R. M. and Oates, W. E., 1991. Community composition and the provision of local public goods: A normative analysis. *Journal of Public Economics*, 44: 217-237.
- [27] Tiebout, C. M., 1956. A pure theory of local expenditures. *Journal of Political Economy*, 64(5): 416-424.
- [28] Wildasin, D., 1986. *Urban public finance*. Harwood Academic Publishers, New York.
- [29] Wooders, M., 1978. Equilibria, the core, and jurisdiction structures in economies with a local public good. *Journal of Economic Theory*, 18: 328-348.
- [30] Wooders, M., 1999. Multijurisdictional economies, the Tiebout hypothesis and sorting. *Proceedings of the National Academy of Science*, 96: 10585-10587.