The desegregating effect of school tracking[*]

Gianni De Fraja[†]

University of Nottingham

Università di Roma "Tor Vergata"

and C.E.P.R.

Francisco Martínez-Mora[‡]

University of Leicester

Universitat Rovira i Virgili

and C.R.E.I.P

Abstract

This paper makes the following point: "detracking" schools, that is preventing them from allocating students to classes according to their ability, may lead to an increase in income residential segregation. It does so in a simple model where households care about the school peer group of their children. If ability and income are positively correlated, tracking implies that some high income households face the choice of either living in the areas where most of the other high income households live and having their child assigned to the low track, or instead living in lower income neighbourhoods where their child would be in the high track. Under mild conditions, tracking leads to an equilibrium with partial income desegregation where perfect income segregation would be the only stable outcome without tracking.

JEL Numbers: I24, H42.

Keywords: Tracking, school selection, income segregation, school choice, Tiebout.

1 Introduction

Tracking is the practice of allocating the pupils of a school to different classes according to their academic ability. Tracking is highly controversial and has generated an often heated academic and policy debate;[1] this has typically focused on the effects of tracking on educational attainment and other students' outcomes, such as post-education earnings.[2] In this paper we look beyond the educational output, and examine the effects of tracking on the degree of income segregation in residential areas and their schools. Central to the paper is the idea that the characteristics of local schools are an important determinant of households' location choices.[3] One such characteristic is whether or not the local school tracks its students. Tracking affects the peer group, an important input in the educational production function, and thereby it becomes one of the determinants of households' location choices and hence of the socio-economic composition of a residential area and its schools. In turn, these choices are relevant to society, because, for example, a residential pattern where households of different socio-economic background live near one another reduces ghettoes, exposes disadvantaged adolescents to lifestyles, behaviours and ambitions typical of classmates and friends from more disparate social backgrounds, and might enhance social mobility. Understanding the nature of the link between schools' policies regarding tracking and residential income segregation in a given geographical area becomes therefore very important. Our paper is a step in this

---

[1]The early analysis of Coleman and his co-authors (Coleman et al 1966) already considers the effects of tracking; the turning point towards "detracking" is discussed in Wheelock (1992) and Argys et al (1996). A comprehensive survey of the initial debate among educationalists is Lucas (1999).

[2]Betts (2011) reviews the empirical literature on the effects of tracking, and Brunello and Checchi (2007) and Hanushek and Wßmann (2006) provide an up-to-date overview of the international differences in extent and implementation of tracking.

[3]Here again the literature is vast; Calabrese et al (2012) build a general equilibrium model to study the welfare implications of Tiebout sorting (1956). Using a similar setting, Nechyba (1999) investigates the effect of private schools on residential segregation and on school segregation in a metropolitan area. De Bartolome and Ross (2003, 2004, 2007) and Hanushek and Yilmaz (2010) analyse the interactions between Tiebout type incentives and the trade-off between geographical access and land space first studied by Alonso (1964). Recent studies of the link between school performance and housing prices are Downes and Zabel (2002), Dhar and Ross (2012), Clapp et al (2008), Gibbons and Machin (2003) and Bayer et al (2007). Black and Machin (2011) is an extensive review of the empirical literature.

hitherto untrodden direction.

Our main result is straightforward and simply stated: tracking may weaken income segregation, that is the tendency of households to cluster according to income and socio-economic status. Our paper contributes to the policy debate on tracking by showing that the trend towards "detracking American schools" (Argys et al 1996) might well have the probably unintended consequence of exacerbating income segregation and of thus hampering social mobility.

Empirically, income segregation has long been observed not to match the complete stratification predicted by a nave Tiebout-type location model (Pack and Pack 1977, Persky 1990, Calabrese et al 2006). Explanations for the high degree of income mixing have ranged from a two-dimensional distribution of households' characteristics (Epple and Platt 1998, Epple and Romano 2003), to the interactions between income differences in commuting costs and the strength of the preference for public goods (de Bartolome and Ross 2003, 2004, 2007), to the way in which the marginal rate of substitution between commuting and housing varies with income (LeRoy and Sonstelie 1983). Our findings add a further possible explanation for the observed level of income mixing, and they are in line with some recent empirical evidence showing that social mobility is larger when schools select students by past performance rather than by residential location (Lee 2011a and 2011b).

We build a simple stylised model. Households choose where to live, and property prices adjust to demand and supply. The quality of the education received by their children is one of the variables which influences households' location choices. We compare two alternative policy scenarios, one where schools track students, the other where the allocation of students to classes is random. We show that when schools do not track students, the equilibrium is such that households are fully segregated by income: all the poor live in one district and all the rich live in the other. On the other hand, when schools track their students, in equilibrium both rich and poor live in both districts, and their children attend the same schools. The intuition for this "desegregation" equilibrium is easily explained. It hinges on two linchpins, both solidly established in the literature: the peer group effect[4] and the positive correlation

---

[4]Intuition and casual observation suggest that children learn from each other, because they help, or hamper, one another, because they stimulate each other, because they compete to do well, and so on. Moreover, when schools track students, classes comprise students of

2

between a child's ability and her socio-economic background.[5] When the peer group a child has at school matters to her parents, and when there is a positive correlation between ability and socio-economic background, parents from a high socio-economic background whose children are of middling ability face a dilemma: they have to choose between living in a district with many other households of good socio-economic background where however their child is likely to be placed in the low track, and living in a district with fewer households of high socio-economic background, but a higher chance that their child is in the top track at school, and thus benefits from a higher ability peer group. This dilemma is similar in nature to the choice that parents face in Epple et al (2002), where profit-maximising private schools compete with homogenous public schools in a given district. When public schools track their students, they attract more high ability students while losing some lower ability ones from richer households to private schools. The children of parents who opt for those private schools have a lower quality peer group than the public school high track but a higher quality peer group than the public school low track.[6] Lest contemplation of this dilemma be considered beyond households' actual behaviour, note the intriguing evidence revealed by Cullen et al (2013) and Estevan et al (2012). These papers show that indeed households do behave strategically to benefit from school policies: students in Texas "trade down",

similar abilities, and teachers are less likely to slow down or repeat their lessons to make sure weaker students keep up, and can press ahead with the syllabus instead. Winkler (1975) Arnott and Rowse (1987), de Bartolome (1990) are early economic analyses of the impact of peer group effects; Astin (1993) an influential education one. There is also ample empirical evidence documenting their importance. Bishop (2006), Sacerdote (2011) and Epple and Romano (2011) are recent surveys of the vast economic literature.

[5]Sirin's meta-analysis of around 75 studies published in the 90's finds an average correlation of 0.299 (Sirin 2005, p 437), in line with the figure of 0.343 in the earlier studies considered by White (1982). An early economic analysis is Perl (1973). Notice that this assumption does not imply a positive correlation between innate ability and socio-economic background, as it could be the consequence of greater pre-school parental investment by better-off parents.

[6]Hidalgo (2010) compares tracking to a comprehensive school system, and, in her main result, finds that tracking may be the system providing greater equality of opportunities in the sense of Roemer (1998). Other theoretical contributions on the relative merits of selective versus comprehensive schooling systems include Brunello et al (2007), Eisenkopf (2009), Takii and Tanaka (2009) and Hidalgo (2011).

that is they choose a school with fewer able children in order to be more likely to be in the "Ten Percent" of ablest children in the school and so gain automatic admission to a state university.

Though the intuition for our main result might appear convincing, it is important to check that it is not unravelled by the simultaneous decisions of all households and by the operation of the property market. In Propositions 2 and 3, therefore, we establish necessary and sufficient conditions on the joint distribution of income and ability such that when schools practice tracking, households residential choices display income desegregation. As we argue, these conditions are not very stringent.

The paper is organised as follows. The model is presented in Section 2: the households in 2.1, the schools in 2.2, the housing market in 2.3. In Section 3, after some preliminaries and definitions, in 3.1, we derive in turn the equilibrium when schools do not track their students, subsection 3.2, and when they do: the desegregation equilibrium in 3.3, and the full segregation equilibrium in 3.4. Subsection 3.5 briefly discusses the intuition underlying our results and their possible consequences, and Section 3.6 carries out limited welfare comparisons. Section 4 concludes, and the Appendices contain some mathematical details.

2 The model

2.1 The households

We study a given population of households, with size normalised to 1, living in a stylised city with two geographically separate neighbourhoods, or districts, labelled 0 and 1. Households differ in income (a shorthand term for socio-economic background) and in the ability of their children. Both income and ability are exogenously given. We assume that income can take only two values, y_R and $y_P < y_R$; a proportion α of the households has income y_R, and the rest has income y_P. For the sake of brevity we will often refer to households with income y_R and y_P as "rich" and "poor". Ability is measured by a uni-dimensional parameter $b \in \left[\underline{b}, \overline{b}\right]$. We choose a simplified model in order both to present our result as starkly as possible, and also to show that the more "complicated" equilibrium where households residential choices lead to mixing of households with different socio-economic backgrounds can emerge even in

4

a highly simplified set-up, with the deck, as it were, stacked against complex outcomes.

Let $\Phi_i(b)$ be the distribution of households with income y_i: $\Phi_i(b)$ is the mass of households with income y_i where the child has ability b or less. $\Phi_i(b)$ satisfies: $\Phi_P(\underline{b}) = \Phi_R(\underline{b}) = 0$ and $\Phi_P(\overline{b}) = 1 - \alpha$ and $\Phi_R(\overline{b}) = \alpha$. We capture the assumption of a positive correlation between ability and income by imposing $\frac{\Phi_P(b)}{1-\alpha} > \frac{\Phi_R(b)}{\alpha}$ for every $b \in (\underline{b}, \overline{b})$: the ability distribution of children in high income households first-order stochastically dominates the ability distribution in low income households.

Households have identical preferences, represented by a utility function increasing in consumption, x, in the child's ability, b, and in the quality of the education the child receives, θ. This functional form is assumed often (e.g. Epple et al 2002; Caucutt 2002) and it captures the natural assumptions that parents care about their children's future prospects, that these are affected by their educational achievement, and that this in turn depends on their ability and the quality of the school attended.[7] Following a practice established since at least de Bartolome (1990), we simplify the utility function by imposing additivity in its three arguments: a household's utility is

$$U(x, a(\theta, b)) = v(x) + \theta + q(b). \tag{1}$$

with $v'(x) > 0$, $v''(x) < 0$, $q'(b) > 0$.

A restriction in (1) is that the marginal impact of an improvement in the peer group on achievement is constant; another is that bright and less bright children benefit equally from peer quality.[8] These restrictions make (1) suitable to the analysis of the paper, both because they make it more tractable, and because, as argued above, a simpler set-up brings out the main result more starkly. One plausible characteristic of (1) is the fact that the marginal rate of substitution between school quality and other consumption goods is increasing

---

[7]Given the static nature of the model we consider, it makes no difference to assume instead that households are affected by the child future income, as in Fernández and Rogerson (1998, p 816), or by her adult utility as in Loury (1981).

[8]Complementarities between ability and school quality would generate income mixing (Epple and Romano 2003). Ruling them out thus avoids confounding that effect with the income mixing effect of tracking. Furthermore, the available empirical evidence is still inconclusive as to whether the benefit of a better peer group is greater for bright or for less bright children (Betts 2011).

in income, which captures the important stylised fact that willingness to pay for school quality is higher for better-off households.

2.2 The schools

There is one school in each district. Attendance to school is free and compulsory, in the sense that children must attend the school in their district, which provides all the education they receive: we therefore rule out private schools and school choice within the state school sector, with the consequence that parents cannot choose a school, and hence an academic peer group, different from that of their local public school. This is clearly in contrast with the practice in many countries, but note that sending a child to a private school or to a faraway state school carries a cost in tuition fees or additional transportation costs (as modelled by, among others, Epple and Romano, 1998 and 2003 and Caucutt, 2002). Therefore it seems logical to postulate that, the higher this cost, the more likely parents are to send their child to the local state school. Our set-up therefore amounts to taking the extreme case for this plausible assumption, that this cost is prohibitively high.[9]

Although each school has a fixed size, this does not mean that children at the school all enjoy the same quality of education: typically, schools have many classes for each school-year, and how the school forms its classes determines a student's peer group, and hence affects his or her learning. "Tracking" is the practice of dividing students into groups of similar ability. Analytically, tracking amounts to having an increasing relationship between a child's ability and the quality of the education she receives. With tracking, this is a straightforward consequence of the fact that abler children enjoy a "better" peer group. To capture the peer group effect, we take θ, the quality of the education received, to be a function of the abilities of all the children in the track. With little further loss of generality, we follow the simplification of much of the lit-

---

[9]In the light of Calsamiglia et al (2013), our set-up is in fact more general: their analysis reveals that two widely used school choice mechanisms –the so-called Boston Mechanism and the Deferred Acceptance (Gale-Shapley) mechanism– generate the same incentives as the purely residence-based admission policies assumed here if schools give priority to local residents. Therefore our results apply equally to settings where school choice is possible but where schools give priority to local applicants. This is usually the case in practice.

erature (e.g. Epple and Romano 1998, Nechyba 2003, Hidalgo 2010)[10] and let θ simply denote the *average* ability of the students in a given student's class. We also assume that each school has two classes of equal size, and that it can observe the ability of its students.[11] Parents understand the school policy, and also know the ability of their child, and, in equilibrium, correctly anticipate which track she would be placed in at each school.[12] We formalise all this as follows.

Assumption 1 *Let B be the median ability of the students in a school. If the school practices tracking, students with ability above or equal to B are assigned to one track; students with ability below B are assigned to the other track. If the school does not practice tracking, the allocation of students to classes is random.*

We label the high and low tracks H and L, for students above and below the median respectively.

Assumption 1 admits a different interpretation. Suppose each neighbourhood has two *ex-ante* identical schools, each of size $\frac{1}{4}$. If school A is believed

---

[10]A slightly more general set-up is in Summers and Wolfe's early empirical analysis (1977). They consider two summary statistics of the distribution, the proportion of children with ability below a certain threshold and the proportion of children with ability above a different threshold.

[11]This assumes that the track placement in a school is fully determined by the ability distribution of the children of that school. In practice of course the "absolute" ability of the children would affect the track placement, thus the top track might be large in schools with more able children: as will become apparent, the result of the paper depends only on there being a sufficient number of children who would be in the top track in one district and in the lower track in a different district, where there are more abler children. The evidence from the "Ten Percent" mentioned in the Introduction (Cullen et al 2013, Estevan et al 2012), does indicate that a child's ability relative to her peers affects her track placement.

[12]We could make alternative assumptions; for example that schools choose the class size, and they observe ability with an error. Such assumptions, however, would simply complicate the model without adding any interesting insight, and so blur the mechanism underlying our main results. We conjecture that other plausible model would also generate segregation without tracking and mixing with tracking: for example, let households have the same income, as well as the same preferences, but different education, and suppose parents observe their child ability with an error, and educated parents make smaller errors. Then, if parents are risk averse, educated parents would have a higher willingness to pay for their children education and the results of our paper would *mutatis mutandis* hold in this set-up.

to attract good students, then, because of the peer group effect, all students will apply to that school, and, if schools select students on the basis of ability, then school A will admit all the ablest applicants. This, from the households' viewpoint, is exactly the same as when a single school in the neighbourhood practices tracking: in both cases, only students with above median ability benefit from the higher quality peer group. Our paper therefore is also relevant to the policy debate, vigorous in many countries, on whether schools should be allowed to select the students they admit.[13]

2.3 The housing market

Households have a simple decision: they choose in which of the two districts to reside. Each household needs a unit of housing. To examine the emergence of segregation, we require that the rich could potentially segregate from the poor, and so assume that housing is available in unlimited supply in district 0, and in fixed supply $\alpha - \varepsilon$, in district 1. District 1, that is, has just fewer houses than there are rich households.[14] This is a technical assumption, ensuring local uniqueness of the equilibrium: suppliers of housing are on the short side in district 1, and this "closes" the model with the determination in their favour of the rental price in district 1. We therefore think of ε as being small, and consider equilibria in the limit as ε tends to 0, though, to lighten notation, we will not declare this formally in the statement of our results: Appendix 2 sketches how the equilibrium price and average peer quality vary with ε. The price of housing in district 0 is given by building costs, which we normalise to 0 to dispense with a redundant parameter. The house price in district 1 is endogenously determined in such a way that the market clears and is denoted by $p \in \mathbb{R}$. The property price can be positive or negative; a negative price simply means a rental price *below* the (flow cost of the) building cost in district 0: given that building costs are sunk, a negative price is not inconsistent with equilibrium.[15] Without further loss of generality, we assume that $\alpha = \frac{1}{2}$: this

---

[13]Recent analyses of school selection from outside the US are Allen (2006) Clark (2010) for the UK and Jackson (2010) for Trinidad and Tobago.

[14]This introduces asymmetry between the districts. An alternative way of doing so would be to assume that district 1 has a desirable amenity (Brueckner et al 1999).

[15]We do not specify how the price is arrived at; it can be reached, for example, if each household bids for a house in district 1, and the highest $(\alpha - \varepsilon)$ bidders all pay the price bid

simplifies the algebra, while maintaining the important assumption that the number of rich households is greater than the housing stock in one district.

2.4 The game

To recap: we study a static game, where the players are the households, and their strategy set is $\{0, 1\}$, the district where they reside. In equilibrium, households correctly anticipate the choices of all other households, and the effect of all these choices on all the payoff relevant variables, the track ability thresholds, the consequent quality of the track their child will be assigned to, and the relative price of housing. Because they are infinitesimally small, households ignore the effect of their own actions on the equilibrium variables, and take prices and school qualities as given. They also take as given the school allocation policy, that is whether or not tracking is practiced. The policy is determined outside the model, for example imposed by government or by a local school board.

Formally, we study and compare two separate games, one where schools track students and one where they do not. We look for Nash equilibria of these games: each household optimises given the choices made by the other households. We allow households to use mixed strategies; in a game with a continuum of players, this is equivalent to requiring that a proportion of players who are indifferent choose one or other of the pure strategies available.

3 Results

3.1 Preliminaries and definitions

In this, like in many other models of its type (e.g. Bénabou 1996a), there exist symmetric equilibria where the households, rich and poor, are distributed across districts in such a way that quality is the same, and the price of housing is 0 in both districts. One such equilibrium arises if each household randomises by choosing to reside in district 1 with probability $\frac{1}{2} - \varepsilon$. This ensures that the two districts end up with exactly the same distribution of ability, and therefore have the same school quality; houses cost the same, so all households receive the

---

by the $(\alpha - \varepsilon)$-th highest bidder. The other households buy a house in district 0 at price 0. Or it could be calculated by a traditional Walrasian auctioneer.

same utility in each district, and no one has an incentive to deviate.[16] Other equilibria, qualitatively similar, are obtained by varying the randomisation, while preserving equality of the mean ability in the two districts. Besides being uninteresting, these symmetric equilibria are extremely fragile: they would be disrupted by a coordinated deviation of a small group (of strictly positive measure) of households whose average ability is higher than the mean. Hereafter we disregard them, and concentrate instead on the asymmetric ones, which instead are robust to these coordinated deviations by a small group of players.

In the rest of this subsection, we collect a number of definitions and variables, useful in the rest of the paper. We begin with the following functions:

$$\Delta_i (p) = v (y_i) - v (y_i - p), \qquad\qquad i = R, P.$$

Intuitively, $\Delta_i (p)$ is the utility cost of paying p for housing instead of paying 0. According to the following Lemma, this is lower for high income households: this is an immediate consequence of their lower marginal utility of income.

Lemma 1 (i) $\Delta_i (p) \gtreqless 0$ according to $p \gtreqless 0$ and (ii) $\Delta_P (p) \gtreqless \Delta_R (p)$ according to $p \gtreqless 0$.

The proofs of all the results are collected in the technical Appendix at the end of the paper. Next define the average abilities of rich and poor households, θ_R and θ_P:

$$\theta_i = 2 \int_{\underline{b}}^{\bar{b}} b\phi_i (b) \, \mathrm{d}b, \qquad\qquad i = R, P,$$

and the average ability in the population:

$$\theta_{av} = \int_{\underline{b}}^{\bar{b}} b \left(\phi_P (b) + \phi_R (b)\right) \mathrm{d}b = \frac{\theta_P + \theta_R}{2}.$$

In the next bit of notation, we divide the entire population into four equally sized groups, as follows. In the first two groups are the households whose children are in the top quarter and in the next quarter of the *overall* ability distribution. Formally, let B^3 be the lower boundary of the highest quartile of
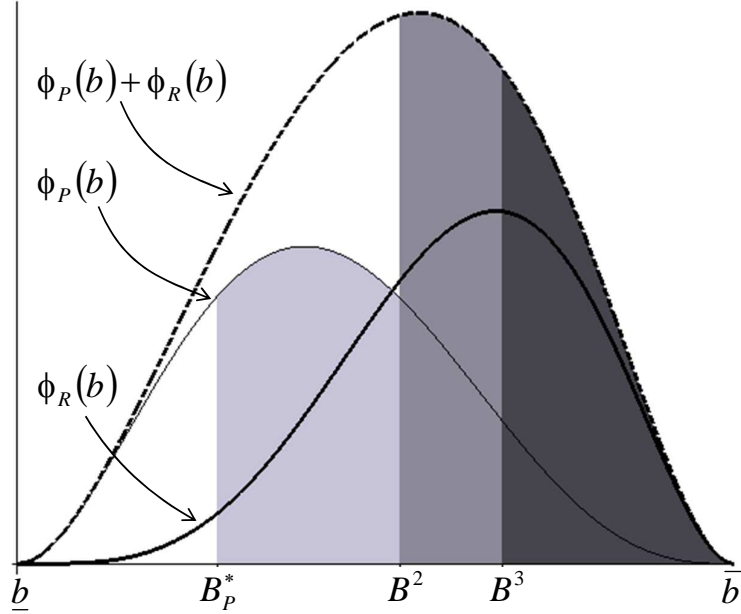

Figure 1: The distribution of abilities in rich (thick line) and poor (thin line) households, and the aggregate distribution (dashed line).

the *aggregate* ability distribution, and B^2 be the median ability in the population: that is, B^3 and B^2 solve:

$$\Phi_P\left(B^3\right) + \Phi_R\left(B^3\right) = \frac{3}{4},$$

$$\Phi_P\left(B^2\right) + \Phi_R\left(B^2\right) = \frac{1}{2}.$$

Figure 1 shows these groups of households: it depicts the densities of the ability distribution in rich (thick line) and poor (thin line) households, and the aggregate density, the dashed line. The darkest area (from B^3 to $\bar{b}$, with measure $\frac{1}{4}$) are the households with the brightest children and the second darkest area (which also has measure $\frac{1}{4}$), the households whose children have ability in the third quartile, that is, above B^2 and below B^3.

The average ability of children in these two groups is given by

$$\theta^3 = 4 \int_{B^3}^{\bar{b}} b\left(\phi_P\left(b\right) + \phi_R\left(b\right)\right) \mathrm{d}b, \tag{2}$$

$$\theta^2 = 4 \int_{B^2}^{B^3} b\left(\phi_P\left(b\right) + \phi_R\left(b\right)\right) \mathrm{d}b. \tag{3}$$

11

The two groups below the median are not the analogous of those above the median. Begin by defining the average ability of the children whose ability is below the median:

$$\theta_m = 2 \int_{\underline{b}}^{B^2} b \left(\phi_P \left(b \right) + \phi_R \left(b \right) \right) \mathrm{d}b.$$

Next, define B_P^* by:

$$\frac{1}{4} = \int_{B_P^*}^{B^2} \phi_P \left(b \right) \mathrm{d}b.$$

In words, B_P^*, which equals to $\Phi_P^{-1} \left(\Phi_P \left(B^2 \right) - \frac{1}{4} \right)$, is the level of ability such that there are exactly $\frac{1}{4}$ poor children between B_P^* and the population median. Their average ability is

$$\theta_P^* = 4 \int_{B_P^*}^{B^2} b \phi_P \left(b \right) \mathrm{d}b.$$

The lightest grey area in Figure 1 shows this group of households. The remaining households, the poor ones with very low ability children and the rich ones with children with ability below the median, form the white area below the dashed curve, which has measure $\frac{1}{4}$.

We end this subsection by defining formally the two types of equilibria that constitute the focus of our paper.

Definition 1 *A* full segregation equilibrium *is a Nash equilibrium of the location choice game such that all rich households live in district 1, and all poor households live in district 0.*

Definition 2 *For given B^2 and B^3, a* desegregation equilibrium *is a Nash equilibrium of the location choice game, such that the allocation of households to districts and of children to tracks is given by*

ability interval	rich households	poor households
$\left(B^3, \bar{b} \right]$	*district 1, high track*	*district 1, high track*
$\left(B^2, B^3 \right]$	*district 0, high track*	*district 0, high track*
$\left[\underline{b}, B^2 \right]$	*all in the same district, low track*	*some in district 1, low track some in district 0, low track*

12

In a desegregation equilibrium, both districts are home to a positive measure of both poor and rich households. We shall show that, both with positive and with negative prices, all rich households with below-median ability children live in the more expensive district, where the quality of the low track is higher. Their children are in the low track in either district, but in the more expensive one, be it 0 or 1, the peer group is higher: this is necessary to ensure that the poor households with low ability children are indifferent. As we mentioned above, a negative price has the natural interpretation of house prices in the district with fixed supply falling below the construction cost of new homes in the district where new construction is possible.

This notation introduced, in the next three subsections we study first the game where schools do not track their students, and subsequently determine the conditions that must hold for the two possible equilibria (mixed and fully segregated) to occur in the game where schools track their students.

3.2 Equilibrium without tracking

Let us suppose that schools do not track students. This corresponds to the model of neighbourhood schooling in Epple and Romano (2003).

Proposition 1 *Let schools assign students to classes randomly. The unique non-symmetric equilibrium is a full segregation equilibrium. The price of housing in district* 1 *is given by*

$$p = \Delta_R^{-1} \left(\theta_R - \theta_P \right) > 0. \tag{4}$$

Intuitively, Proposition 1 holds because, since wealthier households are more willing to pay for school quality, and since they cannot pay for school quality directly, they pay for it through the purchase of a good that is complementary in consumption to school quality, namely housing. Houses in district 1 become sufficiently expensive, pricing lower income households out of the market. Complete income segregation ensues. This result would hold even if some of the fairly specific assumptions under which it is obtained were relaxed; for example, with less than perfect complementarity between school quality in a district and housing in that district. It would hold a fortiori if wealthier households valued

school quality more.[17]

As we show next, tracking may prevent the operation of this mechanism, the reason being that, with tracking, ability to pay is insufficient to ensure the benefit of a better peer group.

3.3 Equilibrium with tracking: (i) desegregation

We begin with the desegregation equilibrium, where each district is inhabited by both poor and rich households. The main results of the paper are contained in Propositions 2 and 3. Proposition 2 provides necessary and sufficient conditions on the ability distribution for existence of a desegregation equilibrium with positive housing price. Proposition 3, the "negative price" counterpart of Proposition 2, establishes that, for any income and ability distribution, there exists of a continuum of desegregation equilibria with negative price.

Proposition 2 *A desegregation equilibrium with positive price exists if and only if*

$$\frac{\theta^3 + \theta_P^*}{2} \geq \theta_{av}. \tag{5}$$

The qualitative feature of this equilibrium is that in district 1 one finds all the households with very high ability children, all the rich households with below median ability children, and some of the poor households with below median ability children.

This is illustrated in Figure 2 . The ability distributions are the same as in Figure 1, but here the ability density of rich households is drawn only as stacked above that of poor households. Households with children with ability in the top quartile, the two darkest areas, all live in district 1, pay the high rental price, and their children study in the top track. The rest of the population of district 1 are the rich households where children have ability below the population median, B^2 – of which there are fewer than $\frac{1}{4}$, the white area below the dashed line –, and, to make up the numbers, some poor households where the children have below median ability. These households, the lightest shade of grey, are indifferent between the two districts, and, as Figure 2 shows, they live in both

---

[17]And it does not depend on there being only two income levels either: with a generic income distribution, there would be a cut-off level of income, with households with higher income than that concentrating in district 1 (Epple and Romano 2003).
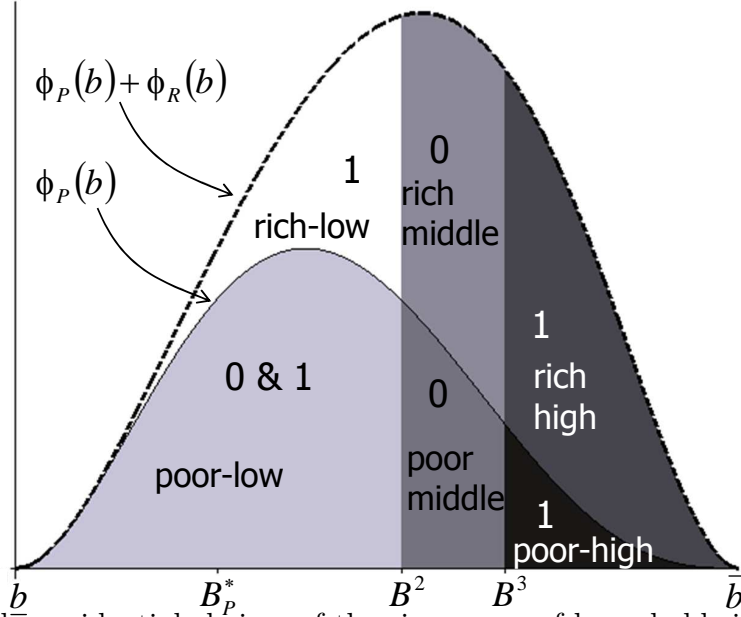
Figure 2: The residential choices of the six groups of households in the deseg-regation equilibrium

districts: because of the positive correlation between income and ability, there are more than $\frac{1}{4}$ of them, and their children fill the low track in the school in district 0. The rest of the households are those in the third quartile of the ability distribution. They are drawn as the intermediate shades of grey, all live in district 0, and are assigned to the top track there.

To illustrate the intuition behind Proposition 2 note that in equilibrium no household must have an incentive to move, and so the equilibrium property price must ensure that poor households with children of below median ability are indifferent between the districts, which they must be, since they are dis-tributed across both. Since they are indifferent, they can randomise, and to each possible randomisation (that is each randomisation which ensures that $\frac{1}{4}$ of the poor households reside in district 0) there corresponds a different qual-ity for the low track in the two districts, and hence a different property price. Thus, the equilibrium is not unique: instead there is a continuum of equilibria with these qualitative features. Since the low track must have higher quality in district 1 (that is, $\theta_1^\mu - \theta_0^\mu \geq 0$), Lemma 1 implies that rich household with children of below median ability prefer district 1 strictly and so do not want to move in equilibrium. The households with high ability children do not want to

move either: the poor prefer to stay put in district 1 if they are willing to pay for a house more than their low ability counterparts, that is, if the quality gap is greater between the high tracks than between the low tracks. That will be the case if

$$\theta_3 - \theta_2 \geq \theta_1^\mu - \theta_0^\mu, \tag{6}$$

a condition which, along with Lemma 1, implies that rich households with high ability children also prefer to stay put. Finally, consider the rich and poor households with children of ability between B^2 and B^3. They live in district 0 where they are assigned to the high track. In district 1, in turn, they would be in the low track *and* pay a positive housing price. Given that the quality of the low track in district 1, θ_1^μ, is always below θ_2, these households strictly prefer to live in district 0.

Existence of this equilibrium therefore depends on the existence of a randomisation function for which (6) holds, which is precisely what condition (5) guarantees: the latter ensures that with the randomisation determining the lowest possible positive value of $\theta_1^\mu - \theta_0^\mu$, condition (6) holds and so that no household has an incentive to move.

To describe the equilibrium more formally and gain a better understanding of condition (5), suppose poor households whose children have ability $b \in [\underline{b}, B^2]$ randomise and go to district 0 with probability $\mu(b) \in [0,1]$ satisfying

$$\int_{\underline{b}}^{B^2} \mu(b)\, \phi_P(b)\, \mathrm{d}b = \frac{1}{4}. \tag{7}$$

Notice that, since positive income ability correlation implies that there are more than $\frac{1}{4}$ poor households below the population median B^2, $\mu(b)$ is well defined. Given (7), the rent premium that ensures that poor households with low ability children are indifferent between the two districts is:

$$p = \Delta_P^{-1} \left(\theta_1^\mu - \theta_0^\mu \right), \tag{8}$$

with

$$\theta_0^\mu = 4 \int_{\underline{b}}^{B^2} b\mu(b)\, \phi_P(b)\, \mathrm{d}b, \tag{9}$$

$$\theta_1^\mu = 4 \int_{\underline{b}}^{B^2} b\left((1 - \mu(b))\, \phi_P(b) + \phi_R(b) \right) \mathrm{d}b. \tag{10}$$

16

provided $\theta_1^\mu \geq \theta_0^\mu$. Clearly, as $\mu(b)$ changes, so do p and the low track qualities in the two neighbourhoods.

According to Proposition 2, a desegregation equilibrium with positive house price exists if condition (5) holds; this condition requires the school quality differential between schools 1 and 0 be larger for students of high ability, who go in the high track, than for students of low ability, assigned to the low track. Note that this is a mild condition: every child in the high track of school 1 has ability above all the children in the high track of school 0. On the contrary, children in the low tracks of the two districts have abilities between $\underline{b}$ and B^2. That is, children in the high tracks are fully segregated by ability, those in the low tracks are mixed.

Put differently, condition (5) requires that children in two groups have, on average, an ability exceeding the average in the whole population. These two groups are the children in the top quartile, from both poor and rich households, and the ablest $\frac{1}{4}$ of the below-the median children from poor households. That this is a mild condition can be gleaned by inspecting Figure 2. The former group, the ablest $\frac{1}{4}$ children, whose ability is above B^3, are in the darkest areas in the figures, and have average ability equal to θ^3. The latter group of children are the ablest below median from poor households, the lightest grey areas in Figure 2; their average ability is θ_P^*. For (5) to hold, it must be that the average ability of the children in these two groups exceeds the overall average ability. In order for (5) to be violated, there must be sufficiently high correlation between income and ability. To see this, note that, with no correlation, the white area vanishes, and so (5) necessarily holds. By continuity, this is also true for sufficiently small correlation.

The multiplicity of mixed strategy equilibria is a direct consequence of the assumption of the separability between a student's ability and her school's quality, measured by peer ability. This ensures that households benefit equally from school quality, and in particular that *all* the poor households with low ability children are indifferent between the districts, and so multiple randomisations are possible. To every randomisation there corresponds a different school quality and consequently a different price ensures that the household who randomise are indeed indifferent.[18]

---

[18]This multiplicity would not occur in a model where ability and school quality are complements, that is one where brighter children benefit more from a high quality peer group.

The next result illustrates that in the desegregation equilibrium there is genuine mixing, not just a few token households moving to the "other" district.

Corollary 1 *In a desegregation equilibrium with positive price, $\left(\frac{1}{4} - \left(\Phi_P \left(B^3 \right) - \Phi_P \left(B^2 \right) \right) \right)$ poor (rich) households live in district 1 (district 0).*

To gain a sense of the extent of the desegregation, consider the following example: let ability be normalised in $[0, 1]$ and distributed in the two population groups according to

$$\phi_i \left(b \right) = \frac{b^{\nu_{1i}} \left(1 - b \right)^{\nu_{2i}}}{2\beta \left(\nu_{1i}, \nu_{2i} \right)}, \qquad\qquad i = R, P, \qquad (11)$$

with ν_{1i} and ν_{2i} taking values $(5, 3)$ for the rich and $(3, 4)$ for the poor (these are the densities depicted in Figures 1 and 2). In this case, in district 1 there live 0.346 rich households and 0.154 poor ones, a ratio of 2.25 (and obviously the reverse in district 0). While, clearly, tracking stratifies children by ability in the classroom, it does not necessarily do so geographically. Indeed, in the numerical example considered here, the average ability in the districts is 0.636 and 0.455 without tracking and 0.563 and 0.528 in the equilibrium with tracking where district 1 has the highest possible average ability:[19] so tracking reduces the gap in average ability in the two districts. The composition of the district also changes in a non-monotonic way: for example, the standard deviation of ability is 0.139 and 0.144 in district 1 and 0 respectively without tracking, and 0.217 and 0.094 with tracking. This is not a general result, and different patterns can easily be found.

We argued above that there is no reason why house prices in district 1 (where new houses cannot be built) could not fall below the marginal cost,

---

If the effect is small enough, it does not change the preferences of the other groups of households, but it breaks the indifference of the poor households with low ability children: all the households below a given cut-off strictly prefer district 0, as they benefit less from school quality, and the households above the cut-off strictly prefer district 1, where school quality is higher. The cut-off is determined to clear the housing market, and the housing price to ensure that the households with children at the ability cut-off are indifferent between the districts (the argument for substitutability is analogous).

[19]That is, we refer to the equilibrium with the largest ability difference between the two low tracks and thus with the greatest housing price within the set of desegregation equilibria with positive price.

normalised to 0, of building new houses in district 0. The next proposition establishes that a continuum of desegregation equilibria with negative prices exists for any parameter configuration. These equilibria are very similar to those identified in Proposition 2, the difference being that all rich households with below-median ability children now live in district 0, where houses are more expensive.

Proposition 3 *There exists a continuum of desegregation equilibria with negative property prices.*

The qualitative difference between an equilibrium with positive property price, which according to Proposition 2 exists only if (5) holds and one with negative prices, which Proposition 3 shows always to exist, is the location of the rich households with low ability children: they are in district 1 when the price is positive, and in district 0 when it is negative. In both cases, poor households with low ability children are indifferent between districts and the relation between price and school quality must be such that they are indifferent, while at the same time making sure that the poor households with high ability children are happy to stay in district 0, and send their children to the high track.

3.4   Equilibrium with tracking: (ii) full segregation

To continue with the analysis of the equilibrium set, the proposition in this subsection provides conditions such that there exists a tracking equilibrium displaying perfect income segregation across districts. The condition requires a very high correlation of ability and income, and substantial income inequality: that is, full segregation can happen even when the schools practice tracking, but for this to occur rich and poor households must be very different.

We need some additional notation. Let B_2^i be the median ability of the households with income y_i (i.e. $B_2^i = \Phi_i^{-1}\left(\frac{1}{4}\right)$); note that because of the correlation between income and ability $B_2^P < B_2^R$. Let θ_H^i and θ_L^i be the average ability of children in the top and in the bottom half of the distribution of ability

for each income group i:

$$\theta_H^i = 4 \int_{B_2^i}^{\bar{b}} b\phi_i(b)\, \mathrm{d}b, \qquad\qquad i = R, P;$$

$$\theta_L^i = 4 \int_{\underline{b}}^{B_2^i} b\phi_i(b)\, \mathrm{d}b, \qquad\qquad i = R, P.$$

Lemma 2 $\theta_j^R > \theta_j^P$ $j = H, L$.

The Lemma is a consequence of the correlation between income and ability.

Proposition 4 *If the schools practice tracking, a full segregation equilibrium exists if and only if*

$$\Delta_P\left(\Delta_R^{-1}\left(\theta_L^R - \theta_H^P\right)\right) \geqslant \max\left\{\left(\theta_H^R - \theta_H^P\right), \left(\theta_L^R - \theta_L^P\right)\right\}. \qquad (12)$$

In this equilibrium, the housing price is

$$p = \Delta_R^{-1}(\theta_L^R - \theta_H^P). \qquad (13)$$

Condition (12) has a natural interpretation. Notice first that for it to hold, its LHS must be strictly positive, that is, θ_L^R must exceed θ_H^P. In words, the average ability of the *low ability half of the children from rich households*, θ_L^R, must exceed the average ability of the *high ability half of the children from poor households*, θ_H^P. This requires the correlation between income and ability to be very high. Moreover, (12) is certainly violated when $\Delta_P = \Delta_R$, that is when income levels are equal, irrespective of the ability income correlation. By continuity, this is also the case when the income levels are similar.

3.5 Discussion

Our formal analysis in the previous subsections can be summarised by saying that tracking can lead to desegregation. Desegregation has manifold advantages, consequences of exposing children from disadvantaged backgrounds to the life-style and ambitions of children from different, more privileged backgrounds; this, however, is not the place to discuss the broader implications of segregation.[20]

---

[20]Segregated cities generate significant differences in the educational inputs available to children of different socio-economic backgrounds (through peer group effects at the classroom,

Opponents of tracking (e.g. Oakes, 2005) argue that tracking is detrimental to social mobility, because, with correlation between income and ability, disadvantaged children tend to be overrepresented in low tracks, where tracking reduces the quality of the peer group: to the extent that high quality education, more available to those in the high track, is a gateway to university, disadvantaged children will be correspondingly underrepresented in higher academic institutions. Countervailing this, however, in a desegregation equilibrium, children in poor households with above median ability are educated in schools with a higher quality peer group than they would if schools did not practice tracking. Thus they not only have classmates and neighbourhood friends from a better social background, but they also receive better schooling: both these factors should increase their likelihood to attend university.[21]

From a different viewpoint, the question is whether, for given correlation between income and innate or pre-school *ability*, tracking increases or decreases correlation between income and *school achievement*: does tracking dampen or heighten social inequality? In the set up of our model, school achievement depends on school quality and ability. In a full segregation equilibrium without tracking, privileged children receive higher quality education than their poor counterparts, due to the peer effect. In a desegregation equilibrium, on the other hand, rich and poor youths of equal ability who are assigned to a high track enjoy the same school quality. Correlation between income and school achievement is therefore lower in the latter equilibrium. To the extent that admission to university depends on school achievement, and that "better" universities require higher achievement, then disadvantaged young people will have a better chance to be admitted to a high quality university in the desegregation equilibrium. This must be balanced with the observation that, as can be seen in Figure 2, the gap in average ability in the four classes (two in each district) is increased by tracking: brighter students receive better education

_____

neighbourhood effects, role and behavioural models and so on). As a result, they tend to polarise educational opportunities and to perpetuate income and human capital inequality (Bénabou 1996a, 1996b). Furthermore, ghettos and poverty traps emerge naturally (Bénabou 1993; Durlauf 1996). Note however, as shown for example by Cutler et al (2008), that segregation is not always associated to negative outcomes.

[21]Indeed, as we noted in the introduction, the empirical evidence on the impact of tracking on social mobility is mixed (Lee 2011a, 2011b; Brunello and Checchi 2007; Pekkarinen et al 2009).

with tracking, weaker students worse education. And since there are relatively more brighter students among the wealthier households, the latter, as a group, appear to benefit more than poorer households from a change to tracking.

3.6 Towards a welfare analysis

The previous section notes the trade-off between providing some children from poor households with the best education, but leaving a majority of them in the low tracks at school. To evaluate this trade-off a welfare analysis is needed; while a complete theoretical study cannot be properly conducted in this highly simplified set-up, some formal consideration can be carried out in conclusion to the paper. We begin by giving conditions ensuring that all households are better-off with tracking than without. These equilibria will necessarily have negative property prices, since the welfare gains of poor households with low ability children are realised at the expense of landlords, and therefore we do not speak of a Pareto improvement.

Define $\overline{\theta}_0^\lambda$ as the maximum possible quality of the low track in district 0 in a desegregated equilibrium with negative property prices. This is given by:

$$\overline{\theta}_0^\lambda = 4 \left[\int_{\underline{b}}^{B^2} b\phi_R(b)\,\mathrm{d}b + \int_{\Phi_P^{-1}\left(\frac{1}{4}\right)}^{B^2} b\phi_P(b)\,\mathrm{d}b \right].$$

This maximum is attained when the brightest of the poor kids with below-the-median ability attend the low track in district 0, along with the rich kids with below-the-median ability. The other poor children, those whose ability is below the median of the poor children, all live in district 1.

The next proposition gives a sufficient condition for all households to be better off in a desegregated equilibria (with negative prices) than in the full segregation equilibrium that results in the absence of tracking.

Proposition 5 *Let*

$$\overline{\theta}_0^\lambda > \theta_P; \tag{14}$$

then there are desegregation equilibria with negative house prices which every household prefers to the equilibrium where schools do not track students.

To interpret condition (14), consider again Figure **??**: it requires that the average ability of those in the white area, plus some (the appropriate number to

22

have $\frac{1}{4}$ students in the low track in district 0) of those in the lightest grey area be higher than the average ability of the poor, those below the thin solid curve. The easiest way to satisfy the condition is if the poor, low ability students who go in district 0 are the best, that is the vertical slice at the right of the lightest grey area: this "slice" must be sufficiently thick, and so condition (14) requires positive correlation between income and ability. In this case, the average ability in the low track in district 0 is $\overline{\theta}_0^{\lambda}$. When condition (14) holds, everybody is better off with tracking and the consequent desegregation in equilibrium: the lowest ability poor households benefit because, though they have lower school quality, they also pay a below-cost price for housing. The low ability rich also have lower quality schooling, but do not pay a positive price for housing, since they live in district 0. The higher ability poor, either live in district 1 and so have negative house price *and* higher school quality than they would in the no-tracking equilibrium, or, if they live in district 0, they have better school quality, since they do not have the least able poor children as peers, but some middle quality rich children.[22]

We can say something more specific for households with children with above median ability.

Corollary 2 *Let $\theta_2 > \theta_P$: then all households above the median ability strictly prefer any desegregation equilibrium to the no tracking equilibrium.*

The condition that ensures that above median households are made better off by tracking is mild. It requires the average ability in the third quartile of the population, which is the quality of the high track in the poor district, to be greater than the average ability of children from poor households. The condition that $\theta_2 > \theta_P$ is sufficient *and* necessary for households with intermediate ability and, therefore, they will prefer no tracking if it does not hold. In that case, moreover, $\theta_3 - \theta_P < \theta_3 - \theta_2$ so that an equilibrium may exist in which

---

[22]With the ability distributions given in (11) in the example following Corollary 1, assuming that the utility of consumption is $v(x) = \ln x$, and $y_R = 20$ and $y_P = 5$, in the tracking equilibrium the price for housing in district 1 would be -0.871, the three groups of rich households would have utility 3.812, 3.599 and 3.441, with the poor households having utility 2.544, 2.213 and 2.055. In the no tracking equilibrium, price is 3.567, and the household utility is 3.424 for the rich and 2.038 for the poor. Details of the calculations are available on requests.

$\theta_3 - \theta_2 > \theta_1^\mu - \theta_0^\mu > \theta_3 - \theta_P$. In this case, the poor households with high ability children will enjoy lower utility with tracking and so will their rich counterparts if the income gap is small enough. Thus, another important implication of our analysis is that no group of households will consistently prefer tracking across all parameter configurations.

Together with the multiplicity of equilibria,[23] at least for some parameter combinations, this illustrates that the ambitious task of the welfare analysis of tracking needs both a more nuanced model and a solid empirical underpinning. The aim of the paper, however, is to illustrate the complex interaction between school policies, the household's residential choices, and the social and demographic environment, and to underline that this interaction cannot be ignored in any meaningful analysis of educational policies.

In this vein, note how the relative price of property in district 1 play a different role with and without tracking: when schools cannot track their students, it must be sufficiently *high* in equilibrium to deter an ε mass of rich households from choosing to move to district 1, where schools are better. When schools can track their students, it must be sufficiently *low* in equilibrium so as to induce some poor households to move to district 1 (see Figure 2). However, this argument cannot be used to infer that housing price is lower with tracking. A simple numerical example shows that the opposite can indeed happen: let ability be again distributed in the two population groups according to (11), and let the consumption component of utility be given by $\ln(y)$. If the parameter ν_1 and ν_2 and y take values $(4, 4)$ and 11 for the poor and $(4, 4.5)$ and 12 for the rich households, then the equilibrium price is 0.3478 without tracking and can take any value in $(0, 0.3506]$ with tracking, depending on the mixed strategy followed by the households with low ability children, who are indifferent between the two districts. So tracking can in fact increase the equilibrium price. This is a rather extreme example, where both the ability distribution and the income levels are close for poor and rich households. For parameter combinations which make the groups more diverse, tracking does reduce price. For example, if the parameter ν_1 and ν_2 and y are $(7, 4)$ and 20 for the rich households (and the same as before for the poor), then the equilibrium price is 2.549 without tracking and can go only up to 1.394 with tracking.

---

[23]Which appears to be a feature in this kind of models, see de Bartolome and Ross (2004).

4 Concluding remarks

This paper makes a simple point. In an environment where households care about the quality of the education their children receive, schools follow a residence-based admission policy by favouring local residents, and the quality of education depends on the peer group, the tendency towards residential income segregation, that is the tendency of households to cluster according to income, is weakened if schools assign their students to classes according to ability. In short, tracking generates residential social mixing, implying that the well-intentioned movement advocating "de-tracking" schools (Argys et al 1996) may end up causing an increase in income stratification.

While we use a highly stylised model, the principle behind our result has a general validity: in a neighbourhood populated exclusively by households of a "good" socio-economic background, when schools track students, it will necessarily be the case that many parents find that their child is allocated to a low track. Similarly, if the local schools are selective, they may find their children are not admitted to the school of their choice. These parents might consider that their children's education is adversely affected and might wish to consider alternatives. If private schools are too expensive or otherwise unavailable, for example because they accept only high ability children, then they could send their child to a school where the average ability of the children is lower, so that they are assigned to the top track in that school.[24] If there is positive correlation between ability and socio-economic background, then schools in districts where children are from a lower socioeconomic background may indeed assign their child to the top track. Moving to that area would then be an option for these parents. If they, and other similar households, move, then children from different socio-economic backgrounds would attend the same school, interact as neighbours, play together in the local sports teams and so on. Note that, if there is school choice and residents are not given priority at their local school, the social mixing effect of school tracking is dampened but not cancelled, as children from different social backgrounds would still spend the school hours together and have the opportunity to strike friendships among themselves.

---

[24]Private schools may also disrupt perfect income stratification (Martinez-Mora 2006), and in general interact with state schools in a complex manner Epple et al (2002).

Income desegregation is a hitherto unnoticed effect of tracking, and may offset some of the distributional adverse effects of tracking pointed out by the literature, which also emerge in our set up, such as the concentration of the beneficial effects on households with high ability children, be they well-off households, who generally enjoy lower residential prices, or poorer households, whose children receive better education.

References

Allen, R., 2006. Allocating pupils to their nearest secondary school: The consequences for social and ability stratification. Urban Studies 44, 751–770.

Alonso, W. A., 1964. Location and Land Use: Toward a General Theory of Land Rent. Harvard University Press, Cambridge.

Argys, L. M., Rees, D. I., Brewer, D. J., 1996. Detracking America's schools: Equity at zero cost? Journal of Policy Analysis and Management 15, 623–645.

Arnott, R., Rowse, J., 1987. Peer group effects and educational attainment. Journal of Public Economics 32, 287–305.

Astin, A. W., 1993. What Matters in College: Four Critical Years Revisited. Jossey-Bass, Michigan.

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

Betts, J. R., 2011. The economics of tracking in education. In: Hanushek, E. A., Machin, S. J., Wßmann, L. (Eds.), Handbook of the Economics of Education. Volume 3. Elsevier, Amsterdam, pp. 341–381.

Bishop, J., 2006. Drinking from the fountain of knowledge: Student incentive to study and learn – Externalities, information problems and peer pressure. In: Hanushek, E., Welch, F. (Eds.), Handbook of the Economics of Education, Volume 2. Elsevier, Amsterdam, pp. 909–944.

Black, S. E., Machin, S., 2011. Housing valuations of school performance. In: Hanushek, E. A., Machin, S. J., Wßmann, L. (Eds.), Handbook of the Economics of Education. Volume 3. Elsevier, Amsterdam, pp. 485–519.

Bnabou, R., 1993. Workings of a city: Location, education, and production. Quarterly Journal of Economics 108, 619–652.

Bnabou, R., 1996a. Equity and efficiency in human capital investment: The local connection. Review of Economic Studies 63, 237–264.

Bnabou, R., 1996b. Heterogeneity, stratification and growth: Macroeconomic implications of community structure and school finance. American Economic Review 86, 584–609.

Brueckner, J., Zenou, Y., Thisse, J.-F., 1999. Why is central Paris rich and downtown Detroit poor? An amenity-based theory. European Economic Review 43, 91–107.

Brunello, G., Checchi, D., 2007. School tracking and equality of opportunity. Economic Policy 22, 781–861.

Brunello, G., Giannnini, M., Ariga, K., 2007. The optimal timing of school tracking. In: Wßmann, L., Peterson, P. E. (Eds.), Schools and the Equal Opportunity Problem. MIT Press, Cambridge, pp. 129–156.

Calabrese, S., Epple, D., Romano, R., 2012. Inefficiencies from metropolitan political and fiscal decentralization: Failures of Tiebout competition. Review of Economic Studies 79, 1081–1111.

Calabrese, S., Epple, D., Romer, T., Sieg, H., 2006. Local public good provision: Voting, peer effects, and mobility. Journal of Public Economics 90, 959–981.

Calsamiglia, C., Martnez-Mora, F., Miralles, A., 2013. School choice in a tiebout model: Boston mechanism vs deferred acceptance. Tech. rep.

Caucutt, E., 2002. Educational vouchers when there are peer group effects – Size matters. International Economic Review 43, 195–222.

Clapp, J. M., Nanda, A., Ross, S. L., 2008. Which school attributes matter? the influence of school district performance and demographic composition on property values. Journal of Urban Economics 63, 451–466.

Clark, D., 2010. Selective schools and academic achievement. The B.E. Journal of Economic Analysis & Policy (Advances) 10, Article 9.

Coleman, J. S., Campbell, E. Q., Hobson, C., McPartland, J., Mood, A. M., Weinfield, F. D., York, R. L., 1966. Equality of Educational Opportunity. US Department of Health, Education and Welfare, Washington DC.

Cullen, J. B., Long, M. C., Reback, R., 2013. Jockeying for position: Strategic high school choice under Texas' top ten percent plan. Journal of Public Economics 97, 32–48.

Cutler, D. M., Glaeser, E. L., Vigdor, J. L., 2008. When are ghettos bad? lessons from immigrant segregation in the United States. Journal of Urban Economics 63, 759–774.

de Bartolome, C. A. M., 1990. Equilibrium and inefficiency in a community model with peer group effects. Journal of Political Economy 98, 110–133.

de Bartolome, C. A. M., Ross, S. L., 2003. Equilibria with local governments and commuting: Income sorting vs income mixing. Journal of Urban Economics 54, 1–20.

de Bartolome, C. A. M., 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, 458–483.

de Bartolome, C. A. M., Ross, S. L., 2007. Community income distributions in a metropolitan area. Journal of Urban Economics 61, 496–518.

Dhar, P., Ross, S. L., 2012. School district quality and property values: Examining differences along school district boundaries. Journal of Urban Economics 71, 18–25.

Downes, T. A., Zabel, J. E., 2002. The impact of school characteristics on house prices: Chicago 1987-1991. Journal of Urban Economics 52, 1–25.

Durlauf, S. N., 1996. A theory of persistent income inequality. Journal of Economic Growth 1, 75–93.

Eisenkopf, G., 2009. Student selection and incentives. Zeitschrift fr Betriebswirtschaft 79, 563–577.

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

Epple, D., Newlon, E., Romano, R. E., 2002. Ability tracking, school competition, and the distribution of educational benefits. Journal of Public Economics 83, 1–48.

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

Epple, D., Romano, R., 2011. Peer effects in education: A survey of the theory and evidence. In: Benhabib, J., Bisin, A., Jackson, M. O. (Eds.), Handbook of Social Economics, Volume 1B. Elsevier, Amsterdam, pp. 1053–1163.

Epple, D., Romano, R. E., 1998. Competition between private and public schools, vouchers and peer group effects. American Economic Review 88, 33–62.

Epple, D., Romano, R. E., 2003. Public school choice and finance policies, neighborhood formation, and the distribution of educational benefits. In: Hoxby, C. (Ed.), The Economics of School Choice. University of Chicago Press, Chicago, pp. 227–286.

Estevan, F., Gall, T., Legros, P., Newman, A. F., 2012. College admission and high school integration. mimeo.

Fernandez, R., Rogerson, R., 1998. Public education and income distribution: A dynamic quantitative evaluation of education-finance reform. American Economic Review 88, 813–33.

Gibbons, S., Machin, S., 2003. Valuing English primary schools. Journal of Urban Economics 53, 197–219.

Hanushek, E. A., Wßmann, L., 2006. Does educational tracking affect performance and inequality? Differences-in-differences evidence across countries. Economic Journal 116, C63–C76.

Hanushek, E. A., Yilmaz, K., 2010. Household location and schools in metropolitan areas with heterogeneous suburbs; Tiebout, Alonso, and government policy. Tech. Rep. 15915, NBER, Cambridge, MA.

Hidalgo, M., 2010. Tracking can be more equitable than mixing: Peer effects and college attendance. Tech. Rep. 2010-162, Society for the Study of Economic Inequality, Palma de Mallorca, Spain.

Hidalgo, M., 2011. On the optimal allocation of students when peer effects work: Tracking vs mixing. Journal of the Spanish Economic Association 2, 31–52.

Jackson, C. K., 2010. Do students benefit from attending better schools?: Evidence from rule-based student assignments in Trinidad and Tobago. Economic Journal 120, 1399–1429.

Lee, Y. S., 2011a. Educational tracking, residential sorting, and intergenerational economic mobility. Brown University.

Lee, Y. S., 2011b. Exam based tracking versus district based mixing: Implications on households sorting and intergrnerational education mobility. Brown University.

LeRoy, S. F., Sonstelie, J., 1983. Paradise lost and regained: Transportation innovation, income and residential location. Journal of Urban Economics 13, 67–89.

Loury, G. C., 1981. Intergenerational transfers and the distribution of earnings. Econometrica 49, 843–867.

Lucas, S. R., 1999. Tracking Inequality: Stratification and Mobility in American High Schools. Teachers College Press, New York.

Martinez-Mora, F., 2006. The existence of non-elite private schools. Journal of Public Economics 90, 1505–1518.

Nechyba, T. J., 1999. School finance induced migration and stratification patterns: The impact of private school vouchers. Journal of Public Economic Theory 1, 5–50.

Nechyba, T. J., 2003. School finance, spatial income segregation, and the nature of communities. Journal of Urban Economics 54, 61–88.

Oakes, J., 2005. Keeping Track: How Schools Structure Inequality, Second Edition. Yale University Press, New Haven.

Pack, H., Pack, J. R., 1977. Metropolitan fragmentation and suburban homogeneity. Urban Studies 14, 191–201.

Pekkarinen, T., Kerr, S., Uusitalo, R., 2009. School tracking and intergenerational income mobility: Evidence from the Finnish comprehensive school reform. Journal of Public Economics 93, 956–973.

Perl, L. J., 1973. Family background, secondary school expenditure and student ability. Journal of Human Resources 8, 156–180.

Persky, J., 1990. Suburban income inequality: Three theories and a few facts. Regional Science and Urban Economics 20, 125–137.

Roemer, J.E., 1998. Equality of Opportunity. Harvard University Press, Cambridge.

Sacerdote, B., 2011. Peer effects in education: How might they work, how big are they and how much do we know thus far? In: Hanushek, E. A., Machin, S. J., Wßmann, L. (Eds.), Handbook of the Economics of Education. Volume 3. Elsevier, Amsterdam, pp. 249–277.

Sirin, S. R., 2005. Socioeconomic status and academic achievement: A meta-analytic review of research. Review of Educational Research 75, 417–453.

Summers, A. A., Wolfe, B. L., 1977. Do schools make a difference? American Economic Review 67, 639–652.

Takii, K., Tanaka, R., 2009. Does the diversity of human capital increase GDP? A comparison of education systems. Journal of Public Economics 93, 998–1007.

Tiebout, C., 1956. A pure theory of public expenditure. Journal of Political Economy 64, 416–424.

Wheelock, A., 1992. Crossing the Tracks: How "Untracking" Can Save America's Schools. The New Press, New York.

White, K. R., 1982. The relation between socioeconomic status and academic achievement. Psychological Bulletin 91, 461–481.

Winkler, D. R., 1975. Educational achievement and school peer group composition. Journal of Human Resources 10, 189–204.

Appendix 1

Since all the arguments are based on utility differentials between districts, the additive nature of the utility function implies that there is no loss in generality in normalising away to 0 the term $q(b)$.

Proof of Lemma 1. (i) follows from $\Delta_i(0) = 0$ and $v'(y_i) > 0$. Consider (ii) next. Write $\Delta_i(p)$ as

$$\Delta_i(p) = \int_0^p v'(y_i - x)\, dx$$

and so

$$\Delta_P(p) - \Delta_R(p) = \int_0^p \left(v'(y_P - x) - v'(y_R - x) \right) dx = \int_0^p \int_{y_R}^{y_P} v''(y - x)\, dy dx$$

Since $y_P < y_R$ and $v''(y) < 0$ the sign of the above is the same as the sign of p, and the Lemma is established. ∎

Proof of Proposition 1. Consider the full segregation allocation of households to districts, with price given in (4). We show that this is an equilibrium. The rich households are indifferent between the two districts, and so a measure ε of them can settle in district 0. A poor household's utility gain from moving from district 0 to district 1 is

$$(v(y_P - p) + \theta_R) - (v(y_P) + \theta_P),$$

which can be written as

$$(\theta_R - \theta_P) - (v(y_P) - v(y_P - p)) = \Delta_R(p) - \Delta_P(p).$$

By Lemma (1), the above is negative, and so no poor household wishes to deviate. Next consider uniqueness. Clearly there cannot be a full segregation equilibrium with a price different from (4): a lower price would ensure that the ε rich households left in district 0 would want to deviate and move to district 1. A higher price would induce all rich households to move to district 0. Similarly, there cannot be another equilibrium apart from the trivial one with $p = 0$ and identical quality. If $p > 0$, then, by Lemma (1), it cannot be that both rich and poor households are indifferent: if the rich are indifferent the poor strictly prefer district 0. ∎

Proof of Proposition 2. Consider an allocation such that poor households whose children have ability $b \in \left[\underline{b}, B^2\right]$ are indifferent between the two districts. Suppose they randomise and go to district 0 with probability $\mu(b) \in [0,1]$ satisfying (7). $\mu(b)$

determines the average ability in the low track in the two districts, (9) and (10). The requirement that these households be indifferent is

$$v\left(y_P - p\right) + \theta_1^\mu = v\left(y_P\right) + \theta_0^\mu; \tag{A1}$$

which gives the price (8), provided the difference $\theta_1^\mu - \theta_0^\mu$ is positive. Next, notice that, given price (8), rich households whose children have ability $b \in \left[\underline{b}, B^2\right]$ prefer district 1:

$$v\left(y_R - p\right) + \theta_1^\mu > v\left(y_R\right) + \theta_0^\mu, \tag{A2}$$

since (A1) and Lemma 1 imply (A2).

Now consider households with high ability children, that is children with $b \in \left(B^3, \overline{b}\right]$. Take poor households first. In the candidate equilibrium they live in district 1, which gives them utility $v\left(y_P - p\right) + \theta^3$, and so they must weakly prefer staying put over moving to district 0:

$$v\left(y_P - p\right) + \theta^3 \geq v\left(y_P\right) + \theta^2. \tag{A3}$$

Given (A1), this will be the case provided

$$\theta^3 - \theta^2 \geq \theta_1^\mu - \theta_0^\mu. \tag{A4}$$

Consider next rich households with high ability children: at the equilibrium allocation they have utility $v\left(y_R - p\right) + \theta^3$. If they deviate and move to district 0, they will be allocated to the high track and have utility $v\left(y_R\right) + \theta^2$. If (A4) holds, by Lemma 1, $v\left(y_R - p\right) + \theta^3 > v\left(y_R\right) + \theta^2$, and so indeed they strictly prefer district 1. Finally, consider households with children with ability $b \in \left(\theta^2, \theta^3\right]$, both rich and poor. They live in district 0, and are assigned to the high track there, obtaining utility $v\left(y_i\right) + \theta^2$. If they moved to district 1, they would be in the low track, which would give them utility $v\left(y_i - p\right) + \theta_1^\mu$; since $\theta^2 > \theta_1^\mu$, they would pay more for housing *and* have a lower peer group, and so clearly they prefer to stay put.

So existence of equilibrium hinges on the existence of a randomisation function $\mu\left(b\right)$ such that (A4) holds. In (A4), the LHS is exogenously given. The RHS varies as $\mu\left(b\right)$ changes, reaching a minimum when $\mu\left(b\right) = \mu_{\min}\left(b\right)$, and a maximum when $\mu\left(b\right) = \mu_{\max}\left(b\right)$, where

$$\mu_{\min}\left(b\right) = \begin{cases} 0 & b \in \left[\underline{b}, B_P^*\right) \\ 1 & b \in \left[B_P^*, B^2\right] \end{cases}, \tag{A5}$$

$$\mu_{\max}\left(b\right) = \begin{cases} 1 & b \in \left[\underline{b}, \Phi_P^{-1}\left(\frac{1}{4}\right)\right] \\ 0 & b \in \left(\Phi_P^{-1}\left(\frac{1}{4}\right), B^2\right] \end{cases}. \tag{A6}$$
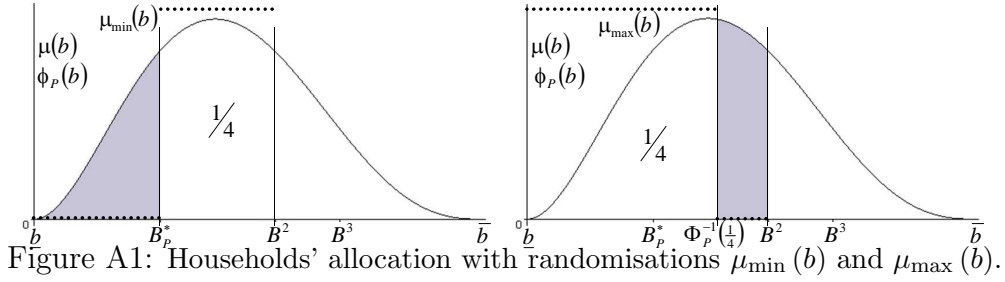
A2

Figure A1: Households' allocation with randomisations $\mu_{\min}(b)$ and $\mu_{\max}(b)$.

That is $\theta_1^\mu - \theta_0^\mu$ in (A4) reaches a minimum when the least able children from poor households all live in district 1, and a maximum when they all live in district 0, as illustrated in Figure A1. This depicts the ability distribution for poor households, and, as a dotted line, the randomisation rules, $\mu_{\min}(b)$ on the LHS and $\mu_{\max}(b)$ on the RHS panel. In each case, households in the grey area live in district 1, as do those where the child has ability exceeding B^2. When $\mu(b) = \mu_{\max}(b)$, then θ_1^μ and θ_0^μ become:

$$\theta_0^\mu = 4 \int_{\underline{b}}^{\Phi_P^{-1}\left(\frac{1}{4}\right)} b\phi_P(b)\,\mathrm{d}b,$$

$$\theta_1^\mu = 4 \int_{\underline{b}}^{\Phi_P^{-1}\left(\frac{1}{4}\right)} b\phi_R(b)\,\mathrm{d}b + 4 \int_{\Phi_P^{-1}\left(\frac{1}{4}\right)}^{B^2} b\left(\phi_P(b) + \phi_R(b)\right)\,\mathrm{d}b.$$

And so the difference $\theta_1^\mu - \theta_0^\mu$ is

$$4 \left(\int_{\Phi_P^{-1}\left(\frac{1}{4}\right)}^{B^2} b\left(\phi_P(b) + \phi_R(b)\right)\,\mathrm{d}b - \int_{\underline{b}}^{\Phi_P^{-1}\left(\frac{1}{4}\right)} b\left(\phi_P(b) - \phi_R(b)\right)\,\mathrm{d}b \right).$$

The first term in the above is $\theta_1^\mu - A$, and the second $\theta_0^\mu - A$, where $A = \int_{\underline{b}}^{\Phi_P^{-1}\left(\frac{1}{4}\right)} b\phi_R(b)\,\mathrm{d}b$. If the second term is negative or 0, then clearly $\theta_1^\mu - \theta_0^\mu > 0$. If the second term is instead positive, then note that, multiplied by $\frac{4}{4-\mu_A}$, where $\mu_A = \Phi_R\left(\Phi_P^{-1}\left(\frac{1}{4}\right)\right) = \int_{\underline{b}}^{\Phi_P^{-1}\left(\frac{1}{4}\right)} \phi_R(b)\,\mathrm{d}b$ is the measure of the rich households with low ability children, each is an average abililty. But the first term is an average of abilities above $\Phi_P^{-1}\left(\frac{1}{4}\right)$, the second term an average of abilities below $\Phi_P^{-1}\left(\frac{1}{4}\right)$, and therefore their difference is positive. Therefore it is always possible to find a randomisation such that the property price in district 1 is positive. However $\theta_1^\mu - \theta_0^\mu$ cannot exceed $\theta^3 - \theta^2$ for every possible randomisation, and so, if $\theta^3 - \theta^2$ is larger than the minimum possible value of the RHS, then there is at least one randomisation function $\mu(b)$ which ensures that (A4) holds, and the Proposition is established. This minimum possible value is

reached when $\mu(b) = \mu_{\min}(b)$, and in this case (A4) becomes:

$$\theta^3 - \theta^2 \geq 4\left[\int_{\underline{b}}^{B^2} b\left(\phi_P(b) + \phi_R(b)\right)\mathrm{d}b - 2\int_{B_P^*}^{B^2} b\phi_P(b)\,\mathrm{d}b\right] = 2\left(\theta_m - \theta_P^*\right). \quad (A7)$$

The average ability in the population, θ_{av} can be written as

$$\theta_{av} = \frac{\theta_m}{2} + \frac{\theta^3 + \theta^2}{4},$$

and so (A7) becomes:

$$\theta^3 - \theta^2 \geq 4\theta_{av} - \theta^3 - \theta^2 - 2\theta_P^*.$$

Rearranging, (5) is obtained. Therefore, if (5) holds, then there is at least one randomisation function $\mu(b)$ which ensures that (A4) holds, and the Proposition is established. ∎

Proof of Corollary 1. This is simple counting: in a desegregation equilibrium with positive price, there are two groups of poor households living in district 1: the black area and all but $\frac{1}{4}$ of the lightest grey area in Figure 2. The black area has measure $\frac{1}{2} - \Phi_P(B^3)$. The lightest area has measure $\Phi_P(B^2)$. ∎

Proof of Proposition 3. Consider an allocation such that households where the child has ability in the upper (in the third) quartile live in district 1 (in district 0), rich households with children whose ability is below the median live in district 0, and poor households whose children have ability $b \in [\underline{b}, B^2]$ are indifferent between the two districts. Suppose the latter randomise and go to district 1 with probability $\lambda(b) \in [0, 1]$, where $\lambda(b)$ satisfies

$$\int_{\underline{b}}^{B^2} \lambda(b)\,\phi_P(b)\,\mathrm{d}b = \frac{1}{4}.$$

For the same reason as $\mu(b)$ in Proposition 2, $\lambda(b)$ is also well defined. $\lambda(b)$ determines the average ability in the low track in the two districts, which we denote θ_0^λ and θ_1^λ (analogously to θ_0^μ and θ_1^μ). The requirement that these households be indifferent is

$$v(y_P - p) + \theta_1^\lambda = v(y_P) + \theta_0^\lambda;$$

which gives a negative price $\Delta_P^{-1}\left(\theta_1^\lambda - \theta_0^\lambda\right)$, analogously to (8), provided the difference $\theta_1^\lambda - \theta_0^\lambda$ is negative. Given this price, Lemma 1 implies that rich households whose children have ability $b \in [\underline{b}, B^2]$ strictly prefer district 0:

$$v(y_R - p) + \theta_1^\lambda < v(y_R) + \theta_0^\lambda.$$

A4

Next, consider households with high ability children, that is with $b \in \left(B^3, \bar{b}\right]$. In the candidate equilibrium they live in district 1, which gives them utility $v\left(y_i - p\right) + \theta^3$, $i = P, R$ and so they must weakly prefer staying put over moving to district 0:

$$v\left(y_i - p\right) + \theta^3 \geq v\left(y_i\right) + \theta^2; \qquad i = P, R.$$

This is clearly the case because in district 1 they enjoy both higher school quality and larger private consumption.

To end the proof, consider poor households with children with ability $b \in \left(B^2, B^3\right]$. They live in district 0, and are assigned to the high track there, obtaining utility $v\left(y_P\right) + \theta^2$. If they moved to district 1, they would be in the low track, which would give them utility $v\left(y_P - p\right) + \theta_1^\lambda$. Since $\theta^2 > \theta_1^\lambda$ and $p < 0$, we need to check that the property price is not sufficiently negative so as to induce them to prefer district 1. The equilibrium price is $p = \Delta_P^{-1}\left(\theta_1^\lambda - \theta_0^\lambda\right)$, while the price that would make these households indifferent between the two districts is $\Delta_P^{-1}\left(\theta_1^\lambda - \theta_2\right)$. Clearly, $\theta_2 > \theta_0^\lambda$ and so $\theta_1^\lambda - \theta_0^\lambda > \theta_1^\lambda - \theta_2$ and $\Delta_P^{-1}\left(\theta_1^\lambda - \theta_2\right) < p$, that is, to be willing to move to district 1, these households would demand a lower property price than the equilibrium one. Hence they prefer to stay put. Since $\theta^2 > \theta_1^\lambda$, by Lemma 1, rich households with children of ability $b \in \left(B^2, B^3\right]$ also prefer to stay put. ∎

Proof of Lemma 2. Begin with $i = L$. We want to determine the sign of $\theta_L^R - \theta_L^P$. So we can write

$$\theta_L^R - \theta_L^P = \left(4 \int_{\underline{b}}^{B_2^R} b\phi_R\left(b\right) \mathrm{d}b\right) - \left(4 \int_{\underline{b}}^{B_2^P} b\phi_P\left(b\right) \mathrm{d}b\right)$$

$$= 4 \left(\int_{\underline{b}}^{B_2^P} b\left(\phi_R\left(b\right) - \phi_P\left(b\right)\right) \mathrm{d}b + \int_{B_2^P}^{B_2^R} b\phi_R\left(b\right) \mathrm{d}b\right).$$

Integrate by parts both integrals, and write the above as

$$\frac{\theta_L^R - \theta_L^P}{4} = B_2^R \Phi_R\left(B_2^R\right) - B_2^P \Phi_P\left(B_2^P\right) + \int_{\underline{b}}^{B_2^P}\left(\Phi_P\left(b\right) - \Phi_R\left(b\right)\right) \mathrm{d}b - \int_{B_2^P}^{B_2^R} \Phi_R\left(b\right) \mathrm{d}b \tag{A8}$$

but B_2^i is the median of $\Phi_i\left(b\right)$, $i = R, P$, and so

$$\Phi_R\left(B_2^R\right) = \Phi_P\left(B_2^P\right)$$

and (A8) can be written as:

$$\frac{\theta_L^R - \theta_L^P}{4} = B_2^R \Phi_R\left(B_2^R\right) - B_2^P \Phi_R\left(B_2^R\right) - \int_{B_2^P}^{B_2^R} \Phi_R\left(b\right) \mathrm{d}b + \int_{\underline{b}}^{B_2^P}\left(\Phi_P\left(b\right) - \Phi_R\left(b\right)\right) \mathrm{d}b$$

$$= \int_{B_2^P}^{B_2^R}\left(\Phi_R\left(B_2^R\right) - \Phi_R\left(b\right)\right) \mathrm{d}b + \int_{\underline{b}}^{B_2^P}\left(\Phi_P\left(b\right) - \Phi_R\left(b\right)\right) \mathrm{d}b$$

the first term is positive because $\Phi_R(b)$ is increasing, the second because of stochastic dominance.

Similarly for $\theta_H^R - \theta_H^P$, which is

$$\theta_H^R - \theta_H^P = \left(4 \int_{B_2^R}^{\bar{b}} b\phi_R(b)\, \mathrm{d}b \right) - \left(4 \int_{B_2^P}^{\bar{b}} b\phi_P(b)\, \mathrm{d}b \right)$$

$$= \left(4 \int_{B_2^R}^{\bar{b}} b\left(\Phi_R(b) - \Phi_P(b)\right)\, \mathrm{d}b \right) - 4 \int_{B_2^P}^{B_2^R} b\phi_P(b)\, \mathrm{d}b$$

Again integrate by parts both integrals, and write

$$\frac{\theta_H^R - \theta_H^P}{4} = \int_{B_2^R}^{\bar{b}} \left(\Phi_P(b) - \Phi_R(b)\right)\, \mathrm{d}b + \int_{B_2^P}^{B_2^R} \left(\Phi_P(b) - \Phi_P\left(B_2^P\right)\right)\, \mathrm{d}b$$

which again is positive. ∎

Proof of Proposition 4. Given $\varepsilon > 0$, at a perfect income segregation equilibrium, the average ability in the four tracks is "close"[25] to the following values:

$$
\begin{array}{llll}
\text{District 1:} & \text{High track:} & \theta_H^R, & \text{Low track:} \quad \theta_L^R. \\
\text{District 0:} & \text{High track:} & \theta_H^P, & \text{Low track:} \quad \theta_L^R.
\end{array}
\tag{A9}
$$

Perfect income segregation characterises equilibrium if and only if there is a price for district 1 housing such that no household wants to deviate, and, moreover, that some rich households are indifferent (given that ε is positive). To check for this, there are six types of households to consider: rich and poor households with high, middle and low ability children. The high (low) ability children would be in the high (low) track in either district. A household with a middle ability child, instead, would be in the high track in district 0 and in the low track in district 1. This is a consequence of correlation between income and ability, which implies that in a full segregation equilibrium, the cut-off ability between tracks is higher in district 1.

---

[25]With strictly positive ε, a perfect segregation equilibrium cannot exist, as it is not possible to squeeze $\frac{1}{2}$ rich households into $\frac{1}{2} - \varepsilon$ houses. Rigorously, (A9) should read

$$
\begin{array}{llll}
\text{District 1:} & \text{High track:} & \theta_H^R + g_H^R(\varepsilon), & \text{Low track:} \quad \theta_L^R + g_L^R(\varepsilon). \\
\text{District 0:} & \text{High track:} & \theta_H^P + g_H^P(\varepsilon), & \text{Low track:} \quad \theta_L^P + g_L^P(\varepsilon).
\end{array}
$$

where the functions $g_i^j(\varepsilon)$, $j = R, P$, $i = H, L$ are all continuous functions, defined in $\left[0, \frac{1}{4}\right]$, which tend to 0 as ε tends to 0. This would be unnecessarily cumbersome, and is left implicit.

Consider rich households first. The following must hold:

$$v\left(y_R - p\right) + \theta_H^R \geqslant v\left(y_R\right) + \theta_H^P,$$
$$v\left(y_R - p\right) + \theta_L^R \geqslant v\left(y_R\right) + \theta_H^P,$$
$$v\left(y_R - p\right) + \theta_L^R \geqslant v\left(y_R\right) + \theta_L^P,$$

with at least one equality. These can be written as:

$$\Delta_R\left(p\right) = \min\left\{\left(\theta_H^R - \theta_H^P\right), \left(\theta_L^R - \theta_H^P\right), \left(\theta_L^R - \theta_L^P\right)\right\}. \tag{A10}$$

The RHS of (A10) is in fact $\theta_L^R - \theta_H^P$ (because $\theta_H^R > \theta_L^R$ and $\theta_H^P > \theta_L^P$), and so (A10) reduces to $\Delta_R\left(p\right) = \theta_L^R - \theta_H^P$, which determines the housing price in district 1, (13). For this to be an equilibrium, all poor households must prefer to stay in district 0:

$$\Delta_P\left(p\right) \geqslant \theta_H^R - \theta_H^P, \tag{A11}$$
$$\Delta_P\left(p\right) \geqslant \theta_L^R - \theta_H^P = \Delta_R\left(p\right), \tag{A12}$$
$$\Delta_P\left(p\right) \geqslant \theta_L^R - \theta_L^P. \tag{A13}$$

Notice that the RHS in (A12) is lower than the RHS in both (A11) and (A13), and so (A12) holds if (A11) and (A13) do, which is the case if (12) holds. This completes the proof. ∎

Proof of Proposition 5. We begin with the following Lemma, which is also of independent interest.

Lemma A1 *If the poor households with low ability children are better off with tracking, then all households are better-off with tracking.*

Proof. Consider a desegregation equilibrium with a negative house price in district 1where the randomisation $\lambda\left(b\right)$ determines low track abilities θ_0^λ and θ_1^λ in districts 0 and 1, and let the price be $p_D < 0$. If poor households with low ability children are better off with tracking, then:

$$v\left(y_P - p_D\right) + \theta_1^\lambda = v\left(y_P\right) + \theta_0^\lambda > v\left(y_P\right) + \theta_P, \tag{A14}$$

that is: $\theta_0^\lambda > \theta_P$. Consider rich households with low ability children. Their payoff in the no tracking case is $v\left(y_R - p_N\right) + \theta_R$, where p_N is the price in the equilibrium with no tracking. In the desegregation equilibrium considered, their payoff is $v\left(y_R\right) + \theta_0^\lambda$. We can write:

$$v\left(y_R\right) + \theta_0^\lambda > v\left(y_R\right) + \theta_P = v\left(y_R - p_N\right) + \theta_R.$$

The first inequality follows from $\theta_0^\lambda > \theta_P$, and the equality at the end follows from the fact that with no tracking, the rich must be indifferent between the two districts. Compare the first and second term to see that rich households with low ability children prefer tracking strictly. Next consider the rest of the households. We have

$$v(y_i) + \theta_P < v(y_i) + \theta_0^\lambda < v(y_i) + \theta_2 < v(y_i - p_D) + \theta_3, \quad i = P, R.$$

The first inequality follows again from $\theta_0^\lambda > \theta_P$, the second from the fact that θ_2 is the average of a subset of above-the-median abilities, θ_0^λ the average of a subset of below-the-median abilities, and the third from the fact that $\theta_2 < \theta_3$ and $p_D < 0$. Comparing the first and the third (fourth) term shows that households with middle (high) ability children are better off with tracking. ■

We can now return to the proof of the Proposition. As (A14) shows, for low ability poor households to prefer tracking, there must exist a randomisation $\lambda(b)$ such that $\theta_0^\lambda > \theta_P$. If this is true at the highest possible value for θ_0^λ, then we are done. Since existence of the equilibrium with a negative price requires $p_D < 0$, and so $\theta_1^\lambda - \theta_0^\lambda < 0$, the highest possible value for θ_0^λ is obtained when $\lambda(b) = \mu_{\max}(b)$ (given by (A5)), or:

$$\lambda(b) = \begin{cases} 1 & b \in \left[\underline{b}, \Phi_P^{-1}\left(\frac{1}{4}\right)\right] \\ 0 & b \in \left(\Phi_P^{-1}\left(\frac{1}{4}\right), B^2\right] \end{cases}.$$

Condition (14) requires that this maximum value of θ_0^λ be greater than θ_P, which concludes the proof. ■

Proof of Corollary 2. Consider first households with middling ability children. They are better off with tracking if and only if:

$$v(y_i) + \theta_2 > v(y_i) + \theta_P; \quad i = P, R. \tag{A15}$$

That is, if and only if $\theta_2 > \theta_P$. Consider next households with high ability children, and again compare utility levels: these households are better off with tracking if and only if $v(y_i - p_T) + \theta_3 > v(y_i) + \theta_P$, $i = P, R$. Using (8) to determine p_T, we have that $\theta_3 - \theta_P > \theta_1^\mu - \theta_0^\mu$ is sufficient and necessary for the inequality to hold for poor households, and sufficient for rich ones. To establish the corollary, note first that $\theta_2 > \theta_P$ implies $\theta_3 - \theta_P > \theta_3 - \theta_2$, and second that existence of the desegregation equilibrium entails $\theta_3 - \theta_2 > \theta_1^\mu - \theta_0^\mu$. ■

Appendix 2

Figure 2 sketches how the equilibrium price (in the case of a positive price with almost all the rich households living in district 1), and the quality of the high track

vary as ε varies. The grey area is the range of possible prices for given $\varepsilon \geqslant 0$, and the hatched area the range of possible qualities if the top track in district 1. The function $\zeta_\varepsilon(b)$ is the density of the top quartile rich households who reside in district 0, clearly, with $\int_{B^3}^{b} \zeta_\varepsilon(b) \, \mathrm{d}b = \varepsilon$.

$$p$$

$$\frac{1}{4 - \varepsilon} \int_{\underline{b}}^{\bar{b}} b \left(\phi_R(b) \, \zeta_\varepsilon(b) \right) \mathrm{d}b$$
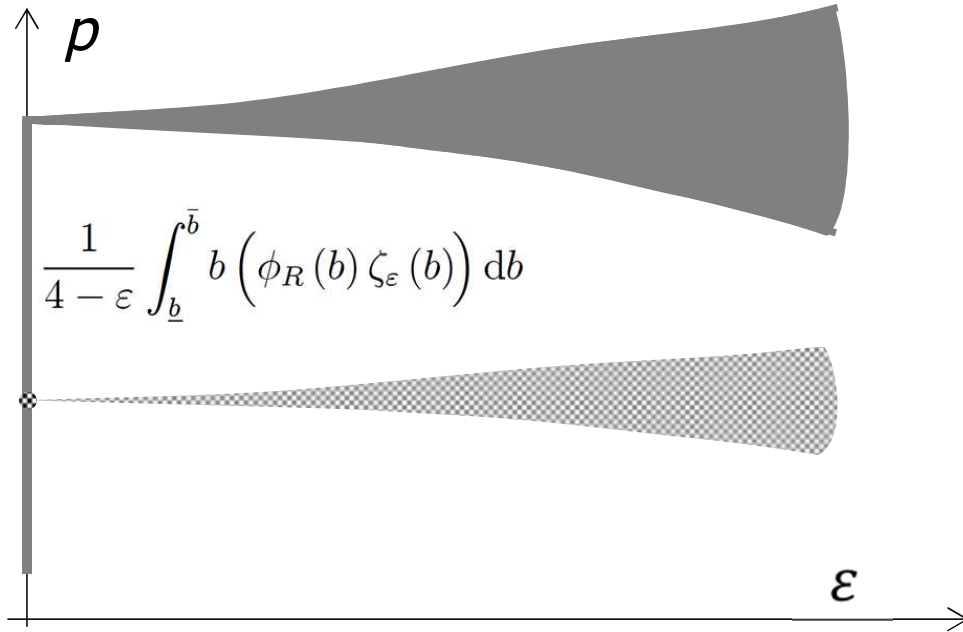
$$\varepsilon$$

Figure 2: The price (solid area) and quality (hatched area) correspondences as ε changes.

To obtain an intuition for the shape sketched in the Figure, note that, for a strictly positive ε there is a multiplicity of equilibrium prices, and consequently quality levels in the two districts: there are multiple equilibria, depending on which of the rich households are "forced" to move out of district 1: all rich households with high ability children have the same willingness to pay for school quality, and so the rationing rule must be random: however, the excluded households bid the price up to the level where all rich households are indifferent between the district. When the excluded households happen to be those with the lowest ability children (among the children with ability in the top quartile, formally children with ability in $\left[B^3, B^3 + 4\varepsilon\right)$), the quality of the top track in district 1 and hence the price are highest; vice versa, if the excluded households are those with the ablest children (children with ability in $[\underline{b} - 4\varepsilon, \underline{b}]$), then the quality of the top track in district 1 and hence the price are the lowest possible. All prices in between these are possible for

A9

some distribution of the "excluded" children, as shown by the grey area in Figure 2. As ε shrinks, the range of possible prices shrinks, and given that

$$\lim_{\varepsilon \to 0^+} \int_{\underline{B^3}+4\varepsilon}^{b} b\Phi_R(b)\, \mathrm{d}b = \lim_{\varepsilon \to 0^+} \int_{\underline{B^3}}^{\underline{b}-4\varepsilon} b\Phi_R(b)\, \mathrm{d}b = \int_{\underline{B^3}}^{b} b\Phi_R(b)\, \mathrm{d}b$$

price and peer quality in the top track in district 1 tend to a unique value. In the above expression, the first term is the maximal peer quality, the second the minimal peer quality, and clearly, both tend to a common value, the third term.

Things are different when $\varepsilon = 0$, however. In this case, all rich households can live in district 1, so it is no longer required that the price is so high as to ensure that they are indifferent between the district, but can come down a bit to ensure that they are strictly better off in district 1. However, the price has a second role to play, namely to ensure that the poor households with low ability children are indifferent: when ε is positive, the price is fixed by the indifference condition for rich households, and so the allocation of poor households to the two districts must be such that the resulting low track qualities ensure that these households are indifferent between the two districts. However, when the price is free to vary, that is for $\varepsilon = 0$, then it can do so, provided the poor households can allocate themselves in such a way that the resulting track qualities in the bottom tracks make them indifferent between districts. If the price becomes too low, then all the poor households would want to move to district 1, which would push price up.

To to sum up, the equilibrium set converges to a point as $\varepsilon \to 0$, but "explodes" when $\varepsilon = 0$ is 0, as sketched in Figure 2.