School choice and neighborhood sorting: Equilibrium consequences of geographic school admissions

Ellen Greaves and Hélène Turon*

Discussion Paper 24/779

28 November 2023

School of Economics

University of Bristol
Priory Road Complex
Bristol
BS8 1TU
United Kingdom

University of Bristol School of Economics

School choice and neighborhood sorting: Equilibrium consequences of geographic school admissions

Ellen Greaves and Hélène Turon*

November 28, 2023

Geographic school admissions criteria bind residential and school choices for some parents, and could create externalities in equilibrium for non-parents through displacement or higher rent. Through a dynamic structural model, we show that the policy decision of geographic versus non-geographic school admissions criteria has important implications for equilibrium outcomes in school and housing markets. Geographic admissions criteria segregate schools, but integrate neighborhoods according to income. Incorporating non-parents into the model challenges the existing understanding of how public schools affect the housing market: non-parent households dampen the equilibrium price premium around popular schools; non-parent households are never better off under geographic admissions.
JEL codes: I21, I24, R21

Keywords: School choice, residential choice, school admissions criteria

---

∗. Greaves: European University Institute, Florence, Italy, ellen.greaves@eui.eu. Turon: University of Bristol, Bristol, UK, helene.turon-lacarrieu@bristol.ac.uk. We are grateful to seminar participants at the European University Institute, University of Bristol, University of Lancaster, the Institute for Fiscal Studies and following conferences for helpful advice: CMPO, EEA, International Workshop on the Applied Economics of Education, Journées Louis-André Gérard-Varet, Royal Holloway PhD students' conference. All errors are our own.

1. Introduction

Wherever location partly determines school access, richer neighborhoods tend to have 'better' schools, as defined by their pupils' educational attainment. This is both because richer pupils have higher attainment, on average, and because richer households can selectively sort into neighborhoods close to 'good' schools. This cyclical process might have long-lasting implications for social mobility and inequality, given the importance of education for individual and societal outcomes. Under a system of school choice, households are able to apply to other schools than their local one, which partly breaks the deterministic school assignment given location. Indeed, one objective of providing school choice is to widen access to 'good' schools to pupils from a greater range of socio-economic backgrounds.[1] But, how pupils are ranked at oversubscribed schools - by location or otherwise - is of first-order importance for the eventual distributional effects of school choice.

Our paper studies school admissions criteria - the rules schools use to rank pupils when they are oversubscribed. Despite the decisive effect these criteria can have on school assignment, they have been largely neglected in existing theoretical and empirical research. We answer the question of how schools' admissions criteria affect the eventual allocation of pupils to schools *and* households across neighborhoods. Recognizing the potential for endogenous residential sorting is important, because failure to account for this channel 'may lead to an incomplete understanding about the distributional consequences of school choice' (Avery and Pathak 2021).[2] We assess the distributional effects of geographic school admissions criteria on households with and without children, to explore the externalities of the public school system for the large group of non-parent households.

The innovative feature of our approach is that we build a dynamic structural model of household choices across different life-stages, allowing for heterogeneity in household types along

1. Among the multiple objectives of school choice, one objective is to encourage competition among schools and thereby increase the overall quality of schooling services. Evidence for this channel of impact is however mixed (see for example Hoxby 2000 and Rothstein 2007), and is not a mechanism we will examine in this paper.

2. Avery and Pathak 2021 model the residential mobility across districts with and without school choice programs. Their central finding is that school choice (with pupils assigned by lottery when oversubscribed) narrows the gap between the highest and lowest quality schools, and therefore equilibrium rents across districts. Then, high income types move away to a 'no choice' district. This is because 'high types' value the provision of a 'high-cost, high-quality public school' that they can gain access to through the housing market in 'no choice' districts.

completed fertility. This model serves two important purposes. First, we use the model to illustrate mechanisms at work in school and neighborhood formation. These insights challenge the existing interpretation of property price premia around 'good' schools. Second, we use the model to simulate the effect of changing admissions criteria from geographic admissions to lottery on school and neighborhood composition.

Key components of our model are that: (i) households make residential choices across two adjacent neighborhoods to maximize their expected lifetime utility; (ii) there is a positive correlation between school and neighborhood quality; (iii) there are different household types, some of which gain utility from schooling and neighborhood amenities (parents with younger children) while others gain utility only from neighborhood amenities (those without children or with post school age children); (iv) households make optimal decisions taking into account their whole lifecycle, incorporating the expected probability of admission to the higher quality school, expected costs of travelling to school, and moving costs across periods, which makes their problem dynamic.

To elaborate, the dynamic components of our framework come from the existence of moving costs, and from the sibling priority rule, where the household's younger sibling is guaranteed a place in the school that the older sibling attends.[3] Households' residential choices therefore exhibit some persistence because households are forward-looking. Although residential location matters decisively in the life-stage when parents apply to a secondary school, forward-looking households that expect to become parents may choose to move close to a 'good' school before they have children, or make their secondary school choices.

The theoretical model alone reveals important channels that have so far been absent in empirical studies of school choice and related property price effects. For example, there is a large empirical literature using boundary discontinuity design, beginning with Black 1999, on parents' willingness to pay for local school quality. These studies typically find that households are willing to pay a property price premium for access to higher performing schools.[4] Our model shows that non-parents and older households dampen the market equilibrium property prices around high performing schools. As such, our model shows that the reduced form es-

3. In our model, having two rather than one child increases the value placed on school quality.

4. See Gibbons and Machin 2008, Black and Machin 2011 and Nguyen-Hoang and Yinger 2011 for detailed summaries of this literature.

timates should be interpreted as a weighted average of demand from parents and non-parents, rather than parents' valuation of school quality (as is typically done). Our model also makes clear that differences in empirical estimates of the relationship between school quality and property prices across contexts are not necessarily due to differences in parents' valuation of school quality, but could be due to differences in the market composition.

Illustrating another unexplored mechanism, our model also shows that households without dependent children also dampen the estimated effect of school choice reforms, for example those studied in previous empirical work (Reback 2005; Bogart and Cromwell 2000; Ries and Somerville 2010; Lee 2015; Machin and Salvanes 2016). This is because households with children in the school choice phase are the only households that are directly affected by the reforms, as they can be allocated to a school further afield. These households must therefore calculate whether their lifetime utility is higher by incurring moving costs but lowering transport costs to school, or vice versa. All other households may be affected by changes in the equilibrium rents or the expectation of allocation to each school. The estimated effect of a school choice reform on property prices is therefore affected by the market composition, travel costs and moving costs.

The model shows the effect of the public school system on the private housing market, in particular, the externalities for non-parents. The model makes clear that non-parents are affected through two channels. First, high demand from parents in the neighborhood around the 'good' school creates additional upward pressure on the rent. The higher rent has a negative impact on the utility of non-parents who choose to live in this neighborhood. Second, some non-parents that would choose to live in this neighborhood in the absence of schooling are displaced to the lower quality area, which decreases their utility.

To quantify the magnitude of these mechanisms, we estimate our model using data from two neighborhoods in one city in England. One more affluent area has the Above Average school, while one less affluent area has the Below Average school. We find that our model replicates patterns of sorting in schools and neighborhood across completed fertility types and life-stages. In our setting, geographic admissions contribute to segregation by household type (age and completed fertility, in addition to income) at the neighborhood as well as school level. These factors have hitherto been unexplored and not considered, but it is reasonable to assume that segregation by household type is as problematic as segregation by income for societal outcomes. For example, under the current system with geographic admissions, the neighborhood with the Above Average school has 10% more households with dependent children, particularly with two or more children (28%) and at school choice (53%) or secondary school (58%) age. Property prices in the neighborhood with the Above Average school are 20% higher, which is due to both school and neighborhood factors. The rate of pupils eligible for free school meals (an indicator for poverty) is 83% lower in the neighborhood with the Above Average school. Although these are not clean causal estimates, the pattern in the data is consistent with the model's predictions that geographic admissions is associated with segregation across neighborhoods as well as schools.

We find, counter to first intuition, that the simulated effect of admissions by lottery has opposing effects on segregation by income at the school and neighborhood levels. As expected from the random assignment, schools become more integrated. Surprisingly, however, neighborhoods become slightly more segregated by income. This is because lower income parents no longer sort into the high quality neighborhood with the Above Average school when the link between school and residential choice is broken. Neighborhoods are therefore perfectly sorted by income, but do become more integrated by household age and completed fertility. These results illustrate the complex interaction between neighborhood and school choices, and the possibly unintended consequences of school choice policies on societal outcomes.

Our first main contribution is to study the effect of school admissions criteria rather than the allocation mechanism on equilibrium outcomes for the school and neighborhood. We complement the model provided by Calsamiglia, Martínez-Mora, and Miralles 2020, who study segregation by income with a focus on equilibrium sorting under alternative allocation mechanisms (Boston and Deferred Acceptance) rather than school admissions.[5] The baseline model of Calsamiglia, Martínez-Mora, and Miralles 2020 excludes residential mobility. In extensions, the authors discuss the role of transport costs and geographic admissions. Strict geographic admissions or prohibitively high transport costs across districts leads to Perfect Assortative Matching. In relation, our dynamic model additionally includes lifecycle considerations such as moving costs and different household types. Calsamiglia and Miralles 2023 assess the role of geographic admissions on preventing school choice (specifically 'ac-

5. See also De Fraja and Martínez-Mora 2014 for a model of how within school-tracking affects school and neighborhood sorting.

cess to better schools') more explicitly than Calsamiglia, Martínez-Mora, and Miralles 2020, but without the addition of endogenous residential mobility. Calsamiglia and Miralles 2023 conclude that the choice of allocation mechanism is marginal compared to the role of school zones, concluding that, 'future studies should incorporate the design of these priorities as a crucial part of the mechanism design problem'.[6]

Our second main contribution is to consider the welfare effects of the school choice environment for households that never have dependent children, or whose dependent children have left home. We find that welfare for households that are always non-parents is weakly lower under a geographic admissions system compared to a lottery system. This is because of higher rents in the desirable neighborhood and displacement across neighborhoods for some households (around the middle of the income distribution). These externalities are important to consider, as non-parent households are a large share of the market (30% of households in our setting). Despite a rich body of theoretical research modelling the joint decisions of school and residential location (see for example Nechyba 2000; Epple and Romano 2003; Ferreyra 2007; De Fraja and Martínez-Mora 2014; Park and Hahm 2023), non-parents have only been incorporated by three (Caetano 2019; Agostinelli, Luflade, and Martellini 2021; Pietrabissa 2023). We complement Caetano 2019 by focusing on the welfare effects for households without dependent children, rather than using these households for identification. We differ from Agostinelli, Luflade, and Martellini 2021 and Pietrabissa 2023 in classifying households by their completed fertility, rather than the current presence of children, which allows us to model dynamic considerations.

Finally, we show that transport costs affect residential choices in addition to school choices. Transport costs have been included by a handful of previous studies (Epple and Romano 2003; Calsamiglia, Martínez-Mora, and Miralles 2020; Agostinelli, Luflade, and Martellini 2021; Park and Hahm 2023; Pietrabissa 2023) but no other theoretical models of school and neighborhood choice.[7] In relation to these papers, we expand on the mechanisms through which

6. In the context we study, the allocation mechanism is truth-revealing, equivalent to the Gale-Shapley deferred acceptance algorithm.

7. Epple and Romano 2003 model the introduction of district-wide open enrollment, from traditional neighborhood enrollment. In the equilibrium with neighborhood enrollment, the model predicts that income stratification implies school quality stratification because the access to neighborhoods with better peer groups is rationed by higher housing prices. In the equilibrium with open enrollment and no transport costs, the model predicts equal school qualities and equal property prices across neighborhoods. Finally, in the equilibrium with open enrollment and transport costs, only higher income households are able to choose schools, while lower income

transport costs affect school and residential choices in a dynamic setting. We show that the direction of the effect of a change in transport costs depends on the initial equilibrium. For example, where the probability of access to the school is initially low, then increasing transport costs reduces demand to live in the surrounding neighborhood. This decreases rents, (in)decreases the share of (non-)parent households, and increases the probability of admission to the school from this neighborhood. Our model reveals that changing transport costs therefore have equilibrium effects in both school and housing markets.

The policy implications of our results are that the design of public services can have wide-ranging effects beyond the immediate recipients. Eventual outcomes will depend on the local context, for example, the correlation between neighborhood and school quality and the local transport service in this case. For school choice, the design of the school admissions criteria affects both the welfare of those directly affected (households with children), and all other households through the housing market. We show that reform to school admissions has unequal effects across household types, and across schools and neighborhoods. There is no Pareto improvement, considering the welfare of all households, by removing or introducing geographic school admissions. Policy-makers concerned with equity and social mobility therefore have trade-offs to make when designing school admissions criteria.

Section 2 provides the context for our study, before section 3 presents our model. Section 4 illustrates the mechanisms of the model, at the same time presenting the intuition of the model. This could be a substitute for the formal details of the model in section 3 for some readers. Section 5 then presents some stylized facts of our chosen neighborhoods, which clearly show the model's predictions. Section 6 shows the results of our estimation and the simulation of an alternative school admissions criteria that breaks the link between residential choice and school access. Section 7 discusses some potential extensions to our model, before section 8 concludes.

2. Context

We focus on households' choice of residential location and public (non-private) secondary school in England, which are similar to high schools in the U.S. Pupils start secondary school

households live and attend the school in the poorer neighborhood.

in September of the academic year they will turn 12, and leave after five years in the academic year they turn 16.

Public schools are centrally funded by Government according to a progressive per-pupil formula (Sibieta 2015). The formula provides additional funding for every pupil with an educational disadvantage, such as eligibility for free school meals, English as an additional language, or special educational needs. In this sense, school quality is not endogenous to financing from the local tax base, as in the U.S.

To access the public school system, all parents/guardians are required to submit school preferences for their favorite public schools to their coordinating body - the Local Authority - in the form of a ranked ordered list.[8] Each Local Authority runs a weakly truth-revealing allocation mechanism, equivalent to the Gale-Shapley Deferred Acceptance algorithm, to assign pupils to schools. The allocation mechanism takes parents' submitted preferences and schools' capacity constraints and admissions criteria as inputs to assign each pupil to their highest ranked school possible.

Secondary schools' admissions criteria typically include giving priority to younger siblings and those geographically close (Burgess et al. 2023). 55% of secondary schools have a pre-defined school zone, while 89% have a distance tie-breaking rule. As such, geography is a deciding factor in school admissions for most secondary schools in England.[9]

3. Model of dynamic neighborhood choice

This section gives a formal description of our model of neighborhood and school choice and its solution. Section 4 describes the intuition of the model discursively and illustrates key mechanisms that shape equilibrium outcomes.

8. There are 152 Local Authorities in England. The 33 Local Authorities in London combine to form the Pan-London Admissions Authority for this process. See Greaves 2023 for more information.

9. A minority of schools in England have other arrangements, for example around 160 selective schools that take only the highest performing pupils. Around 40% of secondary schools have a feeder primary school, in which case the households' decision of primary school is an important deciding factor for access to secondary schools (Burgess et al. 2023). In our empirical application, both schools have standard geographic admissions criteria.

3.1. Environment

We set up a simple environment in order to show the consequences of school admissions criteria on the allocation of school places and dwellings across households of different income, life-stage and completed fertility – i.e. parents vs. non-parents, one-child vs two-children households.

We assume our environment to have reached a steady state where all variables of interest are fixed and no changes, except aging, childbirth and income growth, are anticipated by households. These lifecycle events happen in a deterministic manner.

We consider a universe of two neighborhoods n, denoted H (high) and L (low), offering different levels of amenities, $a_H > a_L$. The housing market consists of D dwellings overall, which are available to rent. The rent in neighborhood H is denoted r_n for $n = H, L$ and adjusts to clear the rental market. The numbers of dwellings in each neighborhood are D_n for $n = H, L$ and are fixed exogenously.

There are N households living across the two neighborhoods, which we follow over the lifecycle. These households are unitary and forward-looking. They can be one of three 'fertility' types f, i.e. their completed fertility can be 0, 1 or 2 children. Households know their type and have no uncertainty over the timing (and number) of births. We rule out divorce and premature death.

Households also differ by income. We assume that income growth only depends on fertility type and life-stage so that each household's rank, r, within their fertility/life-stage specific distribution, $F_{f,t}$, remains constant. Household income, y_t, can therefore be retrieved by:

$$y_t = F_{f,t}^{-1}(r) \tag{1}$$

We define four life-stages, t_t, for households of all fertility types relative to key events for households with children. All children attend a public (non-private) secondary school. The first period, t_0, starts when households enter adult life and need to make a choice of residential location, and ends at the age where they need to apply for a place in a secondary school for their eldest child. The time when households apply for a place in secondary school is denoted t_1. The third period, t_2, relates to the years when the household has children going to a secondary school, and the last period, t_3, is the remaining lifetime of the household after the children have finished secondary school. In our stylized representation of the timing of these events, we assume that the duration of each period is 15, 2, 8 and 15 years, with adult life starting at the age of 25. At this age, we assume households to be formed and education to be completed, since income ranks remain constant thereafter.

Preferences are assumed to be homogeneous across households. They value consumption, school quality, neighborhood amenities and proximity to school in a separable manner. In utility terms, the commute to school across neighborhoods is costly, but the commute to work is the same from both neighborhoods.[10] When moving across neighborhoods between life-stages, households incur a one-off utility cost, m. For simplicity, we assume that households of different ages and sizes rent similar dwellings and all pay the same rent, i.e. r_H in neighborhood H and r_L in neighborhood L.

There is one secondary school in each neighborhood and they differ in quality: the school in the H neighborhood is rated by all as Above Average (AA) while the school in the L neighborhood is Below Average (BA). School quality is therefore correlated with neighborhood 'quality' as measured by amenities. We assume homogeneity of preferences over amenities and schools.

The total number of school places P equals the size of the cohort of children in our environment and each school has a fixed number of places, P_{AA} and P_{BA}, respectively. Since all households value the AA school more than the BA school, some rationing takes place. We will examine two possible school admissions criteria to deal with the excess demand: priority to closer applicants ('geographic admissions') with a random draw among applicants from the same neighborhood, or a random draw ('lottery') among all applicants (regardless of place of residence). We also assume that schools have a 'siblings' criteria, where younger siblings are guaranteed a place in the school of the elder sibling. This will affect the decisions of households with more than one child. The probabilities of being granted a place for the eldest child in the AA school are denoted π_H and (respectively π_L) for households living in neighborhood H (respectively L).

In this framework, households only make one choice per life-stage: the choice of neigh-

10. Pietrabissa 2023 models both the commuting costs for parents to work and children to school, to illustrate how school and labor markets interact. In our model, with only two adjacent neighborhoods, we make the simplifying assumption that travel time to work is equal across neighborhoods.

borhood in which to reside. Saving and borrowing are ruled out. All households with an elder child aged 10 apply to the *AA* school. All agents have perfect information but are not able to coordinate. There are externalities through congestion on the housing market through the rent and on the market for school places through the probabilities of admission to the *AA* school, π_H and π_L. Endogenous outcomes of our model are these two quantities, as well as the composition of each neighborhood in the income distribution, the age distribution and the completed fertility distribution.[11]

Both the housing market and the market for school places clear over the two neighborhoods by definition, since there are N households and $D = N$ dwellings in total, and P children in each cohort for P school places across the two schools.

3.2. Formal definitions

In this section, we provide the formal specification of the model ingredients. The time period is one year. Households are 'born' into our environment with a fertility type f, $f = 0..2$. The probability of being of type f is denoted ϕ_f. f is known to the household and constant over the lifecycle. Households of type $f > 0$ have their first child at age 30. Households of type $f = 2$ have their second child at age 32.[12] When the elder child reaches age 10, i.e. the household reaches age 40, applications to secondary school have to be made. The number of children of secondary school age in a household of type f in life-stage t is therefore deterministic and denoted $k(f,t)$.

Denote τ_t the fraction of total adult life spent in phase t.Since our environment is in steady-state, the fraction of the population of households in the life-stage t and of fertility type f is constant and equal to $\tau_t \cdot \phi_f$.[13] The size of a cohort of children is $N\tau_1(\phi_1 + 2\phi_2)$.[14] Of these, $N\tau_1\phi_2$ are younger siblings, who either have a place reserved in the *AA* school because their older sibling was awarded one, or choose not to apply to the *AA* school because their older sibling already attends the *BA* school.[15]

11. To be precise, income, age and fertility are all exogenous processes, but the selection of different groups across the two neighborhoods is endogenous.

12. We omit the possibility of twins or of having 3 or more children. This does not affect our qualitative results.

13. Note that we have $\sum_{t=0}^{3}\sum_{f=0}^{2}\tau_t\phi_f = 1$.

14. Note that in our empirical application this is divided by two, as we model this life-stage as lasting two years to account for the time of application.

15. We have ruled out households commuting to two different schools by assumption, which is reasonable

The total number of applicants to the AA school without a guaranteed place is therefore A, of which A_H (respectively A_L) reside in neighborhood H (respectively L). The relative sizes of A_H and A_L depend on the rents in the two neighborhoods, $\mathbf{r} = \{r_L, r_H\}$, and on the probabilities of acceptance into the AA school from either neighborhood, $\pi = \{\pi_L, \pi_H\}$:

$$\begin{aligned} A_H &= A_H^1(\mathbf{r}, \pi) + A_H^2(\mathbf{r}, \pi) \\ A_L &= A_L^1(\mathbf{r}, \pi) + A_L^2(\mathbf{r}, \pi) \end{aligned} \tag{2}$$

where A_n^f denotes the number of applications by households of type f residing in neighborhood n for their elder child.

The total number of school places P per cohort across the two schools AA and BA is equal to the size of the cohort of children. The AA school has P_{AA} places and the BA school has P_{BA} places. These numbers are fixed by policy and do not respond to excess demand in our model.

$$P = P_{AA} + P_{BA} = N\tau_1(\phi_1 + 2\phi_2) \tag{3}$$

The number of households of type $f = 2$ who are successful in their application to the AA school with their eldest child and therefore have a place reserved for their second child is denoted S and depends on these households' neighborhood choices:

$$S(\mathbf{r}, \pi) = A_H^2.\pi_H + A_L^2.\pi_L \tag{4}$$

Let us now turn to residential choices. The flow utility for a household of fertility type f, income y, and life-stage t has the following expression when it chooses to reside in neighborhood n and their child(ren) k, if any, attend secondary school s:[16]

$$\begin{aligned} U(f, y, k, s, n) = &\frac{(y - r_n)^{(1-\gamma)}}{1-\gamma} + \alpha_a \mathbb{1}(n = H) \\ &+ \mathbb{1}_{(k>0)}\alpha_s(1 + \mathbb{1}_{(k=2)}\alpha_2)s - \alpha_c c \end{aligned} \tag{5}$$

where c is a dummy variable capturing the need to commute to school across neighborhoods

as there are likely to be large monetary (uniform/travel) and non-monetary (administration/travel time) costs of siblings attending different schools. Our assumption is that these costs outweigh the utility of one child attending the AA school when another child attends the BA school.

16. The neighborhood and school variables are defined as follows: $n = H$ if the household resides in the H neighborhood and L otherwise; $s = 1$ if the child(ren) in the household attend the AA school and 0 otherwise.

defined as:

$$c = \mathbb{1}_{(k>0)}\mathbb{1}\{(n=H).(1-s)+(n=L).s\}$$

In other words, $c=1$ when the child(ren) attend a school in the neighborhood where they do not reside, yielding a disutility of α_c for the household.

We assume a standard utility of consumption, where all income net of rent is consumed. The household derives utility α_a from the amenities in the high-quality neighborhood relative to the low-quality neighborhood. If k is positive, i.e. the household has children of secondary school age, it derives a utility α_s (respectively $\alpha_s+\alpha_2$) from having one (respectively two) child(ren) attending the AA secondary school relative to the BA one.

The two dynamic components of our framework are the following. First, we assume the presence of a cost α_m of moving across neighborhoods between life-stages. Second, for households of fertility type 2, the school for the younger child is determined by the school place granted to the older child, which may depend on neighborhood choice. In addition, household income grows over time in an exogenous process which differs by fertility type, and agents anticipate this.

Households discount the future at a yearly rate of β and choose their neighborhood in order to maximize their lifetime utility:

$$\begin{aligned} V_t(f,y_t,s,n_{-1}) = \max_n \Bigg\{ \left(\sum_{z=0}^{\tau_t-1}\beta^z\right) U\Big(f,y_t,k(f,t),s,n\Big) \\ -\alpha_m m + \beta^{\tau_t} E_{s'}\{V_{t+1}(f,y_{t+1},s',n)\}\Bigg\} \end{aligned} \tag{6}$$

where n_{-1} is the neighborhood chosen in the previous life-stage, if any, and m captures the event of moving, $m=\mathbb{1}\{n\neq n_{-1}\}$. $V_4=0$ is the household utility at the end of life. The uncertainty captured by the expectation factor in the last term above only applies to the choice made in life-stage $t=1$. At this point, households make their choice according to the probabilities that their children will attend the school s in the next life-stage:

$$E_s\{V_2(f,y,s,n)\} = \pi_n V_2(f,y,AA,n) + (1-\pi_n)V_2(f,y,BA,n) \tag{7}$$

In all other life-stages, there is no uncertainty in household choices.

3.3. Equilibrium

The two endogenous outcomes in our framework are the probabilities of admission to the *AA* school from either neighborhood, π_H and π_L, and the rents in each neighborhood, r_H and r_L. In equilibrium, these will be such that both the housing market and the 'market' for school places clear in both neighborhoods.

In the school market, the number of places available in the *AA* school, i.e. not reserved for younger siblings of existing pupils, is $P_{AA} - S$. In the case of lottery admissions, the probability of being granted a place in the *AA* school for households without a sibling priority is constant across neighborhoods:

$$\pi_H = \pi_L = \frac{P_{AA} - S}{A} \tag{8}$$

Under geographic admissions, these probabilities vary by neighborhood, such that the probability of access from L, π_L is only non-zero when there are fewer applicants from H than available school places in AA:

$$\begin{aligned} \pi_H &= min\left(1, \frac{P_{AA} - S}{A_H}\right) \\ \pi_L &= \frac{P_{AA} - S - \pi_H . A_H}{A_L} = max\left(0, \frac{P_{AA} - S - A_H}{A_L}\right) \end{aligned} \tag{9}$$

Lemma 1. *The quadruplet* $\{A_H, A_L, \pi_H, \pi_L\}$ *is such that the market for school places clears in both neighborhoods, for any admissions system.*

Proof. See Appendix A. □

Turning to the housing market, each households' comparison of neighborhoods hinges on the disutility incurred by lower consumption in the H neighborhood (because the rent is higher), relative to the utility derived from amenities. In addition, for parent households, the utility derived from (the higher probability of) their child(ren) attending the *AA* school and not commuting to school. All but the disutility of the higher rent are independent of income, so we have the following result:

Lemma 2. *For each household type f and life-stage t there is a unique rank threshold $R_{f,t}(\mathbf{r}, \pi)$ above which households choose to live in neighborhood H when rent levels are at $\mathbf{r}$ and probabilities of admission to the AA school are at π.*

Proof. See Appendix A. □

Summing up over all household types and life-stages, we obtain the total number of households choosing to live in neighborhood H:

$$N_H(\mathbf{r},\pi) = \sum_{f=0}^{2}\sum_{t=0}^{3} \tau_t \phi_f \left[1 - R_{f,t}(\mathbf{r},\pi)\right] \tag{10}$$

Where $R_{f,t}$ denotes the percentile of income within the group-specific income distribution, where group is defined by type f and life-stage t. $R_{f,t}$ therefore equates to the proportion of households earning less than the income level $F_{f,t}^{-1}(R_{f,t})$.

The number of applicants to the AA school, i.e. all households with children at the time their elder child reaches secondary school age, is, from each neighborhood:

$$\begin{aligned} A_H(\mathbf{r},\pi) &= N\tau_1\Big(\phi_1\left[1 - R_{1,1}(\mathbf{r},\pi)\right] + \phi_2\left[1 - R_{2,1}(\mathbf{r},\pi)\right]\Big) \\ A_L(\mathbf{r},\pi) &= N\tau_1\Big(\phi_1 R_{1,1}(\mathbf{r},\pi) + \phi_2 R_{2,1}(\mathbf{r},\pi)\Big) \end{aligned} \tag{11}$$

We can now define the equilibrium in the set of two neighborhoods $\{L,H\}$.

Definition. *Equilibrium definition*

- *An equilibrium is a set of probabilities π of being offered a place in the AA school for one's elder child, a set of rents in both neighborhoods* $\mathbf{r}$ *and a list of lifetime values $V_t(f,r)$ for households of type f with income rank r in life-stage t.*[17]

- *The number of households in the environment, N, the fractions of completed fertility types, ϕ_f, the fractions of lifetime spent in each stage τ_t, the numbers of dwellings in each neighborhood, D_n, and the number of places in each school, P_{AA} and P_{BA}, are considered fixed and exogenous.*

- *Given $(\mathbf{r},\pi)$, there exists a threshold $R_{f,t}$ for all f,t such that the household with rank r's optimal choice of neighborhood is H if and only if $r > R_{f,t}$.*

17. For ease of exposition, we omit the dependence of the lifetime values $V(.)$ and the thresholds $R(.)$ on school attended s and chosen neighborhood in the previous life-stage n_{-1}. All our results follow when these variables are included.

- *Given the thresholds $R_{f,t}$, the housing market and the 'market' for school places both clear:*

$$\begin{aligned} N_H(\mathbf{r},\pi) &= D_H \\ A_H(\mathbf{r},\pi).\pi_H + A_L(\mathbf{r},\pi).\pi_L + S(\mathbf{r},\pi) &= P_{AA} \end{aligned}$$

Note that the total number of dwellings in the $\{H,L\}$ environment, $D = D_H + D_L$ equals the total number of households, N, so that if the housing market clears in neighborhood H, it does too in neighborhood L. Similarly, on the 'market' for school places, the total number of school places in the $\{H,L\}$ area, $P = P_{AA} + P_{BA}$, equals the total number of children per cohort, so that if the AA school is full, all remaining children have a place in the BA school.

Proposition 1. *For each set of values $\{N, \phi_f, \tau_t, D_n, P_{AA}, P_{BA}\}_{f=0..2,t=0..3,n=H,L}$, an equilibrium $(\mathbf{r},\pi)$ exists and is unique.*

Proof. See Appendix A. □

4. Mechanisms

This section illustrates the mechanisms of the model and shows how our simple framework deepens the understanding of existing empirical reduced form results. This section is intentionally discursive to clearly describe the channels through which residential and school choices affect equilibrium outcomes, such as the rents across neighborhoods and segregation in neighborhood and schools. First, in section 4.1, we explore the large body of literature relating property prices to local school quality, specifically the interpretation that parents' valuation of school quality can be inferred from the causal increases in property prices around popular schools. Second, in section 4.2, we illustrate some underlying mechanisms that affect the new equilibrium when geographic admissions/school zones are replaced with freer school choice, taking Machin and Salvanes 2016 as a case study. Third, we explore how the cost of transport to school interacts with the school choice environment to shape equilibrium outcomes (section 4.3).

This section is intended to stand alone, but key ideas from section 3 are useful. These are that: (i) there are different household types, some of which gain utility from schooling and neighborhood amenities (parents with younger children) while others gain utility only from neighborhood amenities (those without children or with post school age children); (ii) each household type has a rent threshold above which they choose to live in the neighborhood with high rather low quality than amenities (H rather than L); (iii) parents with school aged child(ren) gain utility from attending the Above Average (AA) school compared to the Below Average (BA) school; (iv) households make optimal decisions taking into account their whole lifecycle, incorporating the expected probability of admission to the AA school and moving costs across periods, which makes their problem dynamic.

Throughout this section, we will see that households without children of secondary school age (or below) play a crucial role in accommodating changes in demand for residence in the desirable neighborhood by younger parent households, by reacting to changes in the difference in rent. One overarching intuition is that price responses to changes in the school environment would be larger if the economy consisted of only parent households, as larger price changes would be needed to equalize supply and demand.

4.1. School quality and property prices

The introduction described the large empirical literature on the willingness to pay for local school quality. Since Black 1999, boundary discontinuity design has typically been used to estimate the causal effect of local school quality on property prices, where the identifying assumption is that neighborhood attributes are continuous at the school zone boundary, while school quality jumps discontinuously. Across studies in this literature, there is variation in the property price premium across contexts, from 1.4% in Paris to 3.5% in England.[18]

This reduced form effect has been interpreted as parents' valuation of school quality. For example, Black 1999 states that the 'value that parents place on school quality' is given by 'calculating how much more people pay for houses located in areas with better schools'. Our model reveals three caveats to this interpretation.

The first is context specific, depending on if a tie-breaking rule used in school assignment.

18. The existing empirical estimates from school attendance boundaries/school zones (in order of magnitude) are: Fack and Grenet 2010 (Paris): 1.4% for school test scores, 2.1-2.4% for school peer-group; Bayer, Ferreira, and McMillan 2007 (San Francisco Bay): 1.8%; Black 1999 (Massachusetts): 2.1%; Harjunen, Kortelainen, and Saarimaa 2018 (Helsinki): 3.3%; Davidoff and Leigh 2008 (Australia): 3.5%; Gibbons, Machin, and Silva 2013 (England): 3.5%.

The boundary discontinuity design approach implicitly assumes that residing on either side of the boundary makes a 0/1 difference in the prospect of attending the *AA* school. However, if the *AA* school becomes oversubscribed from within the school zone and a tie-breaking rule is used, then the probability of acceptance to the *AA* school becomes less than one. Maintaining the assumption that amenities are constant over the border, the difference in rent locally on either side of the border captures the amenity of the *AA* school quality *weighted* by the probability of acceptance on the *H* side of the border – assuming that this probability is 0 on the *L* side of the border.

The second caveat is that the methodology assumes that neighborhood amenities are identical on both sides of the border (see Bayer, Ferreira, and McMillan 2007 for an exception). A consequence of this, according to our framework, is that no non-parent households should live locally on the *H* side of the border. Even a small difference in amenities will induce the richest non-parents to live on the *H* side, however, and the difference in rent will then capture a weighted average of the willingness to pay for parent and non-parent households.[19] It follows that the equilibrium rent depends on the share of parent to non-parent households in the economy.

Proposition 2. *An increase in school quality or an increase in the valuation of the AA school by parent households leads to an increase in the rent differential between the two neighborhoods, which reflects:*

- *The marginal utility of attending the AA school times the difference in probabilities of acceptance to AA from residing in H rather L.*
- *A weighted average of the valuation of amenities in H by the fraction of non-parent households living in H and of the valuation of both amenities and of higher school quality by the fraction of parent households living in H.*
- *The increase in the rent differential only equates to the marginal willingness to pay for higher school quality if all households are parents.*

Proof. See Appendix A. □

19. For parent households, the willingness to pay for the *AA* school *plus* the small difference in amenities. For non-parent households, the willingness to pay for the difference in amenities.

The direction of the bias in existing estimates is unclear. The first caveat leads to the difference in rent being an underestimate of parents' willingness to pay for attending the *AA* school, for example. The second caveat leads to the difference in rent being an overestimate, however, assuming a positive correlation between school quality and neighborhood amenities.

In the absence of schooling in our stylized model, neighborhoods would be perfectly sorted according to income. Neighborhood H would be home to the highest income households due to its high-quality amenities. Introducing the school market decreases income segregation across neighborhoods, as parents' threshold to live in H is lower than non-parents'. Non-parents whose threshold is below the new equilibrium rent choose to live in L, while parents (with lower income than these displaced non-parents) choose to live in H.[20]

When schooling is introduced, the equilibrium rent in the neighborhood with the *AA* school increases as the share of parents in the whole population increases. At either extreme, if there are only parents or only non-parents, then the intuition in the previous paragraph follows: neighborhoods would be perfectly sorted according to income. For example, in the absence of non-parents, any shift in the demand for dwellings in H by parent households must be fully offset by an increase in rent so that the housing market clears.

Instead, when there are both types of households, any shift in demand for dwellings in H by parent households causes a rise in equilibrium rent in H, which displaces some non-parent households. The new equilibrium has more parent households and fewer non-parent households living in H.

As the share of parents in the whole population increases, the equilibrium rent in H increases. This is due to two channels. First, there is a higher proportion of households that value both the amenities in H and the *AA* school. This leads to increased pressure upwards on the rent in H. Second, there is a lower share of non-parents that only value the amenities in H and who therefore dampen the equilibrium rent increase.[21]

How does a change in the share of non-parent households affect the school market? As the proportion of parents increases (and the rent in H increases), a smaller *share* of parents choose

20. Lower income parents have a higher marginal utility of consumption than higher income non-parents, but choose to live in H because they gain utility from the school amenity.

21. In the H neighborhood, the marginal non-parent household is richer than the marginal parent household and is therefore less sensitive to a marginal increase in rent. A given rent increase puts off the marginal parent household more than the marginal non-parent household. This mitigates somewhat the dampening effect on rent adjustments by the presence of non-parent households.

to live in H, which increases the probability of admission to the AA school from H.

The AA school becomes more selective by income when parents care less about school quality, which contradicts first intuition. This is because when the utility from attending the AA school is low, only the richest parents choose to live in H (partly due to the high neighborhood amenities) and so the school becomes exclusive. When the utility from schooling increases, parents' income threshold to live in H decreases, so a broader range of parents choose to live in H and gain (a positive probability of) access to the AA school.

Overall, this discussion has shown that the equilibrium effect of local school quality depends on the market composition, as non-parents dampen the price response to local school quality. The effect of population composition may therefore help to reconcile differences in reduced form estimates across contexts. For example, prices rise by 1.4% for a one-standard deviation in test scores in Paris (Fack and Grenet 2010) compared to 3.5% in England (Gibbons, Machin, and Silva 2013). Our model makes clear that these differences across contexts are not necessarily due to differences in parents' valuation of school quality. Differences in the reduced form causal effect could be driven by population composition, even if the valuation of school quality is the same.[22]

It is clear from this illustration that parents' valuation of local quality can not be inferred directly from the causal reduced form increase in price around 'good' schools. This intuition extends to papers which study the effect of school quality information (score cards and/or inspections) on parents' demand, inferred through local property prices. An information shock that changes perceptions about local school quality leads to a shock in the demand for dwellings in H by parent households. The resulting pressure on the housing market in H raises the rent in H, which pushes the poorest non-parent households living in H to move to L. The rent response to the initial information shock is therefore increasingly muted as the share of non-parents initially living in H increases.

22. Other potential mechanisms that reconcile results across studies that are not included in our model are differences in the elasticity of housing supply (Cheshire and Sheppard 2004; Harjunen, Kortelainen, and Saarimaa 2018), the supply of 'family homes' (Ries and Somerville 2010), the availability of private school outside options (Fack and Grenet 2010), the availability of school quality information, the stage of education, the exact empirical specification and the time period.

4.2. Change in the school admission system

We now consider the underlying mechanisms that our model reveals when the school admissions system changes. This change could be from geographic admissions to any other system that weakens the link between home and school attended. For example, Machin and Salvanes 2016 study the property price effects of a reform in Oslo county that moved from a geographic admissions system to allocation by ability (where applicants were ordered by ability in line for their preferred schools).[23] Machin and Salvanes 2016 find that the property price premium around 'good' schools decreased by at least 50% following the reform, which is interpreted as showing that 'parents substantially value better performing schools'.

As discussed in section 4.1, it is problematic to infer this price change directly as parents' valuation of access to local school quality. Primarily, this is because the overall price change is a weighted average of the demand from different groups in the population, for example older and younger households and households with and without children. Only a sub-set of these households receive the flow utility of priority at the local school, and so only a subset of households is directly affected by the reform.

Other households are affected indirectly by the reform. Some households may change residential location in response to changes in equilibrium prices across neighborhoods. Parents of child(ren) who are unsuccessful in admission to their local high school under the ability admissions system could choose to pay higher transport costs or move closer to their child's allocated school. Parents of child(ren) or expectant parents who anticipate being unsuccessful in admission to their local high school may also change residential location in advance. We show the following result:

Proposition 3. *A change of the school admissions system from a geographic to lottery admissions system leads to:*

- *A lower rent in H, with the magnitude of the decrease being smaller as the fraction of non-parent households increases.*

23. We simulate the effect of a change from geographic admissions to lottery admissions in section 6. Rehm and Filippova 2008 study the re-introduction of school zones in New Zealand in 2000, after the country's experiment with school choice. Lee 2015 studies the introduction of school districts in South Korea (away from an exam based entry system), while Reback 2005 studies the introduction of school choice in Minnesota.

- *An increase (respectively decrease) in the fraction of households in H who are non-parents (respectively parents).*
- *Greater adjustments in neighborhood and school composition if the number of parent households living in H was initially close to the number of school places in the AA school.*

Proof. See Appendix A. □

Machin and Salvanes 2016 observe that the property price premium does not fall to zero after the Oslo reforms. The authors suggest that the most likely explanation is that 'persistent neighborhood differences induced by the former school zones remain', such as unobserved differences in neighborhood quality and neighborhood peers from the pre-reform days. Our model is consistent with this explanation, as the presence of moving costs generates persistence in neighborhood composition that the authors describe. This is despite our assumption that households gain no flow utility from neighborhood composition - in our model, moving costs alone generate the persistence.

Our model also offers further potential explanations for the remaining premium. The intuition is that all households maximize their expected lifetime utility. After the reform, households will calculate their expected lifetime utility from moving or remaining in their current location. Moving costs mean that not all households will choose to move from their pre-reform location, even if they lose priority at their local school, or equilibrium rents increase. Households with child(ren) in the school choice phase are the only households that are directly affected by the reform, as they can be allocated to a school further afield. These households must therefore calculate whether lifetime utility is higher by incurring moving costs but lowering transport costs, or vice versa. Households with younger child(ren) or those who expect to become parents must make a similar calculation, additionally accounting for the uncertainty about allocation to school.

Equilibrium rents in H fall most in response to the reform when moving costs are low and transport costs are low. The intuition is that when moving costs are low, it is more likely that households re-optimize location in response to the reforms, with parents moving away from H. When transport costs are low, it is less likely that households relocate to the neighborhood where their child(ren) are assigned. A discussion of external validity of studies of this kind must therefore consider the moving costs and transport costs in the study setting, in addition to the population composition. The next subsection considers changes in transport costs in more depth.

4.3. Changes in transport costs

Transport to schools varies across and within countries, depending on the quality and cost of the public transport system, provision of school buses, and density and location of schools relative to neighborhoods. Differences in the cost of travel to school may affect households' school choices. For example, free bus travel provided to pupils in London may expand the choice set of schools that households consider feasible to attend. Outside of London in England, free transport is only provided to pupils attending their 'nearest suitable school' and living a set distance from it.[24] Differences in the cost of travel to school may be one factor that contributes to differences in the patterns of school choice across England, where households in rural areas are much more likely to choose and attend their closest school (Burgess, Greaves, and Vignoles 2019).

There are few empirical papers that directly assess the role of transport on school choice. Trajkovski, Zabel, and Schwartz 2021 study the provision of free school buses in New York City, using cut-offs in eligibility by distance to school to provide causal estimates of the effect of free transport on choice of school. Trajkovski, Zabel, and Schwartz 2021 conclude that, overall, 'bus eligibility plays a significant role in school choice decisions and increases the likelihood of attending a school'. Agostinelli, Luflade, and Martellini 2021 study the role of transport provision in their model of school and neighborhood choice (in a static setting), exploiting changes in bus transportation in Wake County (North Carolina) over time. They find that households consider transport options in addition to commuting distance when choosing schools, and free bus provision is particularly important to lower income households.

Our model highlights that in addition to affecting school choices, reduced transport costs also affect residential choices. Under non-geographic admissions, reducing transport costs allows more households to choose to live in a lower-cost neighborhood while attending an alternative school. In line with this, Boterman 2021 recognizes the combined role of free

24. These distance cut-offs are 2 miles from the school if a pupil is under 8 and 3 miles from the school if a pupil is 8 or older. Source: https://www.gov.uk/free-school-transport.

school choice, high school density and low transport costs in Amsterdam (in this case commuting by bike) in facilitating residential integration in the city. The intuition is that free (non-geographic) school choice and low transport costs separate the residential choice and school choice, while geographic admissions and high transport costs bind them.

Proposition 4. *The effect of an increase in transport costs depends on the initial probability of admission to the AA school:*

- *If the probability of access to AA from H is one, and less than one from L ($\pi_H = 1$ and $\pi_L < 1$), there is increased demand for H (to reduce transport costs), equilibrium rent in H increases, and non-parents move out of H.*
- *If the probability of access to AA from H is less than one, and zero from L ($\pi_H < 1$ and $\pi_L = 0$), there is decreased demand for H (to reduce expected transport costs), equilibrium rent in H decreases, and non-parents move into H.*

Proof. See Appendix A. □

How do transport costs and the school choice environment affect neighborhood composition? Whether the share of households that ever have children is evenly distributed across neighborhoods strongly depends on the cost of transport and school admissions criteria. Under lottery admissions with low transport costs (most similar to the case described by Boterman 2021 for Amsterdam), the share of households of different types in H is close to the population share. When transport costs are higher, then the proportion of households that ever have children increases in H. This is because households find it optimal to live close to their allocated school (and the *AA* school has a higher share of school places than the *BA* school).

Under the geographic admissions system, the transport costs have a more marginal effect on neighborhood composition. Decreasing travel costs slightly increases the proportion of households that ever have children in H, for the same reason discussed in relation to the equilibrium rents: under the geographic admissions system, households with child(ren) have a strong incentive to live in H to gain admission to the *AA* school. If they are unsuccessful in admission (when the number of children in H is greater than supply of school places), transport costs affect whether unlucky households move to L to be closer to the *BA* school.

In contrast, the households' income in H is largely unaffected by changing the cost of transport. This is because higher income households sort into H because of the high neighborhood amenities, regardless of the school admissions system in place.

Overall, our model highlights that transport costs affect residential choices in addition to school choices. When designing school choice systems, policy-makers should consider the specific context for households' choices, in particular the quality and cost of the local transport system. Transport for London planned to remove free bus travel for school children in London (in response to funding cuts related to Covid-19), which could have dramatically affected households' school choices. London, and other cities worldwide, should evaluate the effect of transport policies on school choices and allocations, and potential spillovers to the housing market. Our model shows that these policies are likely to have the largest effect where school choice is 'free' - that is, not geographically constrained.

5. Data and stylized facts

To describe the neighborhoods' characteristics and create empirical moments to estimate the model, we primarily use publicly available sources of data. These are described in full in Appendix B. In brief, information from the census is available for small neighborhoods that broadly nest within school zones, known as lower level super output areas (LSOAs).[25] Through data from the national Census (2011), we can describe household composition in age, socio-economic status, education and presence of dependent children.[26] For prices, property level data are available. To calculate household income growth over time as an input to the model, we use a nationally representative longitudinal study, classifying households according to completed fertility.

We now present the key area- and school-level characteristics across two contiguous neighborhoods in a city in the South West of England. The neighborhoods are chosen to represent one area containing an above average school and high neighborhood quality (corresponding to school AA and neighborhood H in the model), and one area containing a below average

25. LSOAs are small areas of relatively even size (around 1,500 people).
26. Unfortunately, the most recent Census from 2021, does not require all variables we need to estimate the model.

school and reasonably comparable neighborhood quality (school *BA* and neighborhood *L* in the model). These shorthand names do not truly reflect the school and neighborhood characteristics, but are used for convenience to be consistent with the model. ‘Neighborhood quality’ in this context could best be described as proximity to the city center, and amenities such as tree-lined streets. We choose one school to represent the *AA* school as it has the highest academic attainment (and is the only school far above the national average for state schools in England in our sample).[27]

The characteristics of these schools are presented in the bottom panel of Table 1. Pupils attending *AA* have the highest attainment. For example, 83% of pupils achieved at least 5 GSCEs above grade C, compared to 50% in *BA*. There are similarly large differences in the percentage of pupils achieving the English Baccalaureate (EBACC) with 50% in *AA*, compared to 17% in *BA*. Note that these large differences in the attainment of pupils are not necessarily the result of school quality, as final test scores are the result of school, child and parent inputs.

In practice, parents have preferences for the peer composition as well as school quality (Hastings, Kane, and Staiger 2009; Borghans, H. H. Golsteyn, and Zölitz 2015; Burgess et al. 2015; Abdulkadiroğlu, Agarwal, and Pathak 2017; Glazerman and Dotter 2017; Abdulkadiroğlu et al. 2020; Beuermann et al. 2022). The percentage of peers eligible for Free School Meals is a proxy for the intake of the school. This is lowest in *AA*, where only 4% of pupils in 2011 are eligible for Free School Meals, compared to 22% in the city. *BA* is more comparable with the city average, with 24% of pupils eligible for Free School Meals.

Based on the number of pupils of school choice age living in the school zone, and the number of those predicted to be the eldest child, we calculate that the probability of admission to *AA* from *H* is 0.81. As children within the school zone have priority, and there are many of them, the probability of admission to *AA* from *L* is zero.

The top panel of Table 1 shows some key characteristics of the neighborhoods in both school zones. The overall picture is that these two neighborhoods are relatively affluent compared to the city as a whole, which is particularly true for the *AA* school zone. Both neighborhoods have a much higher percentage of households with a ‘managerial, administrative and professional’ occupation: 69.7% of households in the *AA* school zone and 57.1% of households in the *BA* school zone, compared to 38.4% in the city as a whole. There are also fewer households with

27. In 2011, the national average for state schools in England was 58%.

‘semi-routine and routine’ occupations. Similarly, there are far fewer households with ‘low income scores’ from national classifications, and more households with a degree qualification or higher.

Households in the *AA* school zone are more likely to have dependent child(ren): 49.6% compared to 45.3% in the *BA* school zone, and 43.6% in the city overall. This is driven by households with at least two dependent children: 33.2% in *AA* compared to 25.8% in *BA* and 23.3% in the city.

The number of children is related to proximity to the school choice life-stage. The *AA* school zone has fewer children aged one than the *BA* school zone, and a similar number aged four, but many more children of secondary school choice age (10) and secondary school age (16). This suggests that households sort into H in relation to the timing of the secondary school choice process.

Property prices are around 20 to 30% higher in the *AA* compared to the *BA* school zone for the median price, median price for a five room property and median price for a six bedroom property. To account for differences in the composition of homes by property size across H and L, we concentrate on the prices for six room homes, which is the modal number of rooms in H and L.

Overall, Table 1 shows that the predictions from the model broadly match the characteristics of schools and neighborhoods for *AA* and *BA*. There is evidence of sorting by occupation/income into H, and therefore the *AA* school. Indeed, the percentage of poor pupils in *AA* is very low, suggesting an almost perfect sorting by income as predicted by the model. There is also evidence of sorting by life-stage, with more children around the age of secondary school choice and attendance.

We convert these descriptive statistics into moments to estimate the model. First, we take the probability of admission to *AA* from H and L, and the rent difference between H and L as known equilibrium values. The probability of admission is as presented in Table 1 and described in Appendix B. For the rent difference, we convert the median property prices for six room homes into monthly rents using a mortgage rate calculator, as described in Appendix B.

6. Results

This section presents the results of the model, estimated using the data presented in section 5 for the contiguous neighborhoods.

6.1. Empirical and model moments

We estimate the preference parameters in equations (5) and (6) by minimizing the distance between model-predicted and empirical moments. The moments we match are the proportion of households of each completed fertility type and life-stage that choose to live in neighborhood H, where the denominator is all households that choose to live in H or L.[28] Key auxiliary moments are the mean income at the beginning of the lifecycle, and income growth over each life-stage for all fertility types (shown in Appendix B).

Although there is not a one-to-one mapping between parameters and predicted moments, we can provide some intuition for the estimation of each parameter. Moving costs for each fertility type are estimated from the observed change in the proportion of households of each fertility type that choose to live in H across the lifecycle, relative to their change in income across the lifecycle. The value of H depends on the proportion of fertility type f_0 in H, as they gain no direct utility from the AA school. The value of the AA school is therefore inferred from the residential choices of fertility types f_1 and f_2, with the difference between them providing information for the difference between α_S and α_2. Finally, the costs of commuting to school are inferred from the proportion of households with children in the school choice phase that live in H but are predicted to attend the BA school in the contiguous neighborhood. Note that the auxiliary moments of the income profile by fertility type, and equilibrium values of the probability of admission to AA from H and L are also key to the estimation.

Our estimated parameters are shown in Table 3. All parameters but the moving costs relate to monthly utility flows. For intuition about the magnitude of coefficients presented, the marginal willingness to pay for a 0.2 increase in utility is £35, £115, £232 at (monthly) income

28. Note that these proportions are constructed from area-level census data and will contain some measurement error. We exclude the final life-stage (the age consistent with children having left home) as we can not identify households that once lived with children in the data. We choose to match the proportion of households that choose to live in H to relax the assumption that there is no in or out migration from our neighborhoods across the lifecycle.

levels £2,500, £3,500 and £4,500, respectively, and the monthly rent difference between the two neighborhoods is £242. To households with one child, the utility gain from a place in the *AA* school rather than the *BA* school is similar to the utility gain from better neighborhood amenities in H. Households with two children value access to the *AA* school 1.7 times more than households with one child. The disutility of commuting to school, which we interpret as a combination of the disutility of travelling to school daily with that of living away from the children's school peers, is large, amounting to 58% of the amenity provided by the *AA* school. Finally, the cost of moving, which is a one-off cost that we allow to vary across fertility types, is 22 times as large for households with completed fertility of two children than for childless households and 38% larger than for households with completed fertility of only one child.[29] This, together with anticipated income growth over the lifecycle and the cost of commuting to school, explains why households with children do not commonly move back to neighborhood L in life-stage t_2 after securing a place in the *AA* school.

Overall, the model matches the empirical moments well (Table 4). The exception is that f_0 households do not move over the lifecycle. This is because the value of amenities is constant across the lifecycle, and income growth is observed to be minimal for this fertility type. The model therefore can not generate the movement across the lifecycle observed in the empirical moments for this fertility type.

6.2. Simulation of lottery in admissions

With the estimated preference parameters, we can simulate counterfactuals with respect to the school admissions system. We relax the link between residence in H and admission to the *AA* school and assume that the probability of admission is equal across all applicants living in the two neighborhoods, as in equation (8). The sibling rule still applies. The new equilibrium outcome under admission by lottery and the comparison with the initial equilibrium under geographic admissions is shown in Table 5.

The probability of admission to the *AA* school, which was 0% and 81% respectively for applicants residing in the L and H neighborhoods, is now 59% for all (elder sibling) applicants. Living in neighborhood H therefore loses some of its appeal for households with children f_1

29. The moving cost for fertility type f_0 is estimated to be very small, as income grows slowly across the lifecycle and yet more households move into H at later life-stages.

and f_2, particularly before the school choice phase. Under the geographic admissions system, 15% of f_1 households move to H at this stage (between t_0 and t_1), and 18% of f_2 households. Under the lottery admissions system, this decreases to 13% and 10%, respectively.

We see opposing effects of moving to a lottery admissions system on parents and non-parents, as parents become slightly less likely to live in H while non-parents become slightly more likely to. From the beginning of life-stage t_1 onwards, 55% of non-parent households (f_0) choose to reside in neighborhood H in the lottery system versus 54% in the geographic admissions system; 64% of households with one child f_1 (respectively 68% of f_2) live in H under lottery admissions, versus 63% (respectively 75% to 69%) under the geographic admissions system.

The decrease in applicants to the AA school residing in neighborhood H, in line with Figure B.4, reduces the pressure on the housing market in neighborhood H. This results in a lower rent in H, as shown in proposition 3. The new rent in H is now £1,448, representing a 0.8% decrease. Note that this estimate is smaller than most existing reduced form estimates for the property price premium around 'good' schools, but is similar to the estimated 1.4% premium in Paris, for example (Fack and Grenet 2010).

The rent in H remains higher than in L since households value the neighborhood amenities in H highly, and across the lifecycle.[30] The overall effects on the mean incomes in both neighborhoods is minimal: the income threshold above which households with children decided to live in H under the geographic system was lower than that of non-parents who did not value the increased probability of admission to the AA school provided by residence in H. Under the lottery admissions system and a lower rent, this income threshold rises among households with children and decreases for childless households. So the population of non-parents living in H is now less well-off on average, whereas the population of households with children is now more selected and richer on average. Neighborhoods sorting by income is now almost perfect.[31]

30. Recall that households receive a flow utility of amenities from H in each period, while for schooling only in the t_2 life-stage.

31. The only remaining difference between parents' and non-parents' valuation of the two neighborhoods is the commute to school, but since the direction of commute is not predicted well before school places are allocated –the probability of admission to AA is 59%– there is only a small difference in the residential choices among households with and without children. Households could decide to move once school places are allocated, but the under our estimated parameter values they choose not to, given the large one-off cost of moving relative to the annual transport cost for only life-stage t_2.

On the other hand, the distribution of incomes in the two schools changes sharply: the mean income of children's households in the *AA* school was 49% higher than the mean income of children's households in the *BA* school under the geographic admissions system, whereas these means are equal under the lottery system.[32]

Household composition across neighborhoods changes in the expected direction when the lottery replaces the geographic admissions system, but minimally. The proportion of households living in neighborhood H who never have children increases from 26.9% to 27.4%, a 2% increase, while the proportion of households that ever have two children decreases by 1%, from 53.7% to 53.2%. Changes in the age composition of neighborhoods are more muted: we see a small increase in young households in t_0 under the lottery system (2%), offset by small decreases in stages t_2 and t_3.

We can evaluate welfare for all households as the sum of discounted utility flows over the lifecycle. Moving from geographic to lottery admissions will affect households without children that pay lower rents in H, or change location from L to H. Households with children will be affected by the same factors, and additionally the change in probability of access to AA. This will be a negative effect for richer parent households that previously had priority at *AA* under geographic admissions, and a positive effect for poorer parent households that now have a non-zero probability of admission from L under lottery admissions.

Focusing here on households that ever have children, using our estimated parameter values, we compute that 53.7% of parents with two children choose to live in H under the geographic admissions system, compared to 53.2% under the lottery admissions system. Overall, moving from geographic to lottery systems, rich parents (from the 7th decile upwards) are worse off because of the lower probability of access to the *AA* school (59% under the lottery system compared to 81% previously) but marginally better off because of the slightly lower equilibrium rent in H. Parent households who decide to move to L when the admissions system changes from geographic to lottery (in the upper half of the 6th decile) experience a greater change in rent and the same change in the probability of obtaining a place in the *AA* school. Finally, parents in around the bottom half of the income distribution remain in neighborhood L but experience a sharp increase in their probability of access to the *AA* school, from 0% to

32. Note that households with a place in *AA* never decline it because of transport costs, since α_C is less than α_S.

59%, so are now better off.

6.3. Spillovers to Non-Parents

The inclusion of households who remain without children throughout the lifecycle in our model allows us to consider externalities of the choices of households with and without children on one another as the school system changes.[33] If schooling was irrelevant for all households in our model, neighborhoods would be perfectly sorted by income: all households would prefer neighborhood H because of its better amenities, creating upward pressure on rents. At the equilibrium rent, the housing market clears and only the richer households (of any type) find it worthwhile to choose H. In the following, we call this the benchmark equilibrium.

If, however, the schools in the two neighborhoods are perceived to be of different qualities, where the higher quality school is in H, this neighborhood now offers two advantages valued by households with children but only neighborhood amenities for non-parents. This creates externalities for non-parents in two ways. First, this creates additional upward pressure on the rent in H, since households with children have an additional incentive to live in this area. The higher rent has a negative impact on the utility of non-parents who choose to remain in H. Second, some non-parents, whose threshold for living in H is lower than the new higher rent, are displaced to L and enjoy a lower utility than in the benchmark.

This occurs in the geographic admissions system where the desire to reside in H for households with children is the result of a higher probability of gaining a place in the AA school, a closer proximity to that school –and avoidance of the costly commute to school– and neighborhood amenities. Under the lottery system, the latter two advantages of living in H remain, and the demand for housing coming from households with children is higher than in the benchmark. In both cases, the rent is higher than in the benchmark, with a more marked increase under the geographic admission system.

This is illustrated in Figure B.2 from proposition 2. When the school admissions system is based on a geographic criterion, or when there is an increase in the relative quality of the AA school relative to BA, the demand by school by parent households (denoted P in Figure

33. Caetano 2019 includes non-parents in his model of school and residential choice, but does not use this to consider spillovers between groups and welfare calculations. Agostinelli, Luflade, and Martellini 2021 and Pietrabissa 2023 include non-parents, but defined in a static way: those currently with or without children.

B.2) shifts up. Although the demand by non-parent households does not shift, the housing market now clears at a higher rent, with a move of the equilibrium allocation along the non-parent demand curve. This displaces some non-parents to L. At the new equilibrium, the fraction of parent households in H has increased. As explained above, this affects all non-parent households negatively, since they either pay a higher rent or decide to move to the L neighborhood. This decrease in the number of non-parent households wanting to stay in H eases the congestion in H and means that the market clears for a lower rent than would have prevailed in the absence of non-parent households. Their presence therefore dampens the housing market response to changes in school quality or admissions system and has positive externalities on parent households, not only through a lower rent for all residing in H but also for those parent households who can afford to move to H at this rent and would have been priced out in the absence on non-parent households.

Welfare is unambiguously lower for non-parents under a geographic admissions system compared to a lottery system. Using our estimated parameter values, our model shows that 1% of non-parents are displaced from H when the admissions system changes from lottery to geographic admissions. The displaced non-parents are around the middle of the income distribution (the 6th decile of income) and choose L instead of H when the equilibrium rent in H rises as a result of geographic admissions. Welfare falls slightly for these households, given the small change in equilibrium rents. The size of these welfare effects depends on the specific context. Welfare would be more negatively affected where the value of schooling is stronger than the value of neighborhood amenities, as this would create more displacement and rent increases. For the welfare effects to be reversed, there would need to be a negative correlation between neighborhood and school quality, so that non-parents prefer to live in L.

7. Discussion of potential extensions

Our model provides important insights into the relationship between school admissions criteria, school choice, and neighborhood and school sorting. Extensions for future work are possible, of course. In this section we outline three possible extensions (a private school 'outside option', endogenous school quality, and heterogeneous preferences) and the likely implications of these for equilibrium outcomes.

First, incorporating a private school into the market has been shown to be important for the equilibrium outcomes of school choice systems by providing an 'outside option' to richer households (Epple and Romano 2003; Calsamiglia, Fu, and Güell 2020; Calsamiglia, Martínez-Mora, and Miralles 2020). The relevance of this private school option will depend on its quality (relative to the public schools) and price, and the admissions system used. Given the cost of private school, this will change the behavior of richer households, but will not alter the essence of the mechanisms previously described.

As in our main model, household income is the key deciding factor in the trade-off between the value of the higher school quality and the disutility of foregone consumption. Higher income households will always choose the private school above the public schools, if the quality of the private school is higher than that of the Above Average state school. These households will skip the public school application process, and will behave as non-parents in our model. Middle income households would prefer attending the Above Average state school for free to the private school for a fee, but could also prefer this private school option to sending their children to the *BA* school. These households' eventual enrollment might therefore be sensitive to the school assignment. Finally, lower income households will not consider the private school option, whether they have a place in the Above Average school or not, and so will behave as previously described.

Only the school choices and admissions of middle income households are therefore affected by a change in admissions criteria from geographic to non-geographic.[34] These households compare the two neighborhoods characterized by the trade-off between the probability of obtaining a place in the Above Average state school and the higher rent, which changes under geographic vs non-geographic admissions. This trade-off will be similar to that examined in our model, the difference being, for these households, the 'outside option' is to join the private school.[35]

Ultimately, the way in which the equilibrium rent and allocation of households across neigh-

34. Higher income households always choose the private school while lower income households always choose a state school, in both admissions systems. These households will be indirectly affected by changes in equilibrium rents and probability of admission to the Above Average school.

35. We note, also, that whether the allocation mechanism is truth-revealing or not will also affect the equilibrium responses with a private school outside option. In our setting, where there is a weakly truth-revealing mechanism, parents have no strategic advantage with a private school outside option when submitting their state school preferences, as in Madrid (Calsamglia, Fu, and Güell 2018).

borhoods respond to a change in the school admissions system depends on the quality-price package that the private school offers. Unfortunately, we do not observe applications or admissions to private schools in our market, or their price or quality, so this mechanism remains unquantifiable in our setting.

Second, our model assumes that households have preferences for exogenous school quality rather than the endogenous school peer group. This is justifiable as some aspects of schools, such as infrastructure, teaching and management teams are exogenous, at least in the short run. Previous evidence finds that parents also value the peer group of the school as one component of school quality, however (Fack, Grenet, and He 2019; Abdulkadiroğlu et al. 2020).[36] What effect would endogenous school quality have on the equilibrium outcomes of our model?

Since our model predicts that the admissions system has an effect on the distribution of parental income in each school, school quality as defined by the peer group would change across geographic and non-geographic counterfactuals. If this endogenous school quality component was quantitatively important, it could lead to multiple equilibria, where the quality of a school could switch from being less attractive to more attractive than the alternative following a small change in the groups of households choosing one (or gaining a place in) or the other.

Since the geographic system leads to more income sorting across schools, if households value the average household income of school peers, a switch from lottery to geographic system will amplify the difference in school quality and the upward pressure on rent in the H neighborhood from households that ever have children. The return to equilibrium will therefore see more non-parents move out of this neighborhood and a new equilibrium with a higher rent and a higher fraction of households with children in H than in our main analysis. Allowing for endogenous school quality would therefore predict sharper differences between the two admissions systems in rent differentials and in neighborhood composition.

Finally, our baseline model assumes that all households have the same preferences for school quality. This assumption is supported by some empirical work (Greaves and Hussain 2023) but contradicted by others (Hastings, Kane, and Staiger 2009; Borghans, H. H. Golsteyn, and Zölitz 2015; Burgess et al. 2015; Abdulkadiroğlu, Agarwal, and Pathak 2017; Glazerman and Dotter 2017; Harris and Larsen 2019; Ruijs and Oosterbeek 2019; Ajayi and

36. More generally, peer group preferences have been shown theoretically to weaken schools' incentives to exert effort under school choice (Barseghyan, Clark, and Coate 2019).

Sidibe 2020; Walker and Weldon 2020; Beuermann et al. 2022). It could be possible to extend our baseline model to incorporate increasing utility from attending the Above Average school for higher income households, comparing the equilibrium outcomes under various assumptions about the heterogeneity of households' preferences.

These potential extensions do not detract from the key insights and importance of our model. Our simple dynamic framework incorporates households that are affected indirectly by the public school system. This allows the model to provide understanding and reasoning for the interactions between public and private markets along additional dimensions, and challenges the interpretation of existing reduced form estimates. Only with these insights, can academics and policy-makers begin to model the long-run equilibrium effects of school choice reform on all types of households.

8. Conclusion

School choice has the potential to increase access to 'good' schools for children from less advantaged backgrounds. Whether school choice achieves this aim depends crucially on the design of the system. School admissions criteria, for example giving local pupils priority, have a decisive effect on whether a system of school choice lives up to this potential, but previously have been understudied.

We quantify the relationship between the type of school admissions criteria (geographic admissions versus lottery) and the social mix within schools and within neighborhoods, defined as the mix between households of different ages, different income and different completed fertility. To do this, we build the first dynamic structural model of household choices across different life-stages, allowing for heterogeneity in household types. Incorporating non-parents and older households in our dynamic setting allows the model to calculate the externalities of the public school system to the large share of non-parents in society, finding substantial effects.

The model illuminates important mechanisms underlying reduced form estimates of causal effects. For example, the model shows that it is not possible to interpret the estimated relationship between school quality and property prices as parents' valuation of school quality, as the estimated effect will be dampened by the presence of non-parents and older households in the market. A similar logic applies to reduced form estimates of school choice reforms and improving information about school quality. The model also shows the interacting effects of transport costs and moving costs in school and neighborhood choice, which are therefore factors for academics and policy-makers to consider in evaluation and design of school choice environments.

Comparative statics from the model illustrate the potential trade-off between integration in neighborhoods and schools. When residential location does not affect access to the 'good' school, neighborhoods are perfectly sorted according to income, while the 'good' school has a more integrated composition. In the alternative 'geographic admissions' case, where living in the high quality neighborhood increases the probability of admission to the 'good' school, neighborhoods become more integrated by income, but schools become less so. These results suggest a potentially difficult trade-off for a social planner wishing to increase integration between household types, particularly if segregation by household type (age and completed fertility) also has consequences for society.

Indeed, we find that there is no Pareto optimal school admissions system. Although non-parents are unambiguously better off under non-geographic admissions, there are winners and losers among parent households. Typically, lower income households benefit from access to the 'good' school under non-geographic admissions, while for richer households the effect is the opposite. The correct policy response therefore depends on the social planners' weight on the welfare of different household types.

Our model provides important insights into the relationship between school admissions criteria, school choice, and neighborhood and school sorting. There are some limitations, however. First, the model is estimated using data from two contiguous neighborhoods in England. The model could be re-estimated in other contexts to provide more generalizable results. The estimated parameters are sensitive to the auxiliary parameters of the income distribution, which are estimated from national data, and therefore have some measurement error when applied to our local context. Although these extensions would not change the qualitative implications of our model, future work could include additional elements, for example a private school 'outside option', endogenous school quality dependent on the peer group, and heterogeneous preferences by household income.

References

Abdulkadiroğlu, Atila, Nikhil Agarwal, and Parag A. Pathak. 2017. "The Welfare Effects of Coordinated Assignment: Evidence from the New York City High School Match." *American Economic Review* 107, no. 12 (December): 3635–89.

Abdulkadiroğlu, Atila, Parag A. Pathak, Jonathan Schellenberg, and Christopher R. Walters. 2020. "Do Parents Value School Effectiveness?" *American Economic Review* 110, no. 5 (May): 1502–39.

Agostinelli, Francesco, Margaux Luflade, and Paolo Martellini. 2021. *On the Spatial Determinants of Educational Access.* Technical report. HCEO Working Paper Series.

Ajayi, Kehinde F., and Modibo Sidibe. 2020. *School Choice Under Imperfect Information.* Working Paper 294. Economic Research Initiatives at Duke (ERID).

Avery, Christopher, and Parag A. Pathak. 2021. "The Distributional Consequences of Public School Choice." *American Economic Review* 111, no. 1 (January): 129–52.

Barseghyan, Levon, Damon Clark, and Stephen Coate. 2019. "Peer Preferences, School Competition, and the Effects of Public School Choice." *American Economic Journal: Economic Policy* 11, no. 4 (November): 124–58.

Bayer, Patrick, Fernando Ferreira, and Robert McMillan. 2007. "A Unified Framework for Measuring Preferences for Schools and Neighborhoods." *Journal of Political Economy* 115 (4): 588–638.

Beuermann, Diether W, C Kirabo Jackson, Laia Navarro-Sola, and Francisco Pardo. 2022. "What is a Good School, and Can Parents Tell? Evidence on the Multidimensionality of School Output." *The Review of Economic Studies* 90, no. 1 (June): 65–101.

Black, Sandra E. 1999. "Do Better Schools Matter? Parental Valuation of Elementary Education*." *The Quarterly Journal of Economics* 114, no. 2 (May): 577–599.

Black, Sandra E., and Stephen Machin. 2011. "Chapter 10 - Housing Valuations of School Performance." In *Handbook of the Economics of Education,* edited by Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, 3:485–519. Elsevier.

Bogart, William T., and Brian A. Cromwell. 2000. "How Much Is a Neighborhood School Worth?" *Journal of Urban Economics* 47 (2): 280–305.

Borghans, Lex, Bart H. H. Golsteyn, and Ulf Zölitz. 2015. "Parental Preferences for Primary School Characteristics." *The B.E. Journal of Economic Analysis & Policy* 15 (January).

Boterman, Willem R. 2021. "Socio spatial strategies of school selection in a free parental choice context." *Transactions of the Institute of British Geographers* n/a (n/a).

Burgess, Simon, Estelle Cantillon, Mariagrazia Cavallo, Ellen Greaves, and Min Zhang. 2023. *School admissions in England: The rules schools choose on which pupils to admit.*

Burgess, Simon, Ellen Greaves, and Anna Vignoles. 2019. "School choice in England: evidence from national administrative data." *Oxford Review of Education* 45 (5): 690–710.

Burgess, Simon, Ellen Greaves, Anna Vignoles, and Deborah Wilson. 2015. "What Parents Want: School Preferences and School Choice." *The Economic Journal* 125 (587): 1262–1289.

Caetano, Gregorio. 2019. "Neighborhood sorting and the value of public school quality." *Journal of Urban Economics* 114:103193.

Calsamglia, Caterina, Chao Fu, and Maia Güell. 2018. *Structural Estimation of a Model of School Choices: the Boston Mechanism vs. Its Alternatives.* Working Paper, Working Paper Series 24588. National Bureau of Economic Research, May.

Calsamiglia, Caterina, Chao Fu, and Maia Güell. 2020. "Structural Estimation of a Model of School Choices: The Boston Mechanism versus Its Alternatives." *Journal of Political Economy* 128 (2): 642–680.

Calsamiglia, Caterina, Francisco Martínez-Mora, and Antonio Miralles. 2020. "School Choice Design, Risk Aversion and Cardinal Segregation." *The Economic Journal* 131, no. 635 (August): 1081–1104.

Calsamiglia, Caterina, and Antonio Miralles. 2023. "CATCHMENT AREAS, STRATIFICATION, AND ACCESS TO BETTER SCHOOLS." *International Economic Review* n/a (n/a).

Cheshire, Paul, and Stephen Sheppard. 2004. "Capitalising the Value of Free Schools: The Impact of Supply Characteristics and Uncertainty*." *The Economic Journal* 114 (499): F397–F424.

Davidoff, Ian, and Andrew Leigh. 2008. "How Much do Public Schools Really Cost? Estimating the Relationship between House Prices and School Quality*." *Economic Record* 84 (265): 193–206.

De Fraja, Gianni, and Francisco Martínez-Mora. 2014. "The desegregating effect of school tracking." *Journal of Urban Economics* 80:164–177.

Epple, Dennis N., and Richard Romano. 2003. "Neighborhood Schools, Choice, and the Distribution of Educational Benefits." In *The Economics of School Choice,* by Caroline M. Hoxby, edited by Caroline M. Hoxby, 227–286. University of Chicago Press, January.

Fack, Gabrielle, and Julien Grenet. 2010. "When do better schools raise housing prices? Evidence from Paris public and private schools." *Journal of Public Economics* 94 (1): 59–77.

Fack, Gabrielle, Julien Grenet, and Yinghua He. 2019. "Beyond Truth-Telling: Preference Estimation with Centralized School Choice and College Admissions." *American Economic Review* 109, no. 4 (April): 1486–1529.

Ferreyra, Maria Marta. 2007. "Estimating the Effects of Private School Vouchers in Multidistrict Economies." *American Economic Review* 97, no. 3 (June): 789–817.

Gibbons, Stephen, and Stephen Machin. 2008. "Valuing school quality, better transport, and lower crime: evidence from house prices." *Oxford Review of Economic Policy* 24, no. 1 (April): 99–119.

Gibbons, Stephen, Stephen Machin, and Olmo Silva. 2013. "Valuing school quality using boundary discontinuities." *Journal of Urban Economics* 75:15–28.

Glazerman, Steven, and Dallas Dotter. 2017. "Market Signals: Evidence on the Determinants and Consequences of School Choice From a Citywide Lottery." *Educational Evaluation and Policy Analysis* 39 (4): 593–619.

Greaves, Ellen. 2023. "Segregation by choice? School choice and segregation in England." *Education Economics* 0 (0): 1–19.

Greaves, Ellen, and Iftikhar Hussain. 2023. "Symmetric school choice patterns by income: evidence from nationwide school quality information." *mimeo.*

Harjunen, Oskari, Mika Kortelainen, and Tuukka Saarimaa. 2018. "Best Education Money Can Buy? Capitalization of School Quality in Finland." *CESifo Economic Studies* 64, no. 2 (January): 150–175.

Harris, Douglas N, and Matthew F Larsen. 2019. *The identification of schooling preferences: Methods and evidence from post-Katrina New Orleans.* Technical report. Technical report, Education Research Alliance for New Orleans, URL https ...

Hastings, Justine, Thomas J Kane, and Douglas O Staiger. 2009. "Heterogeneous preferences and the efficacy of public school choice." *NBER Working Paper 2145.*

Hoxby, Caroline M. 2000. "Does Competition among Public Schools Benefit Students and Taxpayers?" *American Economic Review* 90, no. 5 (December): 1209–1238.

Lee, Yong Suk. 2015. "School districting and the origins of residential land price inequality." *Journal of Housing Economics* 28:1–17.

Machin, Stephen, and Kjell G Salvanes. 2016. "Valuing School Quality via a School Choice Reform." *Scandinavian Journal of Economics* 118 (1): 3–24.

Nechyba, Thomas J. 2000. "Mobility, Targeting, and Private-School Vouchers." *The American Economic Review* 90 (1): 130–146.

Nguyen-Hoang, Phuong, and John Yinger. 2011. "The capitalization of school quality into house values: A review." *Journal of Housing Economics* 20 (1): 30–48.

Park, Minseon, and Dong Woo Hahm. 2023. *Location choice, commuting, and school choice.* Technical report.

Pietrabissa, Giorgio. 2023. *School Access and City Structure.* Technical report.

Reback, Randall. 2005. "House prices and the provision of local public services: capitalization under school choice programs." *Journal of Urban Economics* 57 (2): 275–301.

Rehm, Michael, and Olga Filippova. 2008. "The impact of geographically defined school zones on house prices in New Zealand." *International Journal of Housing Markets and Analysis* (4): 313–336.

Ries, John, and Tsur Somerville. 2010. "School Quality and Residential Property Values: Evidence from Vancouver Rezoning." *The Review of Economics and Statistics* 92, no. 4 (November): 928–944.

Rothstein, Jesse. 2007. "Does Competition Among Public Schools Benefit Students and Taxpayers? Comment." *American Economic Review* 97, no. 5 (December): 2026–2037.

Ruijs, Nienke, and Hessel Oosterbeek. 2019. "School Choice in Amsterdam: Which Schools are Chosen When School Choice is Free?" *Education Finance and Policy* 14 (1): 1–30.

Sibieta, Luke. 2015. "School Spending in England 2010–15." *Fiscal Studies* 36 (3): 283–302.

Trajkovski, Samantha, Jeffrey Zabel, and Amy Ellen Schwartz. 2021. "Do school buses make school choice work?" *Regional Science and Urban Economics* 86:103607.

Walker, Ian, and Matthew Weldon. 2020. *School choice, admission, and equity of access: Comparing the relative access to good schools in England.* Technical report.

9. Tables

Table 1: Descriptive statistics for two neighboring secondary schools and school zones

	City		*AA* school		*BA* school	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Qualifications						
% Degree + (level 4 +)	31.87	17.7	61.03	1.52	57.94	1.74
Income and occupation						
% Professional occupation	38.4	17.43	69.71	5.13	57.08	2.7
% Routine occupation	23.95	13.22	5.7	1.15	9.65	2
Low income score	25.51	15.8	8.61	2.15	15.09	5.29
Household composition						
% Dependent child(ren)	43.58	10.26	49.57	6.32	45.25	3.22
% One dependent child	20.3	5.02	16.37	1.59	19.41	2.21
% Two+ dependent children	23.28	7.25	33.19	5.43	25.84	4.02
Number of children						
Age 1	23.38	12.19	18.2	4.95	24.61	9.53
Age 4	20.53	10.25	20.01	5.39	20.86	6.12
Age 10	15.87	8.33	21.22	5.94	13.91	5.77
Age 16	17.06	8.98	18.41	5.49	11.68	3.19
Property prices (£1000s)						
Median price	192.27	88.44	344.50	77.47	251.94	53.42
Median price, 5 room	201.3967	91.35	305.60	34.01	261.50	23.80
Median price, 6 room	247.2907	106.48	374.29	27.13	312.25	37.42
School performance (2011)						
% 5A*-C (inc. English & Math)	53		83		50	
% EBACC	14		50		17	
% Value-Added	999.9		1023.7		1002.8	
% FSM	22		4		24	
School access						
Implied P(admission) to *AA*			0.81		0.00	

Notes: All characteristics are measured at the lower level super output area (LSOA). Columns show the average and standard deviation across LSOAs, weighted by the number of households aged 25 to 65. Columns 1-2 show the average and standard deviation across all LSOAs in the city. Columns 3-4 (5-6) show the equivalent for the 11 (6) LSOAs in the *AA* (*BA*) school zone. '5A*-C' is the percentage of pupils that achieve at least 5 GCSEs at high grades (A*-C) including English and mathematics. 'EBACC' is the percentage of pupils that achieve the English Baccalaureate, which requires at least 5 A*-C grades in English, mathematics, two sciences, a foreign language and history or geography at GCSE level. Value-Added is the average progress made by pupils at the school from the end of primary school to the end of secondary school. 'FSM' is the percentage of pupils eligible for Free School Meals. Occupation classifications are the National Statistics Socio-Economic Classification of the household reference person. Low income score is the 2010 Index of Multiple Deprivation (Income Domain). Property prices are taken from the universal Land Registry Database and aggregated to LSOA level. The number of habitable rooms is merged in from the record of Energy Performance Certificates. The implied probability of admission to *AA* is the calculated probability of admission for non-sibling applicants.

Table 2: State variables and equilibrium values under geographic admissions used as inputs into the model

Variable	Value
# dwellings in H:	4,899
# dwellings in L:	3,307
# school places in AA	216
# school places in BA	160
# siblings with priority at AA	80.0
π_H	0.81
π_L	0.00
Share f_0	0.30
Share f_1	0.20
Share f_2	0.50
Share f_1 to f_2	0.26
Rent in H	£1,459.17
Rent in L	£1,217.33

Note: The total number of households is derived from the total number of people aged 25 to 65, divided by two, as assigned to couples. The number of school places in AA is taken from the school website. The number of school places in BA is calculated so that the number of pupils equals the total number of seats. The number of siblings with priority at AA is inferred from the number of children at the school choice age, accounting for the share of those children with an older sibling, and the probability that older sibling was admitted to AA. The probability of admission to AA from H is the number of uncertain applicants from H divided by the number of seats at AA after the sibling priority is accounted for. The probability of admission to AA from L is assumed to be zero, as there is excess demand from within H. The share of households of fertility type f is the total count of household of type f across both neighborhoods, divided by the total count of households. The derivation of fertility type is described in Appendix B. The share of households of fertility type f_1 relative to f_2 is derived from the total number of households in life-stage 0 that are f_1 relative to the total number of parent households in life-stage 0. This figure of 0.26 is similar to the national average of 0.22. The rent in H and L, and therefore the difference, is calculated by converting the median property price paid for 6 room homes into monthly mortgage payments. See Appendix B for more details.

Table 3: Estimates of preference parameters

		Coefficient estimates	Standard errors
Neighborhood amenities	α_H	0.2307	0.0198
AA school	α_S	0.2363	0.0249
AA school (second child)	α_2	0.7230	0.0410
Commute to school	α_C	-0.1362	0.0178
Moving cost f_0	α_m^0	0.6113	0.0780
Moving cost f_1	α_m^1	9.8625	1.1554
Moving cost f_2	α_m^2	13.6451	2.3572

Note: Standard errors are obtained with 100 bootstrap replications. α_S is the household utility derived from a child from f_1 attending the *AA* school. $\alpha_S(1+\alpha_2)$ is the household utility derived from two children from f_2 attending the *AA* school.

Table 4: Empirical and model moments

	Empirical moments			Model moments		
	t_0	t_1	t_2	t_0	t_1	t_2
% f_0	0.44	0.46	0.58	0.54	0.54	0.54
% f_1	0.55	0.66	0.64	0.49	0.64	0.64
% f_2	0.67	0.76	0.74	0.57	0.75	0.69

Note: Empirical moments are derived from area-level statistics from the 2011 Census (Nomis). Model moments are found through estimation.

Table 5: Simulated moments

	Geographic	Lottery
r_H	1,459	1,448
Probability of admission (*AA* from *H*)	0.81	0.59
Probability of admission (*AA* from *L*)	0.0	0.59
Mean income in *H*	3,878	3,873
Mean income in *L*	1,907	1,904
Mean income in *AA*	4,484	3,560
Mean income in *BA*	3,004	3,560
f_{0H}	0.269	0.274
f_{1H}	0.194	0.195
f_{2H}	0.537	0.532
t_{0H}	0.340	0.346
t_{1H}	0.055	0.052
t_{2H}	0.210	0.209
t_{3H}	0.394	0.392

Note: This table shows the equilibrium outcomes under two alternative school admissions systems. First, the geographic system, where pupils within the high quality neighborhood have priority in admission to the Above Average (*AA*) school. These equilibrium outcomes are derived using the estimated model parameters. r_H denotes the rent in the high quality neighborhood, that is estimated to match the observed premium in the data. The probability of admission to the *AA* school is also estimated to match the observed patterns in the data. f_{0H}, f_{1H} and f_{2H} denote the share of households who choose to live in H with completed fertility size of zero, one and two. t_{0H}, t_{1H}, t_{2H} and t_{3H} are the share of households that choose to live in H in each life-stage.

A. Proofs

A.1. Proof of Lemma 1

Under the lottery admissions system:

- The number of students in the AA school is:

$$S + \pi_H \cdot A_H + \pi_L \cdot A_L = S + A \cdot \frac{P_{AA} - S}{A} = P_{AA}$$

- The AA school fills all its places, and therefore so does the BA school, since as a whole the two schools have just enough places for the children $A + S$ living in the two neighborhoods.

Under the geographic admissions system:

- The number of students in the AA school is:

$$S + \pi_H \cdot A_H + \pi_L \cdot A_L$$

- In the case where $\pi_H = 1$ and $\pi_L > 0$, this number is:

$$S + A_H + \frac{P_{AA} - S - A_H}{A_L} \cdot A_L = S + A_H + P_{AA} - S - A_H = P_{AA}$$

  This means that the AA school fills up and so does the BA school.

- In the case where $\pi_H < 1$ and $\pi_L = 0$, the number of students in the AA school is:

$$S + \pi_H \cdot A_H = S + \frac{P_{AA} - S}{A_H} \cdot A_H = P_{AA}$$

  This means that the AA school fills up and so does the BA school.

In both cases and with both admissions systems, the 'market' for school places clears. QED.

A.2. Proof of Lemma 2

The specification of the utility function in equation (5) shows that the marginal utility of consumption is strictly decreasing. In contrast, the marginal utility of amenities, school quality and travel to school are all independent of consumption. Hence, there exists a unique threshold $R(\mathbf{r},\pi)$ at which households would give up an amount of consumption equal to the difference in rent in $\mathbf{r}$ in order to gain the higher probability π of accessing the AA school. Since life-time values are linear combinations of instantaneous utilities, this threshold property applies to lifetime values too. This threshold depends on the fertility type f and life-stage t and is denoted $R_{f,t}(\mathbf{r},\pi)$.

A.3. Proof of proposition 1

.

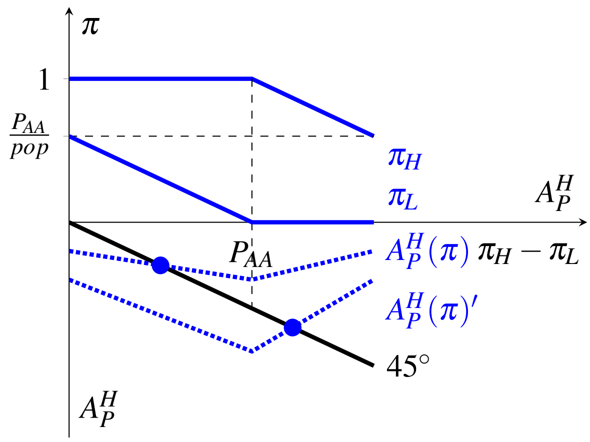

Figure B.1: Equilibrium existence and uniqueness

Figure B.1 shows the unique equilibrium in school and housing markets. The top panel of Figure B.1 shows the probability of access to AA from H and L, as a function of the number of school applicants in H. The bottom panel shows the number of school applicants that choose to live in H as a function of the difference in the probability of accessing AA from H and L. Equilibrium is on the 45° line, where the number of applicants choosing to live in H as a function of the difference in probabilities yields the same probabilities of admission. To expand, first consider the top panel:

- A_p^H is the number of parent households living in H that apply to the AA school each year. pop denotes the whole population of a cohort of children. P_{AA} is the number of school places in AA.

- When there are fewer children in the school choice cohort living in H than the number of school places (P_{AA}), the probability of entering AA from H is one.
- After the point where the number of children in the school choice cohort equals the number of school places in AA, the probability of entry from H declines.
- The probability of entering AA from H is lowest when all children in the school choice cohort live in H: $\pi_H = P_{AA}/pop$
- The probability of entry to AA from L mirrors the probability of entry to AA from H.
- This is positive while there are fewer children in the school choice cohort in H than there are school places, and zero thereafter.
- The biggest difference in the probability from H and L is at the point where the number of children in the school choice cohort in H equals the number of school places in AA.

Next, consider the lower panel:

- The number of parent households living in H (A_p^H) is an increasing function of the difference in probability of admission to AA from H and L.
- The school market will lead to the highest equilibrium demand in H at the point where the difference in probability of admission to AA between H and L is largest.
- Equilibrium is at the 45° line, where the probability of admission implied by the number of parent households choosing to live in H equals the number of parent households choosing to live in H as a function of the probability of admission.
- The two dotted lines in the bottom panel show alternative scenarios where parent households are more and less responsive to the probability of admission to AA.

The supply of dwellings in H is fixed at D_H. Since the total demand for dwellings from parents and non-parents decreases with r_H, this defines a unique rent level $\tilde{r_H}$ at which the housing market clears in both neighborhoods. Note that since the set of the two neighborhoods comprises as many properties as households, when one market clears, the other does too.

A.4. Proof of proposition 2

.

We provide the proof for the static version of the model. In the dynamic version, decisions are also influenced by the anticipation of moving costs, of cost of travel to school and of becoming a parent.

At equilibrium, the housing market in H clears if and only if:

- For the marginal non-parent household, the disutility of loosing an amount of consumption equal to the difference in rent equals the utility gain from amenities a_H.
- For the marginal parent household, the disutility of loosing an amount of consumption equal to the difference in rent Δr equals the utility gain from amenities α_a plus the gain from higher school quality α_s.

The marginal non-parent is therefore richer than the marginal parent, since their marginal utility of consumption is higher.

$$\widetilde{y_P}(\alpha_a + \alpha_s, \Delta r) < \widetilde{y_{NP}}(\alpha_a, \Delta r)$$

Given that the marginal non-parent is richer than the marginal non-parent, there would be fewer non-parents in H than parents *if* parents and non-parent were in equal numbers in the population and had the same income distribution. This observation would be reversed if either:

- There are sufficiently more non-parent than parent households in the population. In this case, there may be more non-parent households in H, despite the higher threshold for the marginal household.
- Non-parent households have sufficiently higher income than parent households. In this case, then the income threshold will correspond to a lower quantile in the income distribution, and therefore more non-parent households choosing H.

Market equilibrium is such that the demand for houses in H exactly equals the number of non-parent and parent households that choose to live in H:

$$D_H = N_{NP} \cdot [1 - q_{NP}(\widetilde{y_{NP}}(\alpha_a, \Delta r))] + N_P \cdot [1 - q_P(\widetilde{y_P}(\alpha_a + \alpha_s, \Delta r))]$$

Where $q(.)$ is the group-specific quantile associated with the income threshold of the marginal household.

Figure B.2 shows the demand for dwellings in H for parent, P, and non-parent, NP households. Note that the y-axis is demand in H for parents, and the total number of properties in H minus demand from non-parents. The demand curve for non-parents is therefore inverted. The orange lines represent an initial equilibrium under parents' valuation of school amenities ε. The equilibrium is where $r_H - r_L = \Delta r^*(\varepsilon)$, where demand from parents and non-parents equals the total number of dwellings in H. Now, consider a rise in the valuation of school quality by parents, to ε^1. This shifts the demand curve for parents, shown by the blue line, but not for non-parents. The new equilibrium difference in rents $r_H - r_L$ is at $\Delta r^*(\varepsilon^1)$. Prices rise to discourage enough parent households from choosing to live in H. As some non-parents are also displaced from H as the price rises, this allows more parent households to live in H in the new equilibrium, relative to an economy without non-parents. If the economy consisted only of parents, then the price rise would need to be larger to discourage more households from living in H, so that supply equals demand.

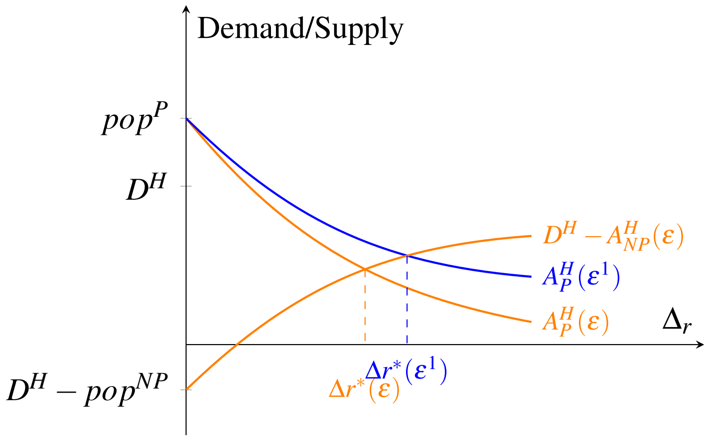

Figure B.2: Housing market in H: Equilibrium rent as value from school increases

Figure B.3 considers a related scenario in which the value of amenities increases. At the first equilibrium, there is no difference in amenities across neighborhoods. The equilibrium price in H is greater than zero, as parents value H for access to the AA school. The equilibrium price must be sufficiently high to discourage enough parents from living in H so that demand equals supply. All non-parents therefore choose to live in L, as they have the same utility from amenities at a lower rental price.

As the value of amenities in H increases to ε^1, parent households compete with non-parent households to live in H. This increases the equilibrium rent, and decreases the number of parent households in H.

As the value of amenities in H increases to ε^2, there is higher demand from both parent and non-parent households to live in H, and the equilibrium rent in H necessarily increases. The overall effect on the number of parent to non-parent households is ambiguous, however, as it depends on the shift of the demand curve for both household types.

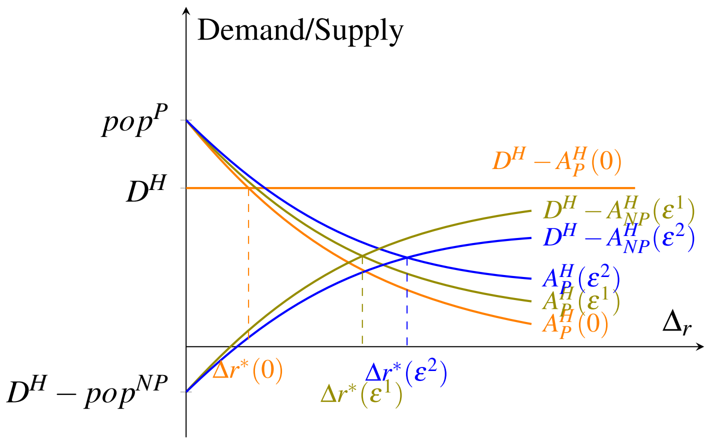

Figure B.3: Housing market in H: Equilibrium rent as amenities increase

A.5. Proof of proposition 3

.

Figure B.4 shows the change in equilibrium probability of admission and demand from parents as the school admissions criteria changes from geographic to lottery.

- Under geographic admissions, the intuition from proposition 1 holds. The probability of admission to AA from H and L affects the valuation of H for parent households, which in turn affects the number of parent households living in H.

- Under lottery admissions, the probability of admission to AA from H and L is equal. This reduces the utility from H for parent households, which:
  - Increases the income threshold of the marginal parent household to live in H.

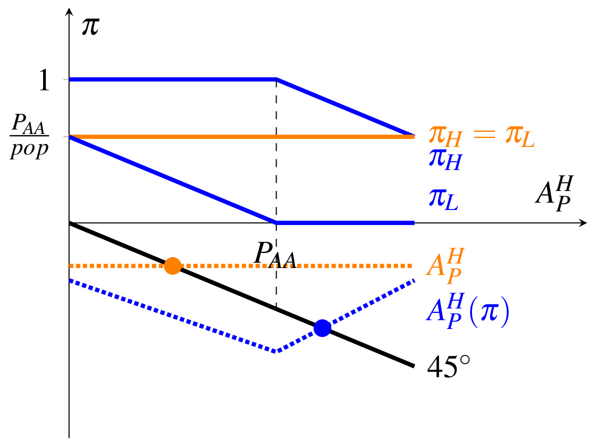

Figure B.4: Equilibrium with lottery admissions

  - Reduces the number of parent households in H in equilibrium.
  - Reduces the rent in H relative to L as there is less demand.

- The difference in equilibrium outcomes is largest when the difference in admission probabilities under the geographic system between H and L is largest.

Note that, following proposition 2, this change in demand from parents will be partly dampened by non-parents.

A.6. Proof of proposition 4

.

An increase in the cost of travelling to school may affect either households who live in H and need to travel to the BA school in neighborhood L, or households who live in L and need to travel to the AA school in neighborhood H. Our description of the equilibrium shows that these two possibilities are exclusive: either $\pi_H = 1$ and $\pi_L > 0$ and we are in the latter case, or $\pi_H < 1$ and $\pi_L = 0$ and we are in the former case.

The effect of an increase in the costs of commuting to school depend on which case we start with, as shown in Figure B.5. In the first case, the only households who need to commute to school are households in neighborhood L travelling to the AA school. An increase in commuting cost therefore makes living in L less attractive. The demand for dwellings in H by parents increases, pushing the rent in H up. The new equilibrium has more parent households living in H and a higher rent in H.

On the other hand, if we start from an equilibrium where only households residing in H commute to school (those who did not get a place in the AA school and commute to the BA school), an increase in commuting costs makes living in H relatively less attractive, therefore decreasing the demand for dwellings in H by parent households. The new equilibrium with be at a lower rent in H and with fewer parent households choosing to live in H.

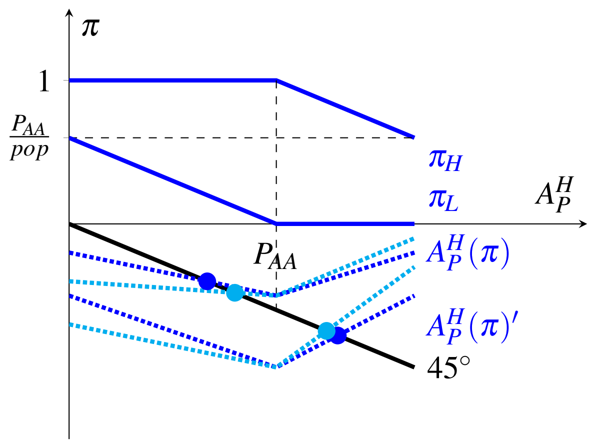

Figure B.5: Equilibrium with increase in travel costs

Last point in the proposition: because when A_H^P is close to P_{AA}, the initial $\pi_H - \pi_L$ is at its greatest.

B. Data Appendix

This appendix provides detail on all the sources of data used for the stylized facts and estimation of the model.

B.1. School quality

Publicly available information on school performance measures is downloaded from the 'Find and compare schools in England' website, for the 2010/2011 academic year. We use the most common school performance metrics for schools in this time period: the percentage of pupils that achieve at least 5 GCSE grades at A* to C (including English and Math) and the percentage of pupils that achieve the English Baccalaureate (5+ A*-C grades in English, mathematics, two sciences, a foreign language and history or geography). We reiterate that these measures of 'quality' reflect both teaching/management quality and peer composition.

B.2. Small area characteristics (Census 2011)

Publicly available information from the 2011 Census at the Lower Level Super Output Area (LSOA) from Nomis (official labour market statistics).

Information used for the stylized facts are:

- % households with 'professional' and 'routine' occupation, taken from dataset QS608EW.
- % households with 'one' or 'two plus' dependent children, taken from dataset QS118EW.
- The number of children with each age, taken from dataset QS103EW.

Information used for the stylized facts and calculating the empirical moments is finer detail about the number of dependent children by age of the youngest child, taken from dataset QS118EW.

B.3. Calculating the empirical moments

Our empirical moments are the proportion of households of each fertility type f and life-stage t who choose to live in H (from the total number in H and L). This process is convoluted, as the area-level data available does not contain precise information on parents' ages and number/age of dependent children together. We therefore use the age of dependent children to infer the households' age/life-stage, based on our model's assumptions. These assumptions are that the first birth occurs at age 30, subsequent children are spaced two years apart, and there is a steady state. The life-stages are from 25 to 39 (t_0), 40 to 41 (t_1), 42 to 49 (t_2), then 50 to 65 (t_3). Under these assumptions, we can calculate the number of households of each fertility type and life-stage from the number of households with the following dependent children (which are the categories available in the data):

1. One dependent child, aged 0-4

2. One dependent child, aged 5-11
3. One dependent child, aged 12-18
4. Two dependent children, youngest aged 0-4
5. Two dependent children, youngest aged 5-11
6. Two dependent children, youngest aged 12-18
7. Three+ dependent children, youngest aged 0-4
8. Three+ dependent children, youngest aged 5-11
9. Three+ dependent children, youngest aged 12-18

To illustrate how we convert these numbers into the number of households of each fertility type and life-stage, first take the simplest case of households with one dependent child. Starting from the households that have one dependent child aged 12-18, these households are assigned to f_1 in period t_2 (secondary school age). Of households with one dependent child aged 5-11, 5/7 are assigned to t_0 while 2/7 are assigned to t_1.[37] Assigning the households with one child aged 0-4 is more complicated, as some of these households will have another child in the future. Under our assumption of two-year spacing between children, households with an only child of 3 and 4 are assumed to have reached completed fertility. We therefore assign 2/5 (corresponding to ages 3 and 4) of the households with one dependent child aged 0-4 to f_1 in period t_0, assuming that these households have completed fertility. Of the other 3/5 (corresponding to ages 0, 1, 2), some will have completed fertility (and so be f_1), while others will go on to have another child (and so be f_2). To calculate this, we use the national proportion of households with dependent children that ever have one child (0.22). The total number of households that are f_1 in t_0 is therefore 0.22*(3/5). The remaining households ((3/5)*0.78) are assigned to f_2 in t_0.

37. These fractions are based on the fraction of the life-stage consistent with our model. Life-stage t_0 is when the eldest child is between 0 and 9, for example, so contains five of the seven years (5, 6, 7, 8, 9) of the child being between five and eleven. The remaining two (10 and 11) are consistent with life-stage t_1, so 2/5 are assigned to this stage.

For households with two or more children, the process is similar. The difference is we assume a two-year age gap, and that households' life-stages are defined according to the age of the oldest child. For example, for the category of two dependent children where the youngest child is 5-11, the oldest child is assumed to be 7 to 13. Ages 7, 8 and 9 would be assigned to life-stage t_0 (3/7) and ages 10 to 11 would be assigned to t_1 (2/7) and ages 12 and 13 would be assigned to t_2. For households with two and three children, the same adjustment for the youngest category is not needed, as the household is already observed to be type f_2.[38]

This process calculates the number of f_1 and f_2 households in each life-stage. The number of f_0 in each life-stage is calculated as the total number of households in that life-stage, minus the number of f_1 and f_2 households.

The empirical moments are then the total count of households of each fertility type and life-stage in H, divided by the total number in H and L.

B.4. Income growth

We calculate income growth by completed fertility type using nationally representative longitudinal data, the British Household Panel Study and Understanding Society, spanning 29 years. Our measure of income is household net income, converted to real terms using the Consumer Prices Index and shown in 2014 values.

Measuring income growth over time requires households to be observed at the correct ages (25 to 65 to be consistent with our model) and have a reliable measure of completed fertility. This measure is constructed as follows. In each wave, the number of dependent children currently in the household is recorded. The surveys also ask questions about the number of children ever had, which we use as 'completed' fertility post age 40. Our final measure of fertility type is primarily based on the answers of households aged 40 or over. Where this is not observable (perhaps as households enter the survey at a young age) we instead use the maximum number of dependent children ever observed in the household.

Our sample of households is:

- All households with age between 25 and 65.

38. Note that this process is not possible for the final life-stage, as our model assumes that children leave the household in t_3, and in practice, that will be true for many households with children of older ages. Our empirical moments therefore do not include the share of each fertility type that chooses to live in H in t_3.

Table A.1: Income growth by fertility type measured in UK household panel data

	Household net monthly income growth (%)			Parameters of log-normal distribution	
	t_0	t_1	t_2	μ	σ
f_0	0.012	-0.002	-0.010	7.76	0.51
f_1	0.024	0.011	0.005	7.55	0.53
f_2	0.017	0.010	0.006	7.67	0.52

Source: British Household Panel Survey and the UK Household Longitudinal Study.

- Where the household is observed after age 40 and answers a question about completed fertility, or is observed at least once in their 30s and 40s with the current presence of children.
- Where income is observed in consecutive waves.
- Where the household is a couple in consecutive waves.
- Where the household is outside London.

The income growth patterns are shown in Table A.1. These are translated into log-normal distributions for each family type, truncated at $0.5r_L$, fitted to the observed distributions (25th, 50th and 75th percentiles).

B.5. Low income score

Low income score is the 2010 Index of Multiple Deprivation (Income Domain), which classifies small areas according to the proportion of the population in an area experiencing deprivation according to low income. This measure captures variation only at the bottom of the income distribution, so is useful for stylized facts but not for estimation of the model, which requires information on income types across the distribution. The 2010 Index of Multiple Deprivation is linked to the 2011 LSOAs by the Consumer Research Data Centre, which we use.

B.6. Property prices

Property prices are taken from the Price Paid Data from HM Land Registry, which covers all property sales in England and Wales that are sold for full market value and are lodged with HM Land Registry for registration. We download the single file (including prices paid between 1 January 1995 to the most current monthly data) and extract all prices paid in 2011. The dataset includes information on the transaction price, type (Detached, Semi-detached, Terrace or Flat/Maisonette), whether the property is newly built and whether the property is freehold or leasehold. Full address information is provided, including the Primary Addressable Object Name (PAON), Secondary Addressable Object Name (SAON) and postcode. The data was merged to Lower Level Super Output Area (LSOA) by postcode using the National Statistics Postcode Lookup for August 2011. 99.27% of postcodes match between the price paid data and this postcode directory.

1.76% of addresses have multiple sales recorded in 2011. For these properties we take the average price. Median prices at the LSOA level are calculated by collapsing the postcode to LSOA. These prices do not take into account the number of rooms in the property. For this information, we use data from the Energy Performance Certificate database, which is described next.

B.7. Property characteristics

The Energy Performance of Buildings Data: England and Wales (EPC) is a database of all Energy Performance Certificates which is provided under Open Government Licence v3.0. Relevant for our purposes, the data contains information on the number of rooms. For most properties, 'address1' in the EPC corresponds to PAON in the prices paid data, and 'address2' corresponds to SAON. This is not the case for flats (or apartments or units), however. The cleaning process is therefore:

- Replacing the SAON with information in the field 'address1' if the property is a flat (for example, 'flat 9').
- Removing information from the PAON now contained in SAON if the property is a flat.
- For properties with a house number, retain the house number only.

87.65% of properties in the city merge perfectly using PAON, SAON and postcode, and so have a complete record of price paid and property characteristics.[39] These properties are used to estimate the price residual for properties across LSOAs, which is explained in further detail below.

B.8. Rents across neighborhoods

We calculate the median property price in H and L, controlling for differences in the number of rooms across neighborhoods. To do this, we use the sample of properties (87.65% of all properties) with a perfect match between the price paid and EPC data, described above, which contains the number of habitable rooms. The modal number of habitable rooms in H and L is six (usually three bedrooms). In effect, we compare the median prices of six room homes in H and L, which are £312,250 in L and £374,290 in H, 19.87% higher (see Table 1).

The 19% difference in property prices between H and L is converted into monthly rents by using a mortgage calculator, assuming a 20% deposit, a 25-year mortgage, and a 3.25% interest rate.[40] We start from the median property price in L (£312,250). This is converted to monthly mortgage/rental value using the assumptions above, which gives £1,217.33 per month. The rent in H is calculated as 19.87% higher than this (based on the difference in median prices of six bedroom properties) which is £1459.17.

39. Properties without a perfect match are then matched by type (flat/house), PAON and postcode (4.23%), type, SAON and postcode (0.17%), type and postcode (6.75%), and finally type and postcode sector (1.19%).

40. The interest rate of 3.25% is around the annual mean in 2011 (FRED Economic Data).