School Choice and Neighbourhood Sorting
WORK IN PROGRESS

Ellen Greaves[1] and Hélène Turon[2]

[1]University of Bristol and Institute for Fiscal Studies
[2]University of Bristol

February 1, 2019

Abstract

We present the first dynamic model of neighbourhood and school choice to illustrate how the design of school priorities - how the school allocates places if over-subscribed - affects sorting into schools and neighbourhoods. Our model shows not only how the design of school choice affects local house prices and school composition, but also the composition of neighbourhoods in household age and completed fertility size. We focus on the comparison of over-subscription priorities between random allocation and catchment areas that give priority to local pupils. We calibrate our model to a city in England, where stylised facts match the key predictions of our model. We find that allocating places by lottery if over-subscribed increases the probability of attending the "Good" school from outside the catchment area, decrease local house prices in this area, while increasing the mixing between household types.

1 Introduction

One objective of providing "school choice", where households can submit preferences for a number of schools, is to widen access to good schools to pupils from a greater range of socio-economic backgrounds.[1] In the alternative "neighbourhood" or "catchment area" system, where schools' intakes only comprise of local children, households' residential and school choices are bound. This yields house price inflation around the highest quality schools and inequitable outcomes in that less well-off households can not afford access to these schools. Under school choice, households are able to apply to other schools than their local one, which partly breaks the deterministic school assignment from location. Over-subscribed schools still need to ration applications to available places, however, through an order of priority of applicants.[2]

[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]In most school choice systems, a central authority then assigns pupils to schools using an allocation mechanism, taking into account parents' stated preferences for schools and schools' ranking of pupils. The focus in our model is on the impact of the school priority order rather than on the allocation mechanism (e.g. Boston or Gale-Shapley) since we only have two schools which can accommodate the total demand in the aggregate.

This paper analyses the impact of school priorities on the sorting of households and pupils across schools and neighbourhoods. In other words, we quantify the relationship between the type of school priority (catchment area versus lottery) and the social mix within schools and within neighbourhoods, defined as the mix between households of different ages, different income and different family size. 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 family types along completed family size.[3] The dynamic components of our framework come from the sibling priority rule applied by schools, whereby the family's younger sibling is guaranteed a place in the school that the older sibling attends, and from the existence of moving costs. Households care for local amenities, school quality and distance to school when there are children travelling to school. Residential location matters most in the life stage when parents apply to a secondary school (if there is neighbourhood sorting) and in the life stage when the household's children travel to school. Because households are forward looking and there is a cost of moving across neighbourhoods, households' residential choices exhibit some persistence.

We calibrate our model to data from three neighbourhoods in the city of Bristol, England, and find that it replicates patterns of sorting in schools and neighbourhood across family types and life stages well. Our results also provide structural estimates of households' willingness to pay for a good school and neighbourhood amenities along the life-cycle and across family size.

A structural model is the appropriate method for our research question for four reasons. First, and crucially, it allows us to simulate counterfactual policy environments. For example, how do choices change if all schools admit pupils by lottery if over-subscribed? The alternative reduced form approach would be to find areas where the school priorities had (exogenously) changed and observe the resulting change in neighbourhood and school sorting, where external validity would be limited. Second, structural models illuminate the mechanisms through which a policy change (here school priorities) influence the outcome of interest (here neighbourhood and school sorting). Our model quantifies the impact on house prices and the overall welfare effect of a change in school priority ordering. Third, we are able to uncover the structural parameters of the underlying theoretical model, i.e. the willingness-to-pay for neighbourhood amenities and the willingness-to-pay for school quality, thereby decomposing the local premium in house prices into these two components. Finally, we make households' choices and incentives explicit, as well as any functional form assumptions and elements that are *excluded* from the model.

1.1 Previous literature

Empirical

There is a large empirical literature, beginning with Black (1999), on the willingness to pay for local school quality. Using boundary discontinuities Black finds that house prices respond to local school quality. Typically, households are willing to pay a premium of around 3-4% for access to a one standard deviation increase in school average test scores. (See Gibbons and Machin (2008), Black and Machin (2011) and Nguyen-Hoang and Yinger (2011) for detailed summaries of recent evidence.) Bayer, Ferreira, and McMillan (2007) aim to disentangle household preferences for school and neighbourhood attributes, while taking into account endogenous sorting across neighbourhoods.

[3] We assume that young households have perfect foresight over their future completed fertility.

They show that previous estimates of willingness to pay for higher school quality are upward biased if households also have preferences for more affluent neighbours.

Our research question is whether the house price premium (and sorting into neighbourhoods and schools) is affected by whether and to what extent location determines access to a good school. Empirical evidence is limited to the effect of admission to local schools on house prices. Machin and Salvanes (2016) use a 1997 reform in Olso county in Norway to estimate the change in willingness-to-pay for houses close to good schools. This reform changed schools' enrolment rule from strict neighbourhood enrolment to an allocation by prior average grade. They estimate a fall in the house price premium near a good school by 50 percent. Bogart and Cromwell (2000) use a change in school districts in Ohio in 1987 and find that house prices respond not just to school quality, but whether the area has a neighbourhood school, and the ease of transportation to school. Ries and Somerville (2010) use a re-zoning reform in Vancouver in 2001 and find that house prices respond to changes in school quality only in the top quartile of the distribution.

These empirical studies provide useful evidence that house prices respond to *how* places at schools are determined, not only the quality of the school, but are silent on the underlying mechanisms and impact on sorting and segregation. The external validity of these studies is of course limited to the specific context of the policy change.

Another strand of literature studies the relationship between neighbourhood and school segregation. In a qualitative study, Coldron et al. (2010) find a strong correlation between residential segregation and school segregation, concluding that residential segregation is the main determinant of sorting into schools in England. Taylor and Gorard (2001) also conclude that "schools remain socially divided chiefly because of the socially divided nature of housing" rather than school choice policies, and recognise the potential for the use of catchment areas and distance to school as criteria for admission to increase social segregation in schools. Johnston et al. (2006) find that the level of segregation in schools is in fact markedly higher than in neighbourhoods in some areas of England when considering the distribution of ethnic groups, although this is not a focus of our paper.

Theoretical

There is existing theoretical research that incorporates the endogenous location of households in response to the design of school choice, and/or explicitly models changes in priority in admissions. We summarise the key ingredients and findings of each paper in turn before highlighting the contributions of our paper.

Nechyba (2000) presents a three-district model of high, middle and low income school districts. The focus of this paper is on the impact of targeting vouchers for private schools on certain school districts or on low-income households. In the model, households have one child, whose ability is correlated with household income and can choose their place of residence, vote on local property tax rate and opt for private schooling. School quality is a function of spending and average peer ability. When calibrated to New York data, the model predicts that the nature of targeting affects the outcome of the vouchers policy substantially and that households' residential mobility plays a key role in the workings of this policy.

Epple and Romano (2003) compare student sorting between two systems of enrolment: one neighbourhood enrolment and one district-wide open enrolment. Each household has one child whose ability is correlated (or not) with household income. They may face transport costs when attending a distant school. Households choose their neighbourhood of residence, school and vote over tax rates. School quality is endogenous. In the equilibrium with neighbourhood enrolment, the model predicts that income stratification implies school quality stratification because the access to neighbourhoods with better peer groups is rationed by higher housing prices. In the equilibrium with open enrolment and no transport costs, the model predicts equal school qualities and equal house prices. Finally, in the equilibrium with with open enrolment and transport costs, only higher income households are able to choose schools, while lower income households live and attend the school in the poorer neighbourhood. The difference in house prices is equal to the transport cost.

Ferreyra (2007) analyses the impact of private school vouchers targeted to non-religious schools relative to vouchers schemes for any private school, religious or not. The outcomes of interest are private school enrolment and location decisions. In the model, households have one child and heterogeneous preferences for the quality of the school and the neighbourhood. Tastes for school depend on households' religion. Households' residential and school choices, school quality and tax rates are all endogenous. The model, when calibrated to the Chicago metropolitan area, predicts that universal targeting of vouchers benefits more households and expands private school intake more than those targeted to non-religious schools.

Calsamiglia et al. (2015)'s research question is close to ours in that is assesses the impact of school allocation rules on socio-economic sorting into schools and neighbourhoods. The allocation mechanisms examined are Boston Mechanism and Deferred Acceptance, with or without neighbourhood priority. Households have one parent and one child and have types defined as either income or child ability. Their preferences are homogeneous and relate only to consumption and the child's human capital. House rents and school quality are endogenous. The main findings are that priority to local applicants lead to segregation, whatever the allocation mechanism. In the absence of neighbourhood priority, different allocation rules do not affect residential sorting but can have large effects on school sorting. Also, the availability of private school can substantially affect the equilibrium.

Calsamglia, Fu, and Güell (2018) examine the relative equity and efficiency of different allocation mechanisms in a context where the population of households comprises both strategic and non-strategic. The former take into account admission probabilities when applying for schools whereas the latter reveal their true preferences. Households also have the option of enrolling in a private school if they value their assigned school less than this outside option.

1.2 Contribution

Unlike all previous papers, we incorporate **different household types**, for example Non-Parents and households whose dependent children have left home. This unique aspect of our model allows us to explore the spillovers of the school choice system to Non-Parents in rents and neighbourhood composition, as well as calibrate the model using these additional parameters. These spillovers are important: only 27% of households have a dependent child in England. Our second unique model component is a **dynamic setting**. Households choose their neighbourhood depending on the environment (priority at the Good school determined by proximity or otherwise) and their life-stage in each period. This allows us to explore neighbourhood formation by age and household-type in addition to income. Due to this feature our model is also able to relate to detailed small area-level characteristics on the age and income composition of neighbourhoods in addition to sorting into schools. We also include sibling priority, which is only considered by Calsamglia, Fu, and Güell (2018) but is relevant in most international environments.

Only Epple and Romano (2003) model the effect of **transport costs** on neighbourhood and school sorting (as we do). This ingredient is realistic, modelling the demand for proximity, and allows us to simulate the effect of transport policies on neighbourhood and school sorting.

Endogenous location is an essential component of a theoretical model of sorting in response to school choice and school priorities. All theoretical papers summarised above have this ingredient, apart from Calsamglia, Fu, and Güell (2018), who leave this to later work.

One model component we currently omit is endogenous school quality, which is a common feature of other theoretical papers. In Calsamiglia, Martínez-Mora, and Miralles (2015) school quality is a function of the peer group ability, while in Nechyba (2000), Epple and Romano (2003) and Ferreyra (2007) school quality is a function of the peer group and school spending. In future work we plan to endogenise school quality as a function of the peer group but include a fixed component (for example due to management quality). As in Calsamiglia et al. (2015), future work will also consider the impact of a private school outside option. Appendix A sketches these extensions.

2 Model dynamic choice of neighbourhood and school

Our environment is composed of N forward-looking households choosing their place of residence and secondary school for their children (if they have any), N landlords renting their properties and S schools allocating their places to applicants. All agents have perfect information but are not able to coordinate, e.g. there is perfect competition in the property market.

Households are modelled over the lifecycle, which is decomposed into four stages T_t defined around the period that is key for our purposes, i.e. the time when households apply for a place in secondary school. This period is denoted T_1. The preceding 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 (if they have any). 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. For simplicity we will refer to all households as a set of one or two adults of the same age and who take decision in a unitary manner. We rule out divorce and remarriage.

We sketch in Appendix A two possible extensions of the model which we have yet to implement where school quality is endogenous and where households have the option to send their child(ren) to a private school.

In our stylised representation of the timing of these events, we assume that the duration of each period is 15, 2, 8 and 30 years, with adult life starting at the age of 25. For households with children, we assume the first child birth to occur at age 30 (and the second two years later), secondary school application to occur 10 years later, and children to leave secondary school when parents are 50 is 8 years. We assume life expectancy to be 80 years and the post-secondary school phase to last 30 years.

We allow households to differ along two dimensions: income and fertility type. The latter is represented by the completed family size denoted f, which can take the values 0, 1 or 2. We assume that households know their fertility type and income y from the start of their adult life. For now, income is assumed constant over the lifecyle; this assumption will be relaxed later on. Note that both these dimensions of household heterogeneity are exogenous.

We model an environment in steady-state, thus the fraction of the population of households in the life stage t and of fertility type f is constant and equal to $\tau_t \cdot \phi_f$, where τ_t is the fraction of total adult life spent in phase t and ϕ_f is the fraction of households of fertility type f.[4]

Households will choose to live in neighbourhood n_t in their life stage t. There is a set of two neighbourhoods to choose from $n = L$ (low quality) or $n = H$ (high quality) which offer different levels of amenities. There is one secondary school per neighbourhood, and school quality[5] is denoted s which can take the values B (bad) and G (good).

Households all rent their property at a yearly cost of r and have to pay a utility cost m to move across neighbourhoods. We rule out saving and borrowing so that households consume all their non-rent income. Preferences are assumed to be homogeneous across households, who value consumption, school quality, neighbourhood amenities and proximity to school in a separable manner. Our specification for the households' yearly utility function is the following:

$$\begin{aligned} U(f,t,y,s,n) \quad &= \quad \log(y-r) + \alpha n \\ &\quad\quad +s \cdot (\theta + \gamma \cdot \mathbb{1}_{(k=2)}) - m \cdot \mathbb{1}_{(n_t \neq n_{t-1})} && (1) \\ &\quad\quad -d \cdot \mathbb{1}_{(k>0)} \cdot \mathbb{1}_{(s=1\&n=L)|(s=0\&n=H)}, && (2) \end{aligned}$$

where $k = k(f,t)$ is the number of children of secondary school age in the current period, α is the utility derived from the amenities in the high-quality neighbourhood relative to the low-quality neighbourhood, θ (respectively $\theta+\gamma$ is the utility of having one (respectively two) child(ren) attending the good secondary school relatively to the bad secondary school and d is the disutility associated with having to commute to a secondary school in a different neighbourhood from the neighbourhood where the household resides. In addition, we have a term for the moving cost potentially incurred at the junction between two life phases: $-m \cdot \mathbb{1}_{(n_t \neq n_{t-1})}$.

In this framework, two markets need to clear: the rental property market and the "market" for school places. The former will clear thanks to the adjustment of the rent level r, the latter will clear thanks to the school allocation rule in case of excess demand. We assume that, across the two neighbourhoods, the numbers of properties and school places are such that the global market always clears.

Starting with the market for school places, we make the following assumptions: First, schools operate a "siblings priority" rule whereby younger siblings are guaranteed a place in the school of the elder child. Second, all households apply to the good school for their eldest child, i.e. the utility derived from the good school always exceeds the cost of commuting to a distant school. However, families with two children apply for their youngest child to the secondary school attended by their eldest child. We justify this assumption by conjecturing that commuting to two different schools incurs a cost that is larger than the benefit of having the youngest child attending the good school

[4]Note that we have $\sum_{t=0}^{3} \sum_{f=0}^{2} \tau_t \phi_f = 1$

[5]Measures of school quality are discussed in section 3.

when the youngest attends the bad school. Third, when a school is over-subscribed, it will use one of two allocation mechanisms to deal with the excess demand: priority to closer applicants ("catchment area") and random draw among applicants from the same neighbourhood, or random draw ("lottery") among all applicants (regardless of place of residence).

The total number of school places P per cohort across the two schools G and B is equal to the number of households with children in stage T_1 of their life. The school G has P_G places and the school B has P_B places. These numbers are fixed by policy and do not respond to excess demand in our model. The total number of applicants to school G, denoted A is all households in the life stage T_1 with children. A_H (respectively A_L) of these live in neighbourhood H (respectively L):

$$A = A_H + A_L = N(\phi_1 + \phi_2) \cdot \tau_1 \tag{3}$$

Among these, some are guaranteed a place because an elder sibling is already attending school G. The numbers of two-children families in this situation in each neighbourhood are denoted S_H and S_L ($S_H + S_L = S$). These will be derived in section 2.2.

The allocation rule means that, depending of place of residence n, the probability $\pi_G(n)$ of being granted a place in the good school for households without a sibling priority is:

$$\pi_G(n_H) = \pi_G(n_L) = min\left(1, \frac{P_G - S}{A - S}\right) \tag{4}$$

in the case of lottery allocation and:

$$\begin{aligned} \pi_G(n_H) &= min\left(1, \frac{P_G - S}{A_H - S_H}\right) &(5)\\ \pi_G(n_L) &= min\left(1, \frac{max(0, P_G - S - \pi_G(n_H) \cdot (A_H - S_H))}{A_L - S_L}\right) &(6) \end{aligned}$$

in the case of allocation with catchment areas.

The above assumptions mean that households do not make a choice with respect to the secondary school they apply to, but that their place of residence may have an impact on their expected success in applying to the good school. We now turn to the property market.

In order to determine the numbers of households wishing to live in each neighbourhood, we examine the cost-benefit analysis of choosing neighbourhood H versus neighbourhood L for all households types as defines by their fertility types and life stage. Given that we have 3 fertility types and 4 life stages, we have 12 different household types for whom the relative benefits of the two neighbourhoods will vary. We denote the rent level in neighbourhood H as r and normalise[6] the rent in neighbourhood L as r_0. Since households are forward looking, they will base their decision on the comparison of lifetime values of choosing either neighbourhood in the current period. These will be defined below.

Looking at the specification of the utility function (2), we see that the only term in which neighbourhood choice interacts with income is in the utility of consumption. Given the strictly diminishing marginal utility of consumption in (2), there is a unique income threshold above which

[6]Without loss of generality.

households will choose neighbourhood H for each of the 12 household types. For the same reason, this threshold will be increasing in r. In fact, these thresholds also depend on the neighbourhood in which the family resided in the previous period since there is a moving cost, and for families with two children in life stage T_2, the threshold will depend on whether the elder child has been given a place in school G or not. For ease of notation, we summarise the demand of properties in neighbourhood H with a single threshold per family type (we relax this in the full model derivation below) $\tilde{y}(f,t,r)$ and can formalise the market clearing of housing in neighbourhood H[7] as:

$$N_H = \sum_{f=0}^{2} \sum_{t=0}^{3} \tau_t \phi_f \bar{F} \left(\tilde{y}(f,t,r) \right) \tag{7}$$

where we denote $F(.)$ the cdf of the distribution of income y across households, which for now we assume is the same across household types and N_H is the number of properties in neighbourhood H, which is exogenous in our framework. Also, $\bar{F} = 1 - F$. Since the right-hand side decreases with r, this define a unique rent level $\tilde{r}$ at which the housing market clears in both neighbourhoods.

The two dynamic components of our framework are the following. First, there is a cost m of moving across neighbourhoods. Second, for families 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 neighbourhood choice.

Note that, within each life stage, yearly utility flows are constant, so the sum of discounted utility flows over the T_t years of any life phase t is equal to this flow multiplied by a factor β_t defined as:

$$\beta_t = \frac{1 - \beta^{T_t - 1}}{1 - \beta}, \tag{8}$$

where β is the yearly discount factor of households.

The dynamic choice of a household of type f and income y over the life cycle is a sequence $\{n_t\}_{t=0..3}$ of neighbourhood choices that maximises:

$$\sum_{t=0}^{3} \beta^{\sum_{t'=0}^{t-1} T_{t'}} \cdot \beta_t \cdot EU(f,t,y,n_t) - m \sum_{t=1}^{3} \mathbb{1}_{(n_t \neq n_{t-1})} \tag{9}$$

where, for ease of notation, we assume that $\sum_{t'=0}^{t-1} T_{t'} = 0$ and:

$$EU(f,2,y,n_2) = \pi_G(n_1) \cdot U(f,2,y,G,n_2) + (1 - \pi_G(n_1)) \cdot U(f,2,y,B,n_2) \tag{10}$$

This dynamic choice problem will be solved by backward iteration (detailed in section 2.1). The outcomes of interest for our purposes are the extent of income segregation among parents ($f > 0$) and non-parents ($f = 0$) in each neighbourhood resulting from the school allocation rule, the relative quality of school G, and the relative amenity in neighbourhood H.

[7] Since the set of the two neighbourhoods comprises as many properties as households, when one market clears, the other does too.

2.1 Optimal choices

In the last stage of life, $T_t = 3$, $s = 0$ for all households as all types are post-children. Neighbourhood choice is therefore independent of f. There is no uncertainty on school allocation.

$$\tilde{n}_3(y, n_2) = \mathit{argmax}_{\ n_3} U(f, 3, y, 0, n_2, n_3),$$

which leads an optimal value $V_3(y, n_2) = U(f, 3, y, 0, n_2, \tilde{n}_3(y, n_2))$.

When $T_t = 2$, $s \in \{0, 1\}$ depending on success in application in $T_t = 1$, but there is no remaining uncertainty.

$$\tilde{n}_2(f, y, s, n_1) = \mathit{argmax}_{\ n_2} [U(f, 2, y, s, n_1, n_2) + \beta^{T_2} V_3(y, n_2)], \tag{11}$$

which leads an optimal value of:

$$V_2(f, y, s, n_1) = U(f, 2, y, s, n_1, \tilde{n}_2(f, y, s, n_1)) + \beta^{T_2} \cdot V_3(y, \tilde{n}_2(f, y, s, n_1)).$$

In the school application life stage, $T_t = 1$, $s = 0$ for all household types as the older sibling is not yet in school. There is uncertainty in school allocation in the next period.

$$\tilde{n}_1(f, y, n_0) \quad = \quad \mathit{argmax}_{\ n_1} \left(\begin{array}{l} U(f, 1, y, 0, n_0, n_1) \\ +\beta^{T_1} \left[\begin{array}{l} \mathbb{1}(f > 0)\,(p_G(n_1) V_2(f, y, 1, n_1)] + (1 - p_G(n_1)) V_2(f, y, 0, n_1)) \\ + \mathbb{1}(f = 0) \cdot V_2(f, y, 0, n_1) \end{array} \right] \end{array} \right),$$

which leads an optimal value of:

$$\begin{array}{rcl} V_1(f, y, n_0) & = & U(f, 1, y, 0, n_0, \tilde{n}_1) \\ & + & \beta^{T_1} \left[\begin{array}{l} \mathbb{1}(f > 0)\,(p_G(\tilde{n}_1) V_2(f, y, 1, \tilde{n}_1) + (1 - p_G(\tilde{n}_1)) V_2(f, y, 0, \tilde{n}_1)) \\ + \mathbb{1}(f = 0) V_2(f, y, 0, \tilde{n}_1) \end{array} \right], \end{array}$$

where $\tilde{n}_1 = \tilde{n}_1(f, y, n_0)$ for ease of notation.

Finally, in the first period, $T_t = 0$, we have $s = 0$ for all household types, as the older sibling is not yet in school. There is no school uncertainty yet and:

$$\tilde{n}_0(f, y) = \mathit{argmax}_{\ n_0} \left[U(f, 0, y, 0, n_0, n_0) + \beta^{T_0} V_1(f, y, n_0) \right]$$

2.2 Model solution

WORK IN PROGRESS

Define seven functions: $\tilde{n}_3, \tilde{n}_2, \tilde{n}_1, \tilde{n}_0$ and V_3, V_2, V_1

Number of applicants to the Good school from the High quality neighbourhood: sum over family types with child(ren) in period 1, and neighbourhood choices in period 0.

$$N_{GH} = \tau_1 \sum_{f=1,2} \left[\phi_f \sum_{n_0} \int_y \mathbb{1}(\tilde{n_1}(f,y,n_0) = H)\, \mathbb{1}(\tilde{n_0}(f,y) = n_0) dF_f(y) \right] \tag{12}$$

In the steady state this is equal to the number of family type 2 successful in application to the Good school for their first child in the current period.

$$S = N\phi_2\tau_1[\bar{F}(\tilde{y}_{22}p_{GH} + F(\tilde{y}_{22})p_{GL}] \tag{13}$$

Where N is the total number of households, $\bar{F} = 1 - F$, and $\tilde{y}_2$ is the threshold type for $f = 2, t = 2$ that chooses to live in the High quality neighbourhood.

To solve the model, find the values of $\tilde{n}_3, \tilde{n}_2, \tilde{n}_1, \tilde{n}_0, V_3, V_2, V_1$ where the school and housing markets clear. That is, find r_H (r_L fixed) such that the housing markets clear and the probability of admission to the Good school is consistent with demand for housing. The structure of the code is as follows:

- Create vector of exogenous p_{GH} and p_{GL} to iterate over.
- Given exogenous p_{GH} and p_{GL} calculate r_H.
- Given r_H calculate $\tilde{y}_{ft}$ for all f and t.
- Calculate the endogenous p_{GH} and p_{GL} given $\tilde{y}_{ft}$.
- Find where the exogenous p_{GH} and p_{GL} equals the endogenous p_{GH} and p_{GL}. This is the equilibirum outcome.

Figure 1 shows the results of the iteration where p_{GH} and p_{GL} are in relatively coarse bins (of 0.1 intervals). The darker blue area shows where the absolute difference is smallest, where $p_{GL} = 0$ and p_{GH} is around 0.6-0.7. The code then iterates over smaller intervals of p_{GH} around this level (0.01 intervals) where $p_{GL} = 0$. Figure 2 shows there is a unique equilibrium where $p_{GL} = 0$ and $p_{GH} = 0.61$

3 Stylised facts

This section presents the key features of differences across three neighbourhoods in a city in the South West of England, Bristol. The neighbourhoods are chosen to represent one area containing a Good school and high neighbourhood quality (corresponding to school G and neighbourhood H in the model), one area containing a less-Good (Bad) school and comparable neighbourhood quality (school B and neighbourhood H), and one area containing a less-Good (Bad) school with lower neighbourhood quality (school B and neighbourhood L). These short-hand names do not truly reflect the school and neighbourhood characteristics, but are used for convenience to be consistent with the model. "Neighbourhood quality" in this context could best be described as proximity to the city centre and amenities. We choose one school to represent the "Good" school as it has the highest academic attainment and the most school choices per places observed in administrative data. All schools have comparable levels of "value-added" or average pupil progress, however, and

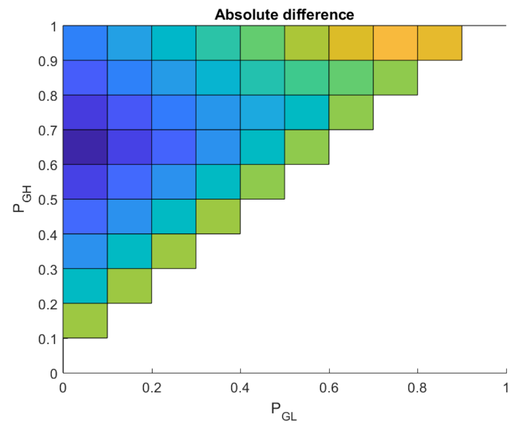

Figure 1: Finding equilibrium: Absolute difference between exogenous and endogenous p_{GH} and p_{GL} with intervals of 0.1

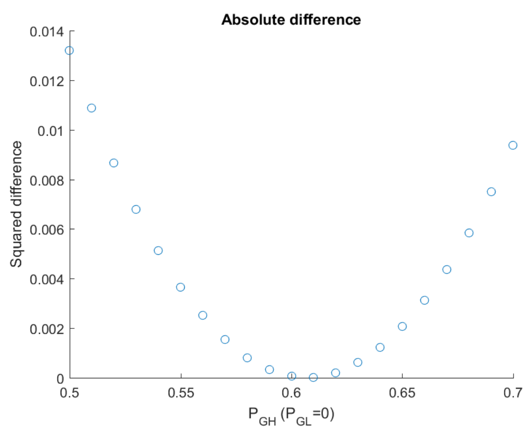

Figure 2: Finding equilibrium: Absolute difference between exogenous and endogenous p_{GH} and p_{GL} with intervals of 0.01 in p_{GH}

indeed one “Bad” school had a higher official grading by independent inspectors than the “Good” school in our chosen time period.

The characteristics of these schools are presented in Table 1. Pupils attending Redland Green have the highest attainment in both 2011 and 2012[8]. For example, in 2011, 83% of pupils achieved at least 5 GSCEs above grade C, compared to 66% in Cotham and 50% in Fairfield. There are similarly large differences in the percentage of pupils achieving the English Baccalaureate (EBACC) with 50% in Redland Green in 2011, compared to 23% in Cotham and 17% in Fairfield. These large differences in the attainment of pupils are not necessarily the result of school quality, and final test scores are the result of school, child and parent inputs. Measures of Value-Added, the average progress pupils make from entry to exit, is therefore not always consistent with the attainment data. In 2011 Redland and Cotham have similarly high Value-Added, while in 2012 Fairfield has the highest. The most recent Ofsted grade, which provides an independent “snapshot” summary measure of school quality is higher in Cotham than in Redland Green[9], and lowest in Fairfield.

In practice, parents have preferences for the peer composition as well as school quality (Burgess et al. (2015); Borghans et al. (2015)). The percentage of peers eligible for Free School Meals (FSM) is a proxy for the intake of the school. This is lowest in Redland Green, where only 4% of pupils in 2011 are eligible for Free School Meals, compared to 22% in Bristol. Cotham and Fairfield are more comparable with the Bristol average, with 18% and 24% of pupils eligible for Free School Meals, respectively.

Redland Green is most likely to be chosen by pupils living in the catchment area. 86% of pupils living in the Redland Green catchment area name the school as first choice, compared to 58% in the Cotham catchment area and only 18% in the Fairfield catchment area. There is a higher flow of choices from the Cotham catchment area to Redland Green than vice versa: 12% of pupils choose Redland Green from Cotham catchment versus 1% in the other direction. Pupils in Fairfield are most likely to choose Cotham than Redland Green, with 15% and 3%, respectivly, choosing the out-of-catchment school. These are not perfect measures of parents’ demand for schools, as first choices may reflect the perceived chance of admission (wanting to avoid “a wasted choice”) or the presence of siblings at the school (which almost guarantees the probability of admission). These patterns suggest, however, that Redland Green is the most popular school, retaining the majority of pupils residing in the catchment area.

[8]This pattern holds up until the latest year of attainment data

[9]This was reversed in the next inspection round, when Redland Green became Outstanding and Cotham became Good

Table 1: Descriptive statistics for three neighbouring secondary schools in Bristol

	Bristol Mean	Redland Mean	Cotham Mean	Fairfield Mean
School performance (2011)				
% 5A*-C (including English and Maths)[1]	53	83	66	50
% EBACC[2]	14	50	23	17
% Value-Added[3]	999.9	1023.7	1023.5	1002.8
Ofsted grade[4]		Good	Outstanding	Satisfactory
% FSM[5]	22	4	18	24
School performance (2012)				
% 5A*-C (including English and Maths)	55	84	56	52
% EBACC	13	56	18	11
% Value-Added	1002.9	1024.0	1012.9.5	1031.4
Ofsted grade		Good	Outstanding	Satisfactory
% FSM	37	6	34	46
School choices (2014)				
% Choose catchment school		86	58	18
% Choose Redland Green from catchment		86	12	3
% Choose Cotham from catchment		1	58	15
% Choose Fairfield from catchment		3	15	18

[1] 5A*-C is the percentage of pupils that achieve at least 5 GCSEs at high grades (A*-C) including English and mathematics. This was the benchmark measure of attainment used to compare schools until 2016.

[2] 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.

[3] Value-Added is the average progress made by pupils at the school from the end of primary school to the end of secondary school.

[4] Ofsted is the Office for Standards in Education, Children's Services and Skills. Ofsted provides inspection reports for schools at regular intervals, with a summary measure of school quality.

[5] FSM is the percentage of pupils eligible for Free School Meals, which is a proxy for income disadvantage.

Data to illustrate the differences across neighbourhoods are primarily from the 2011 Census, measured at a low level of geography (lower level super output area).[10] To create a proxy for household income, the number of household reference people of each National Statistics Socio-Economic Classification is combined with the median total net weekly earnings of the group observed in the Labour Force Survey. Property prices are from the Land Registry database of all sale prices. The variation in property prices due to neighbourhood and school (and all other) factors is calculated by taking the residual of house prices conditional on the total floor area, presence of an open fireplace and total floor area interacted with the number of habitable rooms and whether the property is a flat/maisonette. The observable characteristics of the properties are taken from the 87% of properties with a perfect match between the price paid dataset and Energy Performance Certificate database. See Appendix B for full details on the data construction.

[10] LSOAs are homogeneous small areas of relatively even size (around 1,500 people)

Figure 3 shows the selected neighbourhoods, with Redland (GH) shown in yellow, Cotham (BH) in green and Fairfield (BL) in blue. The orange lines show boundaries in school catchment areas. Lower level super output areas within these catchment areas are chosen to represent the neighbourhoods with priority at particular schools. The area close to Redland is chosen for Cotham so neighbourhood attributes are comparable. The area further from Redland is chosen for Fairfield so neighbourhood attributes are different.

These differences are summarised in Table 2. The following figures show the variation in household composition, income and residual property prices between LSOAs in these three neighbourhoods. In Bristol as a whole, 15% of households have low income. This varies across our chosen neighbourhoods, from 4% in Redland and Cotham to 23% in Fairfield. Imputed average income of the household reference person (in £ per week) shows a similar gradient. The average in Bristol as a whole is £305 per week, compared to £379 in Redland, £299 in Cotham and £285 in Fairfield. Figure 4 shows the variation in average income across LSOAs in our three areas, with particularly high income in the LSOA to the North-West of Redland Green school. Imputed average incomes of the household reference person are derived from the National Statistics Socio-Economic classification of the household reference person. It is therefore unsurprising to see similar variation in the presence of occupation types across areas. For example, the mode group in Redland and Cotham catchment areas is NSSEC 1 (higher managerial, administrative and professional) with the highest percentage in Redland. In Fairfield, the mode group is NSSEC 2 (lower managerial, administrative and professional) but with a comparable number in NSSEC 6 (semi-routine) occupations.

The household life-stage is classified according to the age of the household reference person and our modelling assumptions. $t = 0$ refers to the period before the secondary school choice phase (when adults are aged between 25 and 40). $t = 1$ refers to the secondary school choice phase which lasts only two years (when adults are aged between 40 and 42). $t = 2$ refers to the secondary school phase which lasts eight years (when adults are aged between 42 and 50). $t = 3$ refers to the post secondary school phase (when adults are aged between 50 and 80). There is a clear concentration of households of secondary school choice and attendance phase in Redland. For example, 24% of households are at an age consistent with the secondary school attendance phase in Redland, compared to 16% in Cotham and 21% in Fairfield. Redland also has a low share of households of an age consistent with the post-secondary school attendance phase (23%) compared to 28% across Bristol as a whole, 17% in Cotham (which has a predominantly young population) and 27% in Fairfield. In line with this, Redland has the highest share of households with dependent children (35% compared to 30% in Bristol as a whole) and households where the youngest dependent child is of secondary school age (8% compared to 6% in Bristol as a whole). Cotham has a low share of households with dependent children (14%) while Fairfield has a slightly lower share than Redland (32%). Figure 5 and Figure 6 show the share of households with dependent children and dependent children of secondary school age across LSOAs, respectively. The share is higher for the majority of LSOAs inside the Redland catchment area than outside, but particularly so for the LSOA to the North-West of Redland Green school where 45% of households have a dependent child and 12% of households have a youngest dependent child of secondary school age.

The housing stock is more suited to large family homes in Redland and Fairfield, where around 64% of homes have at least three bedrooms, compared to 59% of homes in Bristol overall, 39% in Cotham and 60% in Fairfield. Properties in Redland are also markedly larger, with 56% with at least six rooms, compared to 39% in Bristol overall. Again, there is variation across LSOAs within catchment areas, shown in Figure 7, although the LSOA to the North-West of Redland Green school is not a particular outlier in this case.

Due to variation in property sizes across catchment areas we consider the residual property price as our main measure of demand. The residual property price abstracts from variations in price that are due to total floor size, the number of habitable rooms and the presence of an open fireplace (as a proxy for period property features). The residual price best captures the property price due to neighbourhood and school characteristics, with the caveat that other features unobservable to us, such as the decoration and presence/size of garden, may also affect property prices. As expected, the mean residual price in Bristol as a whole is close to zero[11]. The mean residual property price in Redland is £71,700, compared to £42,000 in Cotham and -£34,000 in Fairfield. At face value, this suggests that households value properties in areas with Good school quality and High neighbourhood amenities more than both areas with Bad school quality and High neighbourhood amenities and Bad school quality and Low neighbourhood quality.

[11]It is not exactly zero as the initial regression was at the property level. The residuals were then aggregated to the LSOA level. The table shows the average of these LSOA averages

Table 2: Descriptive statistics for Bristol and three specific catchment areas

	Bristol		Redland		Cotham (close)		Fairfield (far)	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Income and occupation[1]								
Low income score[2]	0.15	0.11	0.04	0.01	0.04	0.01	0.23	0.08
Imputed average income (£ per week)[3]	304.82	46.14	378.99	12.22	298.95	44.69	284.75	26.48
% NSSEC 1. Higher managerial admin. and prof.	15.23	11.22	36.19	2.91	28.47	5.14	8.35	3.69
% NSSEC 2. Lower managerial admin. and prof.	21.69	6.91	30.11	3.32	23.34	3.31	20.08	5.94
% NSSEC 3. Intermediate	10.28	2.62	7.79	1.54	7.04	1.46	9.63	0.95
% NSSEC 4. Small employers and own account workers	9.87	2.95	8.42	0.93	5.91	1.77	12.22	3.09
% NSSEC 5. Lower supervisory and technical	7.96	3.6	2.76	0.58	2.54	0.83	7.63	1.35
% NSSEC 6. Semi-routine	13.21	6.24	4.46	0.85	4.22	1.16	17.2	4.6
% NSSEC 7. Routine	12.44	7.57	2.33	0.65	1.9	1.09	15.11	4.93
% NSSEC 8. Never worked and long-term unemployed	4.36	3.68	1.15	0.78	1.23	0.38	6.71	2.29
Household lifestage[4]								
% t=0: pre-school choice phase	46.17	9.98	47.11	7.48	62.52	6.42	46.52	6.99
% t=1: school choice phase	5.22	0.68	5.99	0.86	4.08	0.54	5.28	0.19
% t=2: secondary-school phase	20.89	2.71	23.96	3.44	16.31	2.16	21.12	0.74
% t=3: post-secondary-school phase	27.72	8	22.94	4.04	17.09	4.11	27.08	6.34
% dependent children	30.49	9.11	34.5	8.37	14.14	4.59	32.41	3.53
% dependent children (youngest age 11-18) (HRP 25-54)	6.06	2.26	7.71	2.32	2.76	1.22	6.34	1.19
Housing stock								
% 3 or more bedrooms	58.99	19.06	64.25	16.71	38.75	9	60.49	16.69
% 6 or more rooms	38.64	16.26	56.01	16.84	30.19	9.6	37.05	12.26
Property prices[5]								
Mean price per LSOA (£1,000)	199.62	95.36	342.83	51.22	325.99	69.4	151.49	32.04
Mean price residual per LSOA (£1,000)	-6.79	49.12	71.73	25.01	42.04	27.56	-34.02	29.98
Median price per LSOA (£1,000)	189.64	86.91	330.38	75.89	273.96	77.55	153.71	31.82
Median price residual per LSOA (£1,000)	-3.88	44.75	68.93	36.19	42.33	35.07	-26.48	29.71

All characteristics are measured at the lower level super output area (LSOA). Columns 1-2 show the average and standard deviation across all LSOAs in Bristol. Columns 3-4 shows the equivalent for the 8 LSOAs in the immediate Redland Green catchment area, Columns 5-6 the 9 LSOAs in the Cotham catchment area (close to Redland Green) and Columns 7-8 the 6 LSOAs in the Fairfield catchment area (further from Redland Green).

[1] Occupation classifications are the National Statistics Socio-Economic Classification of the household reference person (HRP).

[2] 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.

[3] Imputed income is derived as the average median income within NSSEC for the household reference person. Median income is calculated from the Labour Force Survey.

[4] Household life-stage refers to the age of the household reference person, where t indicates time period consistent with the age of households in our model.

[5] Property prices are taken from the universal Land Registry Database and aggregated to LSOA level. The residual prices are the variation in price not accounted for by total floor area, the number of rooms and presence of an open fireplace.

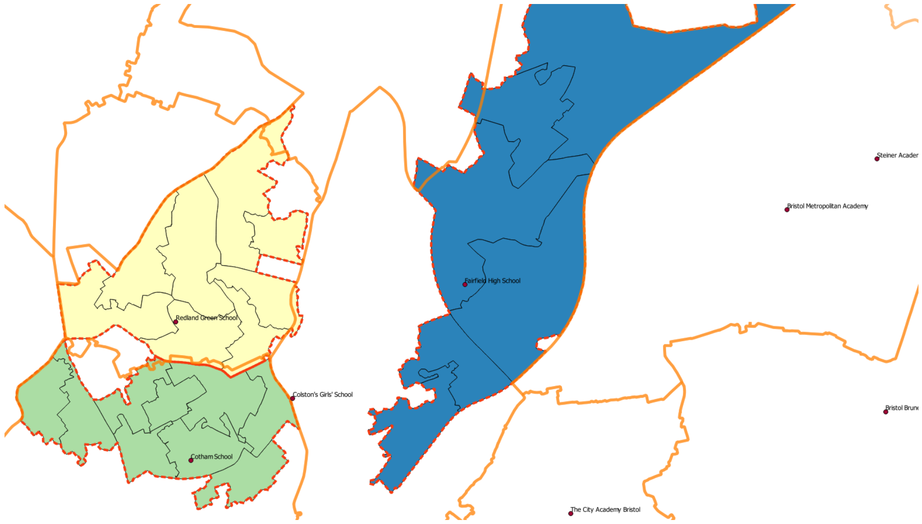

Figure 3: Catchment areas

4 Comparative Statics and Results

This section shows the comparative statics from the model primarily comparing the equilibrium outcomes under the lottery and catchment area cases. We also show the equilibrium outcomes when varying three key parameters (all else equal). These are: the proportion of Parents (those with completed family size 1 or 2) in the population; the utility from living in the High quality neighbourhood; the utility from attending the Good school. In each case, the dotted lines represent the equilibrium under the lottery assignment and solid line the equilibrium under catchment area. In a later version of the paper we will also present results of the model calibrated to match the stylised facts Section 3.

4.1 Comparative Statics

Figure 8 describes the equilibria under lottery and neighbourhood assignment as the proportion of Parents in the population increases from 0.1 to 0.9. We specify the number of school places to stay equal to the number of relevant children in the population, maintaining 60% of school places in the Good school.

The first panel shows the thresholds above which households choose to live in the High quality neighbourhood in the school choice period (period 1). Family type is represented by $F0$, $F1$ and $F2$ for family type with completed fertility 0, 1 and 2 children, respectively. Previous neighbourhood is denoted by $H0$ for the High quality neighbourhood in period 0 and $L0$ for the Low quality neighbourhood. A lower threshold implies a higher proportion of the group choose to live in the High quality neighbourhood. For each household type, this threshold is higher (and so the number of households is lower) if the previous location choice was the Low quality neighbourhood due

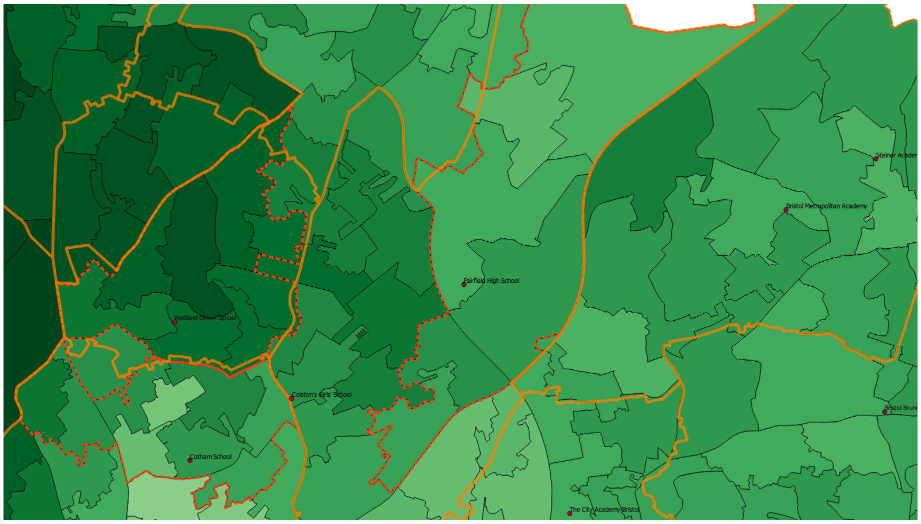

Figure 4: Average income imputed by household NSSEC

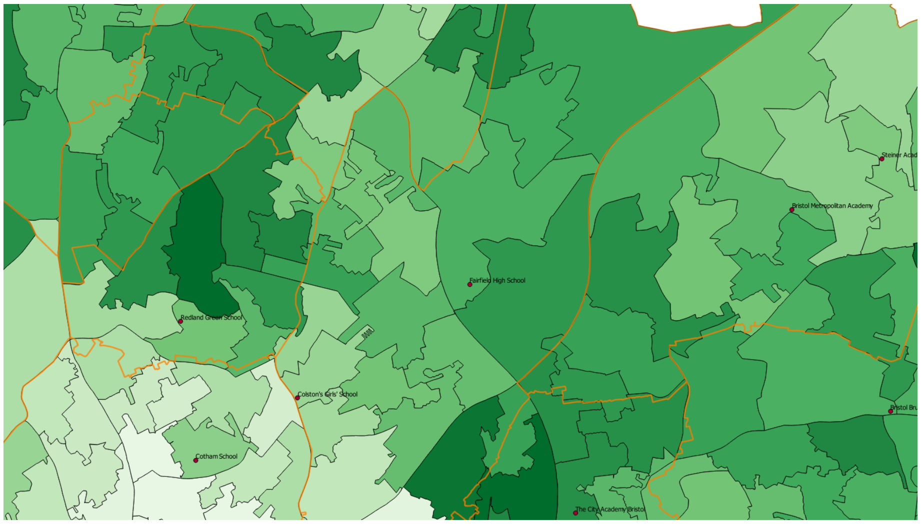

Figure 5: Share of households with dependent child(ren)

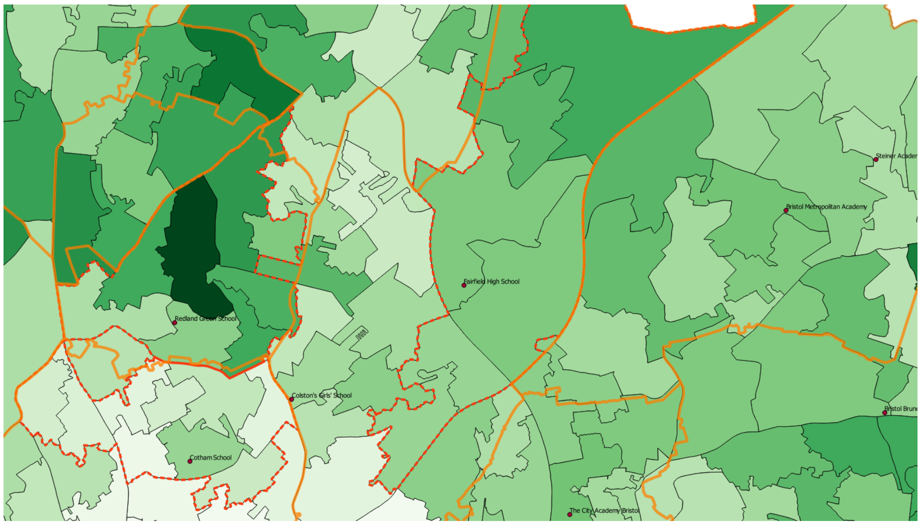

Figure 6: Share of households with dependent child(ren) where the youngest is of secondary school age

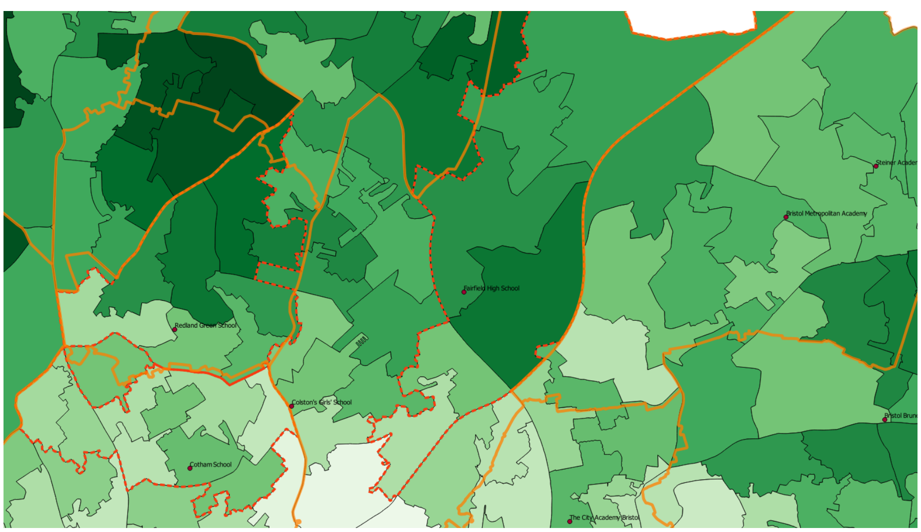

Figure 7: Share of households with at least three bedrooms

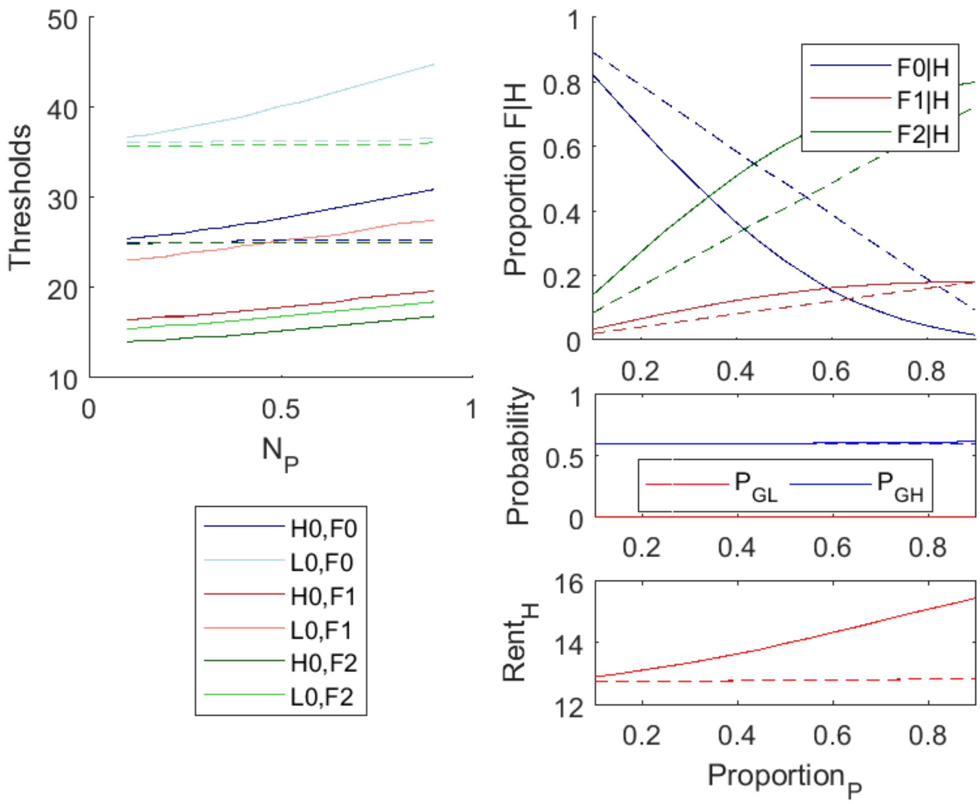

Figure 8: Comparative statics: Equilibrium outcomes as the proportion of Parents in the population increases

to moving costs. Under the neighbourhood allocation, given previous neighbourhood, households with completed fertility of two children have a lower threshold than those with completed fertility of one child, which is lower again than those that never have children. This is because the utility from the High quality neighbourhood is highest because they gain most utility from their children having priority at the Good school.[12] Indeed, for those with two children the threshold for those previously in the Low quality neighbourhood is almost as low as for those previously in the High quality neighbourhood, as the cost of moving is almost outweighed by the higher utility associated with priority at the Good school. Under the neighbourhood allocation the thresholds increase as the proportion of Parents increases. This is because, given preferences for priority at the Good school, demand in the High quality neighbourhood increases and the equilibrium rent increases. Under the lottery system, thresholds are constant as the proportion of Parents in the population increases, and overlap for family types with the same previous neighbourhood. This is because family types with children gain no additional benefit from living in the High quality neighbourhood above utility from amenities.

The second panel shows the proportion of each family type living in the High quality neighbourhood, combining life-stages (or periods). $F0|H$ denotes the proportion of households in the High quality neighbourhood wit completed family size of 0, for example. Mechanically, as the proportion of Parents in the population increases the proportion of those without children decreases in both the lottery and catchment area case. In the lottery case, the proportion of each family type in the High quality neighbourhood is equal to the proportion in the whole population. This is because the distribution of income for each household type is the same, rent is constant, and therefore thresholds are constant. Comparing to the catchment area case, as the proportion of Parents increases, the proportion of family types with children is above the linear lottery case, and the proportion of family type without children is below. This illustrates the externalities that the criteria for allocating pupils to schools have on those without children.

The third panel shows the probability of admission to the Good school from the High quality neighbourhood P_{GH} and Low quality neighbourhood P_{GL}. This is 0.6 from both High and Low quality neighbourhoods under lottery allocation. Under catchment area it is around 0.6 from the High quality neighbourhood and 0 from the Low quality neighbourhood.

The final panel shows the equilibrium rent in the High quality neighbourhood, relative to a rent of 2 in the Low quality neighbourhood. In the lottery case the rent is constant as the proportion of Parents in the population increases, as the premium in the High quality neighbourhood is due to neighbourhood amenities only, which all family types value equally. Under the catchment area case the rent increases as the proportion of Parents increases, as a larger fraction of the population derives additional utility from the priority at the Good school.

Overall, comparison between the lottery and catchment area case shows that the lottery leads to an even mix between household types in the High and Low quality neighbourhoods. The Low quality neighbourhood contains lower income households as the rent in the High quality neighbourhood, although constant, is higher than the Low quality neighbourhood as households have a concave utility function in consumption.

Figure 9 shows the equilibria as the utility from living in the High quality neighbourhood in-

[12]Note that our utility function currently specifies linear utility in the number of children attending the Good school. We will explore the best specification for this, either convex or concave, in future work.

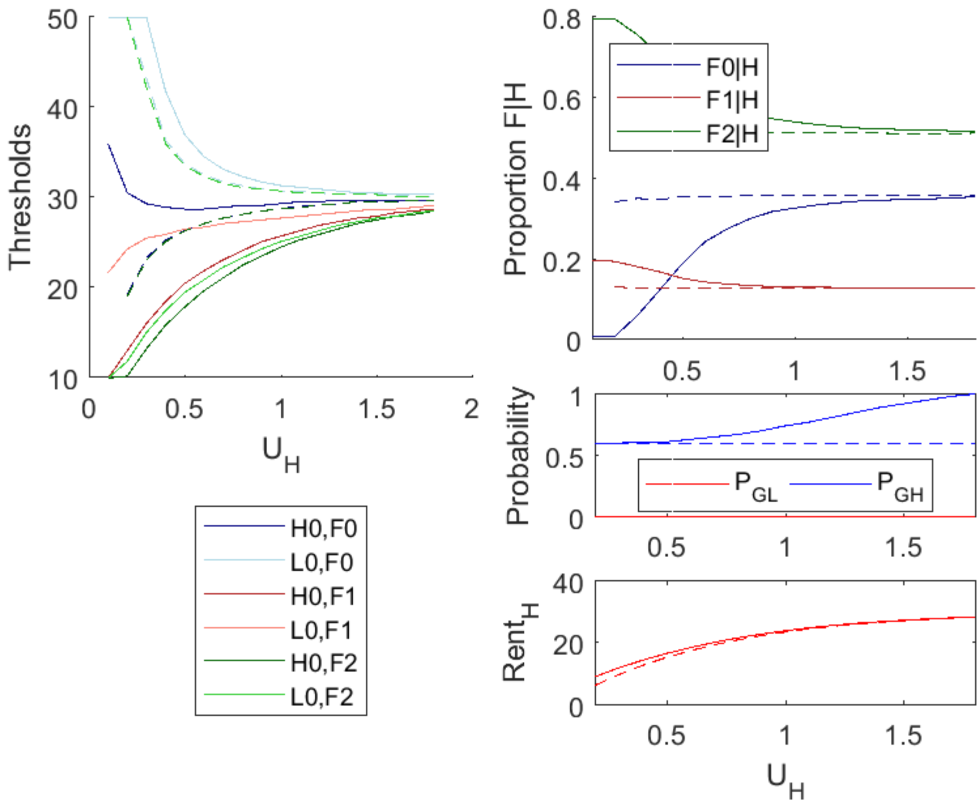

Figure 9: Comparative statics: Equilibrium outcomes as the utility from living the "High" quality neighbourhood increases

creases relative to the Low quality neighbourhood, all else equal. This demonstrates that the equilibrium outcome depends crucially on the relative weights that households place on neighbourhood quality compared to school quality (if relevant).

The obvious pattern in the first panel is that the thresholds for all family types and previous neighbourhood converge as the utility from the High quality neighbourhood increases, as preferences for neighbourhood quality come to dominate all other considerations. For family types with children, the utility from the High quality neighbourhood makes the utility from attending the Good school secondary. For households that previously lived in the Low quality neighbourhood the moving costs become trivial. The High neighbourhood therefore becomes exclusively for households with income above around 30 while those with income lower than this threshold (whatever their family type) live in the Low quality neighbourhood.

The second panel shows the implications of this pattern for the proportion of each family type (across all lifestages) in the High quality neighbourhood. As the utility from the High quality neighbourhood increases the proportion of each family type in the High quality neighbourhood converges to the proportion in the population. This is because the distribution of income is the same across family types. When the utility from the High quality neighbourhood is low the proportion of households with children in the High quality neighbourhood is larger than the population, as they gain utility from priority at the Good school. These panels show an interesting contrast between sorting by income and family type. When the utility from the High quality neighbourhood is low, there is greater mixing by income type but little mixing by family type. The reverse is the case when the utility from the High quality neighbourhood is high: there is perfect sorting by income and perfect integration by family type.

The third panel shows that the probability of admission to the Good school increases above 0.6 in the catchment area case only when the utility from the High quality neighbourhood increases and the proportion of households with children therefore decreases. Entry to the Good school is only guaranteed when the utility from the High quality neighbourhood is at the maximum value considered. The probability of entry from the Low quality neighbourhood is zero in the catchment area case but equal to the probability from the High quality neighbourhood in the lottery case.

The final panel shows the that, as expected, the equilibrium rent in the High quality neighbourhood increases with the utility from the High quality neighbourhood. A difference between the lottery and catchment area case is observed only when the utility from the High quality neighbourhood is low. When it is high, this factor dominates all others so all households have a strong preference to live in the High quality neighbourhood independent of school allocation mechanism.

Finally, Figure 10 shows the equilibria as the utility from the Good school increases. Intuitively, family types with children want to maximise the probability that their child(ren) attend the Good school. The premium they are willing to pay depends on the additional utility they derive. The first panel shows that as the utility from attending the Good school increases the threshold for living in the High quality neighbourhood decreases under the catchment area system. This is because households are willing to sacrifice more consumption to increase the probability of attending the Good school. Mixing across income types therefore increases as the utility from the Good school increases. As in previous figures the thresholds for households that previously lived in the Low quality neighbourhood are lower due to fixed moving costs. Under the lottery system thresholds by family type are the same, conditional on neighbourhood in the previous period.

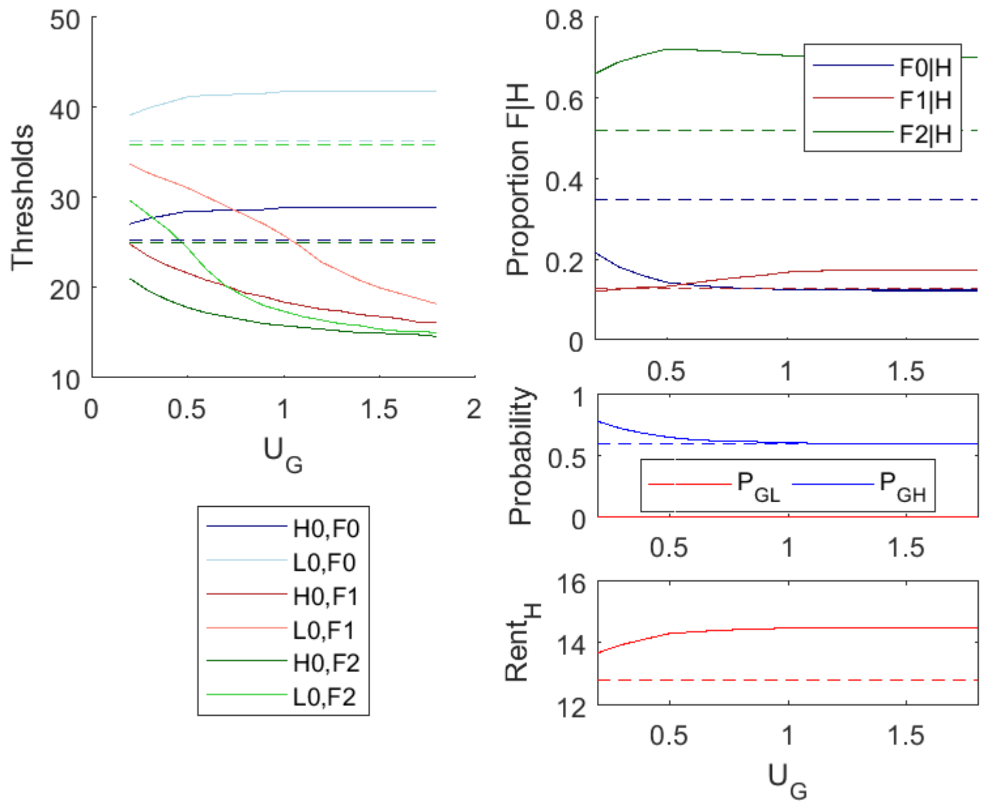

Figure 10: Comparative statics: Equilibrium outcomes as the utility from attending the "Good" school increases

The second panel shows the proportion of each family type living in the High quality neighbourhood across all lifestages. Given the constant thresholds in the lottery case, the proportions of each family type are also constant and equal to the proportion in the population. Under the catchment area system there is a higher proportion of households with one or two children. The proportion of households without children therefore declines as the utility from attending the Good school increases.

The probability of admission to the Good school declines from around 0.75 when the utility from the Good school is low to 0.6 when the utility from the Good school is around 1. Note that there is over-demand for the Good school in the High quality neighbourhood even when the utility from the Good school is relatively low.

Comparing across these comparative statics it is clear that the rent premium in the High quality neighbourhood is highest when the utility from living in the High quality neighbourhood increases (with similar predictions under the lottery and catchment area systems). In this case there is perfect sorting by income and perfect integration by family type. The rent premium is much lower when instead the utility from attending the Good school dominates. This leads to more integration by income and more sorting by family type. Comparing the lottery and catchment area systems, there is a much larger difference as the utility from attending the Good school increases. Moving from a catchment area to lottery system in this case would decrease integration by income type and increase integration by family type. Appropriate policy responses therefore depend on the utility function of the social planner. What weight should society place on integration by family type relative to integration by income across neighbourhoods, in turn relative to income integration in schools?

4.2 Results from dynamic model

WORK IN PROGRESS

This section will show to equilibrium outcomes and overall welfare calculations if Bristol schools moved from catchment areas (neighbourhood priority) to lottery if over-subscribed. We will calibrate the model to fit the observed stylised facts shown above under the catchment area system, and then simulate the effect of moving to a lottery system. The model could be calibrated to other cities in England and elsewhere in the world.

5 Conclusion

WORK IN PROGRESS

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 school choice. The innovation of this paper is to consider school priorities (the ranking of applicants if oversubscribed) rather than the allocation mechanism (used by the central authority to assign places) on the equilibrium outcomes for schools and neighbourhoods. We quantify the relationship between the type of school priority (catchment area versus lottery) and the social mix within schools and within neighbourhoods, defined as the mix between households of different ages, different income and different family size. We build the first dynamic structural model of household choices across different life stages, allowing for heterogeneity in family types along completed family size. Households care for local amenities, school quality and distance to school when there are children travelling to school. Residential location matters most in the life stage when parents apply to a secondary school (if there is neighbourhood sorting) and in the life stage when the household's children travel to school. Because households are forward looking and there is a cost of moving across neighbourhoods, households' residential choices exhibit some persistence.

We calibrate our model to data from three neighbourhoods in the city of Bristol, England, and find that it replicates patterns of sorting in schools and neighbourhood across family types and life stages well. The area around the "Good" school in high demand is characterised by a larger proportion of households of secondary school phase age, a larger proportion of households with dependent children (particularly with the youngest dependent child of secondary school age) and a lower proportion of households in the final stage of life. Residual house prices are higher than in the area with similarly high neighbourhood quality but lower school quality (according to attainment data) and the area with lower neighbourhood quality and lower school quality. In future work our results will provide structural estimates of households' willingness to pay for a good school and neighbourhood amenities along the life-cycle and across family size.

Comparative statics from the model illustrate the potential trade-off between integration in neighbourhoods and schools. When residential location does not affect priority to the "good" school neighbourhoods are perfectly sorted according to income, with an equal proportion of family types, while the "good" school has a more integrated composition. In the alternative "catchment area" case where living in the High quality neighbourhood increases the probability of admission to the "good" school, neighbourhoods become more integrated by income but less integrated by family type (with a higher concentration of households with children) and schools become less integrated according income. These results suggest a potentially difficult trade-off for a social planner wishing to increase integration between household types.

Future work will consider more precisely the effect of moving from a catchment area to lottery system for admissions in our case study area, with the potential to apply the model elsewhere. We will also develop extensions to the model, namely making school quality partly endogenous and incorporating a private school outside option. However, we believe our existing model provides unique insights into the relationship between school priorities, school choice and neighbourhood and school sorting.

References

Patrick Bayer, Fernando Ferreira, and Robert McMillan. A unified framework for measuring preferences for schools and neighborhoods. *Journal of Political Economy*, 115(4):588–638, 2007. doi: 10.1086/522381. URL `https://doi.org/10.1086/522381`.

Sandra E. Black. Do Better Schools Matter? Parental Valuation of Elementary Education*. *The Quarterly Journal of Economics*, 114(2):577–599, 05 1999. ISSN 0033-5533. doi: 10.1162/003355399556070. URL `https://dx.doi.org/10.1162/003355399556070`.

Sandra E. Black and Stephen Machin. Chapter 10 - housing valuations of school performance. In Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, editors, *Handbook of the Economics of Education*, volume 3, pages 485 – 519. Elsevier, 2011. doi: https://doi.org/10.1016/B978-0-444-53429-3.00010-7. URL http://www.sciencedirect.com/science/article/pii/B9780444534293000107.

William T. Bogart and Brian A. Cromwell. How much is a neighborhood school worth? *Journal of Urban Economics*, 47(2):280–305, 2000. URL https://EconPapers.repec.org/RePEc:eee:juecon:v:47:y:2000:i:2:p:280-305.

Lex Borghans, Bart H. H. Golsteyn, and Ulf Zölitz. Parental preferences for primary school characteristics. *The B.E. Journal of Economic Analysis & Policy*, 15, 01 2015. doi: 10.1515/bejeap-2014-0032.

Simon Burgess, Ellen Greaves, Anna Vignoles, and Deborah Wilson. What parents want: School preferences and school choice. *The Economic Journal*, 125(587):1262–1289, 2015. doi: 10.1111/ecoj.12153. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/ecoj.12153.

Caterina Calsamglia, Chao Fu, and Maia Güell. Structural estimation of a model of school choices: the boston mechanism vs. its alternatives. Working Paper 24588, National Bureau of Economic Research, May 2018. URL http://www.nber.org/papers/w24588.

Caterina Calsamiglia, Francisco Martínez-Mora, and Antonio Miralles. School choice mechanisms, peer effects and sorting. Discussion Papers in Economics 15/01, Department of Economics, University of Leicester, 2015. URL https://EconPapers.repec.org/RePEc:lec:leecon:15/01.

John Coldron, Caroline Cripps, and Lucy Shipton. Why are english secondary schools socially segregated? *Journal of Education Policy*, 25(1):19–35, 2010. doi: 10.1080/02680930903314285. URL https://doi.org/10.1080/02680930903314285.

Dennis N. Epple and Richard Romano. *Neighborhood Schools, Choice, and the Distribution of Educational Benefits*, pages 227–286. University of Chicago Press, January 2003. URL http://www.nber.org/chapters/c10090.

Maria Marta Ferreyra. Estimating the effects of private school vouchers in multidistrict economies. *American Economic Review*, 97(3):789–817, June 2007. doi: 10.1257/aer.97.3.789. URL http://www.aeaweb.org/articles?id=10.1257/aer.97.3.789.

Stephen Gibbons and Stephen Machin. Valuing school quality, better transport, and lower crime: evidence from house prices. *Oxford Review of Economic Policy*, 24(1):99–119, spring 2008. URL https://ideas.repec.org/a/oup/oxford/v24y2008i1p99-119.html.

Caroline M. Hoxby. Does competition among public schools benefit students and taxpayers? *American Economic Review*, 90(5):1209–1238, December 2000. doi: 10.1257/aer.90.5.1209. URL http://www.aeaweb.org/articles?id=10.1257/aer.90.5.1209.

Ron Johnston, Simon Burgess, Deborah Wilson, and Richard Harris. School and residential ethnic segregation: An analysis of variations across england's local education authorities. *Regional Studies*, 40(9):973–990, 2006. doi: 10.1080/00343400601047390. URL https://doi.org/10.1080/00343400601047390.

Stephen Machin and Kjell G Salvanes. Valuing school quality via a school choice reform. *Scandinavian Journal of Economics*, 118(1):3–24, 2016. URL `https://EconPapers.repec.org/RePEc:bla:scandj:v:118:y:2016:i:1:p:3-24`.

Thomas J. Nechyba. Mobility, targeting, and private-school vouchers. *The American Economic Review*, 90(1):130–146, 2000. ISSN 00028282. URL `http://www.jstor.org/stable/117284`.

Phuong Nguyen-Hoang and John Yinger. The capitalization of school quality into house values: A review. *Journal of Housing Economics*, 20(1):30 – 48, 2011. ISSN 1051-1377. doi: https://doi.org/10.1016/j.jhe.2011.02.001. URL `http://www.sciencedirect.com/science/article/pii/S1051137711000027`.

John Ries and Tsur Somerville. School Quality and Residential Property Values: Evidence from Vancouver Rezoning. *The Review of Economics and Statistics*, 92(4):928–944, November 2010. URL `https://ideas.repec.org/a/tpr/restat/v92y2010i4p928-944.html`.

Jesse Rothstein. Does competition among public schools benefit students and taxpayers? comment. *American Economic Review*, 97(5):2026–2037, December 2007. doi: 10.1257/aer.97.5.2026. URL `http://www.aeaweb.org/articles?id=10.1257/aer.97.5.2026`.

Chris Taylor and Stephen Gorard. The role of residence in school segregation: Placing the impact of parental choice in perspective. *Environment and Planning A: Economy and Space*, 33(10): 1829–1852, 2001. doi: 10.1068/a34123. URL `https://doi.org/10.1068/a34123`.

A Possible extensions to the model

A.1 School Quality is endogenous

The current model assumes that school quality is exogenous; it is not affected by the composition of the school or local neighbourhood. This means that school quality is entirely a function of management and teacher quality and resources. This unlikely to be true in most school systems in the world. If school quality is defined as test scores, then pupil composition directly affects school quality, as pupils from more affluent backgrounds have higher test scores, on average. If school quality is defined as a progress measure (such as contextual value added), then pupil composition affects school quality indirectly. For example, there is evidence that pupil composition has a causal affect on teacher applications and retention. An extension to the model is to therefore consider the case that school quality is a weighted function of pupil composition and a fixed component of school quality (such as management quality).

The implications for the model is that households make neighbourhood and school choices with expected school quality. Modelling the long-term equilibrium implies that households' expectations converge to perfect information. In this setting, school quality is a weighted function of pupil composition (increasing in y) and a fixed component of management quality.

$$SQ_i = \alpha Q_i + (1 - \alpha)\bar{y} \tag{14}$$

Where Q_i is fixed and $Q_G > Q_B$ and $\bar{y}$ is the average type in the school. α denotes the weight to fixed versus variable school quality, where $\alpha = 1$ corresponds to the benchmark case and $\alpha = 0$ corresponds to the case where school quality depends entirely on the average pupil composition.

The utility function for Parents is modified so that utility is strictly increasing in a continuous measure of school quality. Decreasing marginal returns to school quality seems reasonable.

Equilibrium now depends on the equilibrium school quality of the (now notional) "Good" and "Bad" school. Equilibrium is r_L^* and r_H^* such that given t_{NP}^* and t_P^* (as a function of p_{GH} and p_{GL} and SQ_G and SQ_B) clear the house and school markets. In practice this would mean another function in the code to calculate school quality and update utility from the High and Low neighbourhoods. It relies on the assumption of perfect information to find the equilibrium values.

A.2 Private school outside option

WORK IN PROGRESS

B Data Appendix

WORK IN PROGRESS