Centre for Market and Public Organisation
Bristol Institute of Public Affairs
University of Bristol
2 Priory Road
Bristol BS8 1TX

Tel: (0117) 33 10799
Fax: (0117) 33 10705
E-mail: cmpo-office@bristol.ac.uk

School Assignment, School Choice and Social Mobility

Simon Burgess and Adam Briggs

November 2006

Working Paper No. 06/157

ISSN 1473-625X

CMPO Working Paper Series No. 06/157

School Assignment, School Choice and Social Mobility

Simon Burgess[1]
and
Adam Briggs[2]

[1] *University of Bristol and CMPO*
[2] *CMPO*

November 2006

Abstract
We estimate the chances of poor and non-poor children getting places in good schools, analysing the relationship between poverty, location and school assignment. Our dataset allows us to measure location and distance very precisely. The simple unconditional difference in probabilities of attending a good school is substantial. We run an analysis that controls completely for location, exploiting within-street variation and controlling for other personal characteristics. Children from poor families are significantly less likely to go to good schools. We show that the lower chance of poor children attending a good school is essentially unaffected by the degree of choice.

Keywords: School assignment, social mobility, school choice

JEL Classification: I21, I28

Acknowledgements
Thanks to DfES for providing the PLASC/NPD data, which underlies this project. Thanks to ESRC and Leverhulme Trust for funding this project through CMPO. Thanks for comments to Paul Gregg, Rob McMillan, Carol Propper, and Deborah Wilson. None of these are responsible for the ideas in this paper. Conversations with Deborah Wilson informed some of the analysis in this paper.

Address for Correspondence
CMPO, Bristol Institute of Public Affairs
University of Bristol
2 Priory Road
Bristol
BS8 1TX
Simon.burgess@bristol.ac.uk
www.bris.ac.uk/Depts/CMPO/

CMPO is jointly funded by the Leverhulme Trust and the ESRC

1. Introduction

A big part of getting on in life is doing well at school, and doing well at school is helped by attending a good school. Since not all schools are good schools, places at good schools need to be allocated. These two assumptions – that schools matter and schools differ – mean that education markets have to solve an important assignment problem. In particular, the allocation across pupils from differing family backgrounds is an issue of interest for social mobility. This case is made stronger by recent research showing the increasing importance of family income in influencing life outcomes[1]. Education markets in England can be characterised as utilising a mix of choice-based schooling and neighbourhood-based schooling, so this study also relates to the debate on the effects of school choice.

In this paper we estimate the chances of poor and non-poor children getting places in good schools[2]. In a general sense this depends on demand expressed by parents and children, the availability of such places and the nature of the assignment mechanisms that resolve conflicting choices. In England, the details of these mechanisms are given by Local Education Authority's (LEA) and schools' admissions procedures. One of the key factors is location – distance between school and home, with those living nearest having priority. West *et al*. (2004) show that 86% of secondary schools (grammar schools excluded) have distance as part of their admissions criteria. This has a direct influence on the relationship between poverty and school assignment. The spatially concentrated demand pushes up house prices[3] and generates a correlation between poverty and distance from a good school. Our dataset allows us to measure distance very precisely and characterise the pupil's very local area. This allows us to analyse the relationship between poverty, location and school assignment.

This issue is not straightforward. The composition of the school affects the publicly available measure of school quality, which in turn influences parental choice of school. Non-poor children tend to score more highly all else equal, so schools with

[1] For example, Blanden, Goodman, Gregg and Machin (2004) show decreasing intergenerational mobility in the UK.
[2] We define all these terms properly below. We use 'poor' as a short hand for 'eligible for Free School Meals', which derives in turn from eligibility for certain welfare payments. A 'good' school is in the top third of published league tables for performance.
[3] For evidence see G&M. others

high proportions of poor children will tend to produce lower average scores. Thus a simple correlation between school quality and the proportion of poor children will not necessarily be informative about the assignment mechanism and school admissions. We adopt a number of different approaches to deal with this problem.

The simple unconditional difference in probabilities of attending a good school is substantial. A pupil from a poor family is 17 percentage points less likely to go to a good school, 14 points once we include other individual controls, but not location. This is relative to an overall ratio of 29% of places in good schools. If we add controls for characteristics of the pupil and her neighbourhood, both very local (a street – less than twenty dwellings) and broader (about 12,000 people), the difference falls to around 5 percentage points. However, it may be that these controls do not fully capture the features of a location and its spatial relationship to the surrounding schools. So we run an analysis that controls completely for location. We compare pupils living in the same place, neighbours, but varying in poverty status[4]. A full unit postcode contains on average about 15 (contiguous) dwellings, although obviously only some of these will house families with secondary school-aged children. We exploit this within-street variation and also control for other personal characteristics including prior test scores. Children from poor families are significantly less likely to go to good schools. The difference is 2 percentage points, relative to an overall difference in that sample of 14 percentage points. This result, which is likely to be a lower bound on the effect, shows that location is not the only factor reducing the chances of poor children from attending good schools. This difference captures the widely discussed "working the system" by more affluent parents.

We focus in on a particular case of the pupil-school assignment to examine whether a pupil attends her nearest school, as a function of its quality and her characteristics. The striking result is that while non-poor families exhibit the expected behaviour – a higher probability of attending the school the higher its quality – this is not true for pupils from poor families. This result is robust to controlling for a very wide set of controls for distance and neighbourhood.

The paper makes two contributions to the evidence on school choice. First, we re-run the within-postcode analysis by decile of the feasibility of choice. This is measured by

[4] We discuss the selection issues involved below.

each pupil's minimum distance to reach three schools. We show that the lower chance of poor children attending a good school is essentially unaffected by the degree of choice. It neither improves with choice as supporters would suggest (through the lesser importance of location), nor deteriorates as opponents fear (through more covert selection by schools). Second, we can compare the importance of location in explaining the relative chances of children from poor families. To be precise, we ask how important is "choice of location" relative to "working the system" in the strategies used by the non-poor to get their children into good schools. It is important to be clear on the interpretation of this finding that location matters. We are not claiming that location is randomly allocated across families, and that this is a causal relationship between location and school assignment. It is not surprising that distance/location matters as it is in admissions procedures. Rather, the analysis describes the strategies of parents to achieve the school outcome they want, given the assignment rule and their demand for a school place. The optimal responses to assignment rules that privilege location and (implicitly) some pupil attributes, are to acquire the right location, or to "work the system" to make the pupil attributes clear. This split between location and other factors tells us about the main strategies that parents employ.

Concerns about equity in the schooling system are of course very long standing. Furthermore, the idea that school choice might be good for the disadvantaged is also far from new (see for example Jencks, 1970). But more recently most of the choice research has been focussed on the impact on average attainment, or sorting of pupils across schools[5]. Recent empirical contributions divide into those studying specific targeted choice schemes (for example, Cullen, Jacob and Levitt, 2006), and those examining a generalised system of school choice (for example, Hoxby, 2000, in the US and Sandstrom and Bergstrøm, 2002, in Sweden). We can also characterise studies as analysing the partial equilibrium impact on the pupils making choices (for example, Howell, 2004 on the New York school choice programme), or focussing on the impact on the school system as a whole (Hoxby, 2000, Hoxby, 2003b, and the controversy between Rothstein, 2005, and Hoxby, 2005). Bayer and McMillan (2005)

[5] Recent contributions surveying the field include Howell and Peterson (2002), Hoxby (2003a), Ladd (2002) and Neal (2002).

model the general equilibrium of residential and school choice. Lavy (2006) seems to be the only study looking at both effects on individuals and on the system as a whole.

There are very few studies on choice for Britain. While Bradley et al (2000) find some positive impact of competitive links between schools' exam results, three recent studies find little evidence of a strong impact of choice on overall standards. Clark (2004) finds that schools located near to "opting-out" schools, arguably facing increased competition, did not respond by improving outcomes. Gibbons, Machin and Silva (2006) use a cross-section of primary schools and instrument their measure of competition using distance from the market boundary; they find little effect of competition on outcomes. Burgess and Slater (2006) exploit a change in the boundaries of local education markets to generate an exogenous change in the degree of choice. They find some hint of an effect of the change on standards in the expected direction, but not a statistically significant effect.

Looking specifically at educational outcomes for children from poor families, there has been considerable work in the US on the impact of voucher schemes for disadvantaged children (eg Howell, 2004, and some of the papers collected in Hoxby, 2003a). Much of the recent work in the UK on the importance of socio-economic background has focussed on higher education (for example, Galindo-Rueda et al, 2004, and Vignoles and Machin, 2006). In secondary schools, there has been interest in the role of ability selection in grammar schools (Galindo-Rueda and Vignoles, 2005) and the relationship between income sorting between schools and choice (for example Burgess et al, 2006, Allen and Vignoles, 2006). But there appears to be no analysis of the role of family poverty in the assignment of children to schools in the majority of education markets in England which do not select on ability.

The next section sets out an economic modelling framework for our approach, and the following section details the data used. Section 4 presents the results and section 5 concludes.

2. Modelling Framework

We model the assignment of children to schools, as a function of the characteristics of the school and of the children. The observed assignment is a realisation of an underlying process composed of two decisions: applications by parents and children for places in particular schools (demand), and the administrative procedures that allocate children to schools given their choices (assignment rule). This is like many matching problems – how does the labour market allocate workers to firms; how do particular matches form in the marriage market? We discuss the economic analysis of these processes below. For studies of school assignment explicitly based on a two-sided matching approach see Abdulkadiroðlu et al (2005) on Boston and Abdulkadiroðlu et al (2005) on New York.

a) Notation

In an area, there are *S* schools denoted *s*, and *P* children denoted *i*. A child's poverty status is measured by her Free School Meals (FSM) eligibility, denoted f_i. The school average FSM eligibility is $\overline{f}_s$. A child's score in the tests at the end of compulsory schooling (known as GCSEs) is q_i, and her score in the test just before secondary school (Key Stage 2) is our measure of prior ability, denoted k_i. The average GCSE score of school *s* for time/cohort *t* is $q_{s,t}$. We take this as the public quality score. We define a 'good school' as one in which the average quality score is above some threshold. In the empirical work, we take the top third of the distribution of $q_{s,t}$.

A pupil's location is L_i. Denote pupil *i*'s nearest school as *n(i)* and pupil *i*'s actual school attended as *a(i)*, The distance between pupil *i* and school *s* is d_{is}. This paper analyses the outcome of the processes that map the characteristics of child *i* to the characteristics of school *Z* if $a(i) = Z$.

In particular, we focus on the quality of school assigned to pupil *i*. As noted above, the quality score for a school *s* at time *t* is the school mean GCSE score for the cohort finishing in *t*, $q_{s,t}$. We write the quality of the school to which pupil *i* from cohort *t* is assigned as $q_{a(i,\, t),\, t\text{-}6}$. This emphasises that the quality score was achieved by a prior cohort of pupils – given the timing of decisions and the duration of compulsory secondary schooling, a cohort 6 years older than the entering one.

This matters because of the way that the quality score is generated. We assume that an individual pupil's score depends on her own characteristics and the value-added provided by the school. The most important characteristics are likely to be the pupil's prior attainment and poverty status. So for a pupil: $q_i = a.f_i + b.k_i + v_s + u_i$, where v_s is the school's value-added and u_i is unobserved factors and testing noise. Adding up over the school for pupils in cohort t gives: $\overline{q}_{st} = v_s + a\overline{f}_{st} + b\overline{k}_{st} + \overline{u}_{st}$. This quality measure therefore reflects both the school's value added and its peer group (as measured by $\overline{k}_{st}$ and $\overline{f}_{st}$). In terms of the child-school assignment mechanism, the school composition that the quality score and therefore assignment depends on is one from six years before the assignment decision.

b) Demand and the Assignment Rule

Modelling the general equilibrium of a process involving simultaneous choice of location and school is complex, see for example Bayer and McMillan (2005). Furthermore, if peer groups are important in influencing educational outcomes, then analysing school assignment means searching for an equilibrium in a complex game (see Epple and Romano, 1998, 2003). Typically, simulations are required to capture the salient features of the model (Nechyba, 2003, 2004). These studies, modelling the US education system, necessarily ignore important features of the school market in England. Perhaps the key feature is that popular schools in England cannot expand much to deal with a lot of applications, and so over-subscribed schools have to choose pupils. Thus we cannot simply import these models to study the school market in England. We sketch out the nature of the processes underlying the reduced form assignment function that we estimate below.

We assume a family's utility depends on the present value of future income generation from educational achievement, on income left for consumption after school-associated costs, and on other aspects of the school experience. Thus demand for a school place reflects a demand for teaching quality, school ethos, peer group characteristics, distance from home, and other facilities. The resulting choice of school will depend on the family's preferences, family structure (how many children for example), the child's ability, income, and the prices of complements (school uniform, travel time) and substitutes (private schooling). State schooling itself in England is free. The demand is expressed as applications to schools: students make applications to a small number of different schools (typically 3 to 6).

The administrative procedures are complicated in England in that some schools act as their own admissions authority[6], but most schools have their admissions administered

[6] Foundation and Voluntary-Aided schools.

by the Local Education Authority (LEA). Schools cannot change size very rapidly, so a popular school with more applications than places cannot expand within the period to provide more places. In this case the admissions authority (school or LEA) has to choose the pupils to accept. Schools have incentives to try to pick the more able pupils, since as we have noted above, the quality score for schools is affected by their pupil intake.

Demand and the assignment rule together deliver an allocation of children to schools. This is the map from pupil characteristics to school characteristics. We do not in this paper attempt to separately identify the role of demand and the assignment rule. So a finding that characteristic *x* influences the outcome could arise because that characteristic influences demand and/or the assignment rule.

We characterise the outcome of the allocation in a general form as:

$$q_{a(i,t),t-6} = f\{g(L_i), f_i, k_i, Z_i\} + e_i \qquad (1)$$

where Z_i captures other characteristics of the pupil such as ethnicity, and e_i denotes unmeasured factors, discussed below. The term $g(L_i)$ represents all the relevant features of the location:

$$g(L_i) = g\left(\mathbf{W}_i; d_{1i}, d_{2i}, .., d_{Si}; q_{1,t-6}, q_{2,t-6}, .., q_{S,t-6}\right) \qquad (2)$$

where **W** measures all the characteristics of the area. Such factors include the nature of the neighbourhood and the neighbours, transport links, the reputation of an area, and the presence of local amenities that may induce unmeasured differences in the characteristics of the families living there. *g()* also contains the set of distances from pupil *i*'s home to each school in the area, and the corresponding quality score of each school. Distance may well influence demand in terms of minimising travel time. It also plays a prominent part in the assignment rules. Most schools and LEAs have school/home proximity as a key criterion for allocating scarce places. One potential outcome is that pupils are allocated to their nearest school: $a(i, t) = n(i, t)$. This is the essence of the neighbourhood schooling assignment rule, as a leading alternative to a choice-based assignment rule. In this case the role for demand is in the choice of location. The role of distance in influencing school admissions is often expressed in terms of catchment areas. However, these are not very useful for analysis as realised catchment areas are endogenous. A school in high demand will have to draw a catchment area very tight around the school, whereas a less popular school will have a much broader area. We rely on measured distances.

The remainder of *f()* contains characteristics of the pupil. Again, these may influence both demand and assignment. Explicit selection of pupils on the basis of ability or income is not permitted in the admissions code, but it is well established[7] that schools engage in certain practices (such as interviewing) to implicitly establish ability or family background. The measure of ability we use is performance in the Keystage 2 test taken just before entry into secondary school (details in the data section). The timing of this is such that the outcome cannot be known before the school assignment decision is made. Thus its role is as a measure of general ability, not an explicit factor in the allocation decision.

Finally, the term e_i captures other factors that influence demand or assignment. For example, admissions authorities use the presence of siblings in a school as an important criterion in allocating subsequent children. This factor is assumed to a first approximation to be orthogonal to ability and poverty.

The terms on the right hand side of (1) will likely be correlated amongst themselves. We know that family background affects exam outcomes, so that will generate a correlation between f and k. Distance from good schools is likely to be correlated with f through the working of the housing market, so this is a little more complex. All this implies that not controlling appropriately for k and L will generate a correlation of $q_{a(i, t), t-6}$ with f.

c) Reverse causation?

In the empirical work below, we model the outcome of the assignment process given by (1). We interpret the estimated relationship between the school's quality score $q_{a(i, t), t-6}$ and a student's personal characteristic, f_{it}, as representing the outcome of the assignment process. But we address the possibility that there is an alternative basis for the correlation, namely from student characteristics to the outcome score.

[7] For example, West (2003): "Ongoing analysis suggests that one in five secondary schools used overtly selective criteria (e.g. partial selection on the basis of ability/aptitude, primary school record) or potentially discriminatory criteria (e.g. priority to children of school employees/former pupils/governors) or subjective criteria/practices allowing for administrative discretion (e.g. interviews, compassionate/pastoral factors)."

It is clear that there is no straightforward reverse causation because the quality score relates to the performance of a cohort of children in the school six years before the cohort studied here: we are not simply regressing the score of a group of pupils on their own characteristics.

But there may be a lot of persistence in school attendance. There are two ways in which this might arise. First, suppose that schools were located on "islands", with little or no mobility between them. All students from succeeding generations therefore go to the school on their island. If this were true, then the allocation outcome would be trivial: all would go to their island school ($a(i, t) = n(i, t)$) and $q_{a(i, t), t-6} = q_{n(i, t), t-6}$. In terms of relating pupil characteristics to school characteristics, everything would hinge on which island people lived on.

However, this does not describe the situation for secondary schools in England, though it may be a closer approximation to the way that school districts operate in the US. We have shown (see Burgess et al, 2006) that 54% of children do not in fact attend their nearest school. Furthermore, 28% do not attend one of their nearest three schools. Thus there is a huge amount of mobility from neighbourhoods to different schools. The excess distances travelled to attend non-local schools are not trivial. On average this is 2.3km, relative to a mean distance to nearest school of 2.8km (median 1.7km). This complex pattern of school attendance can be illustrated by an example in Birmingham. We take five regular schools (comprehensive, mixed gender, non-faith). Using precise data on pupil location (see below), we can rank pupils in terms of distance from their school attended. In Figure 1, we plot the 50% distance contour for each school – that is, the line that encompasses the nearest 50% of that school's intake. The key point to derive from the Figure is that these lines overlap to a very considerable extent – the "island" story does not describe England. There may be rural areas where it is more applicable, but even in rural areas, only 59% of pupils attend their nearest school.

Persistence in school attendance may arise in a second way. This might be described as a "dynasties" argument. Even without the rigidity of the "island" model, pupils living in particular locations always go to the same school. Furthermore, because of the operation of the housing market and the established persistence in area poverty, particular locations always house poor families. The poverty of succeeding generations is correlated and the exam score of one generation of pupils drawn from that area is correlated with the poverty of the next. Thus a correlation is generated between the poverty status of a pupil in cohort *t – 6* going to the focus school, that school's score and the poverty status of the pupil in cohort *t*. In econometric terms, estimating:

$$q_{a(i,t),t-6} = a + b.f_i + c.k_i + \mathbf{d.Z}_i + u_i \tag{3}$$

will yield a biased estimate of *b* because the nature of the location is un-controlled for, is correlated with f_i, and with the poverty of the previous cohort of pupils who generated the school quality score. The answer therefore is to control for the characteristics of the area, L_i. Given the "dynasties" argument, this necessarily captures the nature of the locations that the previous attendees of the focus school came from. If they came from different locations, then the correlation does not arise in the first place.

d) Good schools and local schools

We specialise the analysis of the $q_{a(i,\,t),\,t\text{-}6}$ -- f_i assignment map to the question of whether a pupil attends her nearest (local) school. The key issue is the quality of the school, and the pupil's FSM status. In particular, we are interested in the interaction of the two – whether the impact of FSM differs depending on the quality of the local school. The probability that $i \rightarrow n(i,t)$ we write as p^*_{it}, and set out as follows:

$$p^*_{it} = a + b.f_i + c.q_{n(i,t),t-6} + f.f_i.q_{n(i,t),t-6} + \mathbf{d.Z}_i + g(L_i) + e_{it} \tag{4}$$

Again, we include controls for the local area and neighbourhood, as well as other individual pupil controls. As with the analysis discussed above, it remains true that we do not attempt here to distinguish demand for school places from the assignment rule. Thus a finding that particular pupils do not attend their local school may indicate demand or may indicate the operation of the assignment rule through the rationing of the available places. To reiterate, equation (4) represents the outcome of the assignment rule and parental response to that, not a causal relationship between location and school entry.

3. Data

We combine administrative data from the Department of Education and Skills (DfES) with geo-coded data from commercial and government sources. We focus on the secondary school that children join at age 11[8].

a) Data on Pupils

Our core dataset is the Pupil Level Annual School Census (PLASC), kindly provided by the DfES. This is a census of all children in state schools in England, taken each year in January. This was first carried out and made available to researchers in 2002, and we use the three PLASCs currently available. PLASC provides a number of personal characteristics, and can also be linked to other items from the DfES's National Pupil Database (NPD), including each pupil's test score history.

Data on pupils includes the following characteristics: gender, within-year age, ethnicity, and an indicator of Special Educational Needs (SEN, which measures learning or behavioural difficulties). The key variable for our purposes is an indicator of family poverty, the eligibility for Free School Meals (FSM). This is dependent on receipt of state welfare benefits: eligibility for Income Support or Unemployment Benefit brings eligibility for FSM. Income Support largely captures single parents but also includes support for disability. Of course, welfare receipt and so FSM status is a crude, dichotomous measure of poverty, and we should be cautious about simply comparing pupils with FSM across different areas of the country and different micro neighbourhoods. The neighbourhood data discussed below helps here. Measurement error may also be an issue. Our FSM variable is probably a very good measure of the FSM status of children[9], though the FSM variable is a noisy measure of poverty status[10]. The proportion of children eligible for FSM is less than the proportion counted as poor using standard measures, so FSM children are likely to be poorer than the average poor child.

We use test score data from one of the Key-stage tests that students take throughout schooling. We use the Key-stage 2 test taken at age 11 as the pupils finish their spell

[8] There are a few areas with a middle school structure, which we omit.
[9] See Barker, 2006, for a discussion of the multi-stage checking of the schools' data: http://www.bris.ac.uk/Depts/CMPO/PLUG/userguide/anna.ppt
[10] For some evidence, see http://www.bris.ac.uk/Depts/CMPO/PLUG/events/130906/vignoles.ppt

in primary school. This is a nationally set group of tests (in English, Maths and Science), marked outside the school. This represents a measure of ability, and will also be correlated with parental resources.

b) *Data on Schools*

To characterise the quality of secondary schools, we choose the publicly available and widely quoted measure of the proportion of a school's pupils achieving grades A* to C in at least 5 GCSE exams at age 16. These exams are important, are nationally set and come at the end of compulsory schooling. Typically a pupil takes exams in 8 – 10 subjects. Tables showing each school's score are published in the national and local press each year. Until recently, these were the only real quality information available, but a number of value-added tables are now also published (see Wilson, 2003). These exams scores reflect both the teaching of the school (value added) and the composition of the school (peer group). Thus basing a decision on the %5A*-C can be thought of as basing it on a weighted average of value added and peer group. We define a good school as being in the top third nationally of the distribution of %5A*-C scores. The dating of this score is important – we use the score for each school from the time that the cohorts made their decisions on school applications, so deriving from the results of a cohort of pupils 6 years older. We consider alternative definitions of good schools as a robustness check.

c) *Location and Distance*

Crucially for this analysis, we have access to each pupil's full postcode. This locates them quite precisely, to within 100m. We also have the coordinates of the school, which locates it exactly. We rely on the postal geography of the UK[11] for this analysis. Overall, there are about 1.78m unit postcodes covering 27.5m addresses[12]. On average, there are 15 addresses in a unit postcode. A subset of these addresses will house families with children attending state secondary schools. In our data, in any one cohort we have over 333,000 distinct unit postcodes, covering half a million children, and 544,320 distinct unit postcodes in all.

---

[11] For further details see http://www.statistics.gov.uk/geography/postal_geog.asp (accessed 24/3/2006)
[12] As of May 2005, in the UK as a whole.

Using pupils' postcodes, we match in data on neighbourhoods. These measures of neighbourhood fulfil two roles: they measure the deprivation of the neighbourhood and home peer group, and also provide an additional factor proxying the individual's own household context. We have data at two different scales. Firstly, we have matched pupil's postcodes to the Mosaic classification[13] of that address. Mosaic classification is a postcode level dataset that categorises each postcode in the UK into one of 61 different types on the basis of demographics, socio-economics and consumption, financial measures, and property characteristics and value. Over 400 variables are used to construct these classifications and as such this provides a rich picture of pupils' neighbourhoods at a very local level. Secondly, we use the Index of Multiple Deprivation (IMD) produced from administrative data[14]. This ranks every ward[15] on a range of criteria (income, employment, health, education and skills, housing, and geographical access to services); we use the overall weighted index.

Distance can be measured in a number of different ways. Given the size of this dataset, it is only computationally feasible to use straight-line distances. These are computed from coordinates and the application of Pythagoras' theorem. This is inferior in the sense that it will fail to take account of natural barriers such as rivers or hills, and it will also misrepresent distances where the road network is not very dense, but some experimentation suggested it was not too inaccurate[16]. We use this information to identify each pupil's nearest school.

d) Sample selection decisions

We take the cohort of new entrants into state secondary schools from each PLASC, that is, pupils in their first year of secondary school. There are roughly 0.5m pupils in each cohort, so our full sample is 1,566,415 pupils. We specialise this to particular sub-groups in the analyses below. We distinguish between selective and non-selective LEAs, defining the former as having at least 10% of pupils in grammar schools, and

[13] This is commercial geo-demographic data, kindly provided to us by Experian. For more information see http://www.experian.co.uk/business/products/data/113.html.
[14] For more information see http://www.odpm.gov.uk/stellent/groups/odpm_urbanpolicy/documents/page/odpm_urbpol_608140.hcsp.
[15] A ward is a small geographical unit, containing on average around 12,000 people.
[16] In fact, we compared this straight-line method with the travel distance method for three areas – a rural area, an urban area outside London and a London LEA. We identified each pupil's nearest school using both methods. In each of these areas, the correspondence was around 85%. It therefore seems that the approximation given by the straight-line method is reasonably accurate.

omit selective LEAs. This cuts out 13.4% of the pupil total, leaving just over 1,356,000. We omit pupils from some special schools, and pupils are omitted if they have missing values for data. The sample for the overall regressions in Table 2 is 1,239,888 – some 91% of the available total in non-selective LEAs.

The means and some indication of variation of all the variables used are in Table 1.

4. Results

We present four sets of results. First, we quantify the relative chances of children from FSM-eligible families attending a good school. Second, we exploit within-postcode variation to control completely for location and to isolate any remaining influence of poverty status. Third, we focus on one aspect of location – the feasibility of school choice. Fourth, we analyse which pupils attend their local school, as a function of their own characteristics and the school's characteristics.

a) Overall Picture

Overall, 29% of children have places in good schools using our definition[17]. Define for each LEA the ratio of FSM-eligible children in good schools to all FSM-eligible children, denoted g^f. Similarly, the equivalent ratio for non-FSM-eligible pupils is g^{nf}. The means and quantiles of these variables are in panel B of Table 1. In Figure 2 we cross-plot g^{nf} and g^f for the non-selective LEAs. We see that in all LEAs the latter is below the former (bar Sandwell, just above), and that there is a strong correlation between the two. There is also considerable variation in the g^f / g^{nf} ratio. A simple regression across LEAs on g^f / g^{nf} shows it to be lower in poorer areas and lower in London and other urban areas.

This establishes at LEA level that children from poor families as a group have (unconditionally) lower probabilities of attending a good school, we need to use the micro data to understand where this comes from. Our econometric strategy is not to identify school demand in this paper. Given the basic structures of the problem (parents' demand for school places, and an assignment rule), parents then formulate their response strategy to the assignment rule in the light of their preferences and characteristics. The most relevant part of the explicit assignment rules is the role of location. Other factors that may implicitly matter to schools include a pupil's ability

[17] We take the top one third of <u>schools</u>, which does not translate exactly into one third of <u>pupils</u>.

and background. Parents will attempt to choose location and to make any implicit advantages of their children visible to the admissions authorities. Our strategy is to isolate how much of the difference in outcomes works through location, and how much through other channels, controlling for location.

We estimate the likelihood that a child is assigned to a good school using the three combined entry cohorts and the definition of a good school as discussed above. Table 2 reports the probit results for three specifications. The first specification simply controls for FSM status and LEA and cohort dummies. FSM status is negative and strongly significant. Computing the marginal effects and taking the average, we see that children from poor families are 17.3 percentage points less likely to attend a good school. This is equivalent to about half the chance of non-poor families.

Column 2 adds in personal characteristics. The coefficient on FSM status falls substantially to 0.443, and to an average marginal effect of 13 percentage points. We also see that pupils scoring highly at the KS2 are more likely to be assigned to a good school. The interpretation of this is not that there is explicit selection on the KS2 scores, as these are not published when school assignment decisions are made. This effect is capturing correlation of KS2 with general ability and/or family background. We discuss the gender coefficient in the next sub-section, and the ethnicity coefficients shortly.

Turning to column 3, we add a rich set of neighbourhood controls. We use the Mosaic classification for a characterisation of the very immediate neighbourhood, and the ward level IMD to describe the broader area. We also include the full set of interactions of these two. These all add significantly to the explanatory power of the model (raising the pseudo-R^2 from 0.123 to 0.228). Notably for our purposes the coefficient on FSM status more than halves to 0.178. The mean marginal effect of FSM status falls from 17.3 points in column (1) to 4.8 points in column (3)[18]. Finally we return to the role of ethnicity in school assignment. Column 2 shows that holding other personal characteristics fixed, pupils of Black Caribbean heritage, other Black heritage, Bangladeshi or Pakistani ethnicity are less likely to be assigned to a good

[18] Using the column 3 estimates, we have explored some of the variation in the marginal effects, available from the authors. The smallest impact of FSM is for an FSM-eligible child in a poor neighbourhood in a poor LEA. There are larger impacts for pupils in richer areas, where being one of the fewer poor children has a greater effect on your likelihood of going to a good school.

school. However, once we control for neighbourhood characteristics, all of these effects reverse, apart from that for Bangladeshi ethnicity children.

In table 2b we re-run these regressions with the additional variable of distance from the pupil's home to the nearest good school. As with the other location variables, we are not claiming that this is exogenous. In fact, its inclusion reduces the coefficient on FSM status a little, but not a huge amount given the inclusion of other location controls.

Returning to the main argument of this section, a comparison of columns 1 and 3 suggests that about two thirds of the unconditional effect of FSM is related to location. Note further that since all these analyses include LEA dummies, this is all about within-LEA variation. Children from poor families face a reduced chance of being assigned to a good school in large part because of where they live.

b) Within-postcode variation

It is clear that distance and location matter for school assignment. The only way to control fully for location is to compare individuals living in the same place – that is, holding all location-related factors constant. We can do this in a very precise way with our data. The full unit postcode locates individuals to just a handful of dwellings, and this allows us to exploit within-postcode variation in school assignment and FSM status. Sample selection issues are discussed below.

About two thirds of secondary school pupils share their unit postcode with at least one other. Table 3 compares the sample of such pupils with the full dataset. They are marginally more likely to be poor, and have slightly lower KS2 scores. This is because single-pupil-postcodes are more likely to be found in richer, rural areas. Table 4 shows the structure of the data. Over 400,000 live in a 2- or 3-pupil postcode, in some 200,000 postcodes. Even in the postcodes with just two children, they both attend the same school in only 61% of cases. For the others, the cases where all attend the same school is the minority. This gives further illustration of the highly complex spatial patterns of school attendance (see Burgess et al 2006). The picture of clear and distinct catchment areas does not describe the situation in England at all well. There is less but still considerable variation in poverty status in multi-child postcodes; there are around 280,000 postcodes (involving over 1m pupils) with variation in FSM eligibility. Overall, in 40% of postcodes with 2 or more pupils in there is variation in FSM status.

Using this data, we ask: for two pupils who are next door neighbours, one from a poor family and one not, are they likely to go the same quality school or not? Before presenting regression results, Figures 3 and 4 provide a view of the data. The top left panel of Figure 3 takes all two-pupil postcodes. Dividing these locations into those where both pupils are FSM-eligible, where neither are, and where one of the two is, we plot the mean school quality separately for FSM and non-FSM pupils. The remaining panels of the figure repeat this exercise for 3-, 4- and 5-pupil postcodes, providing more overlap of the two groups. A number of patterns are very clear across these groups. First, within each postcode type, FSM-eligible pupils attend lower scoring schools than do ineligible pupils. The gap is consistent across all the panels at about 3 percentage points. Second, the poorer the postcode, the lower the average school quality that its inhabitants attend. Crudely comparing the two relationships, it appears that the difference in mean school quality between poor and non-poor students within a postcode type is about equivalent to the difference in mean school quality between postcodes that are 40% poor and those that are 60% poor. Third, the average school quality of the data point at 100% poor declines marginally as we look across 2-pupil, 3-pupil etc postcodes. This is because higher-pupil postcodes are denser, more urban, and therefore on average poorer.

Figure 3 compares across postcodes; Figure 4, like the following regressions, exploits within-postcode variation. The top left panel takes all 2-pupil postcodes. Within each postcode we order the pupils and compute the difference in their FSM status and equivalent difference in the quality of school assigned to them. In this case the former can take the values –1, 0 or 1. So for each postcode we have the within-postcode FSM difference and the within-postcode quality difference. In the figure we plot the distribution of these differences. The panel shows that poor children are assigned to less good schools, both at the mean and the quantiles shown. The bivariate regression coefficient is –0.024. The next panel take 3-pupil postcodes. For each postcode we order pupils, take all pairwise differences and average. The other panels use 4- and 5-pupil postcodes. The story is the same across all panels, showing a negative relationship.

The regression results pooling using all the observations are in Table 5. Column (1) includes just FSM status and column (2) adds other pupil characteristics. We see that children from poor families do go to lower scoring schools on average. This effect is around 1.5 percentage points, around 9% of the standard error of school quality in this dataset. Standard errors are clustered at postcode-cohort level. In terms of the other variables, differences in gender have no effect. This is interesting, and contrasts with the estimated lower chance for a boy of attending a good school found above, though it may be due to a small number of single-sex schools being dropped in this analysis. Comparing these results suggests a correlation between where families with boys and girls live relative to the location of good schools. Age has no effect, and children from minority ethnic backgrounds tend to attend better schools.

It is useful to compare the impact of FSM eligibility on the chance of attending a good school across the different specifications. In Table 6 we run OLS regressions with fixed effects on the earlier (dichotomous) dependent variable: attend a good school or not. Whilst not perfect, this seems the most transparent technique given the need to deal with over 100,000 postcode fixed effects. We successively add personal and neighbourhood characteristics alongside LEA fixed effects, and then add in postcode fixed effects. The first 3 columns are OLS-FE versions of Table 2 (on a slightly different sample because of missing data). Accounting only for personal characteristics and LEA dummies, poor children face a lower probability of around 10 percentage points. Once we add in postcode dummies capturing all location effects, this falls to 2.2 percentage points. We take two points from this – most of the reason for poor children's lower chances is accounted for by where they live, but not all of it.

Using within-postcode variation potentially brings a selection effect, but one for which the bias can be signed. Assume dwelling-specific house prices within a unit postcode are the same. We would expect FSM-eligible households living in the same street as ineligible households to be among the better off of such households. Similarly, FSM-ineligible households living next door to FSM-eligible families are likely to be relatively poor compared to other FSM-ineligible households. Thus the income differences between households of different FSM status and living in the same street are likely to be lower than unconditional income differences between households of different FSM status. If the link between FSM status and school assignment is estimating a relationship between household income and school assignment, our estimated differences are likely to be an underestimate of the true relationship.

Summarising, these results control completely for location and suggest that children from FSM eligible families on average attend schools that rank less highly in the school league tables. This analysis takes into account all aspects of a location, including distance from schools, and neighbourhood characteristics. Reverse causation seems very unlikely, because the measure of quality used is essentially unrelated to the performance of the children in the postcode. First, the measure relates to a cohort of children passing through the school six years previously. Second, even if generations of children from the same house went to the same schools, in these within-postcode comparisons the focus children clearly constitute a negligible fraction of the actual attendees of the schools (around 5%). Third, the use of within-postcode variation controls for any location effects.

c) Robustness: different definitions of good schools

For robustness, we consider alternative definitions of a good school. First, we take the top third of the distribution within an LEA rather than nationally; second, we define good in terms of value-added rather than GCSE league table performance; and third, we combine both of these. The first is largely about whether parents look at the national or LEA comparators. The LEA effects included in the earlier results would have taken out any first order linear effects, but there may be important interaction effects too. Changing to define a good school based on value-added is more of a substantive difference, and implies slightly different questions. The broad question we are addressing here is to what extent children from poor families attend good schools. Using GCSE league table scores as the basis for defining a good school includes the two aspects of what schools provide: teaching effort (value added) and peer group quality (prior attainment). Clearly, a definition based on value added alone will produce a different ranking, and may change the findings. The question is also different in another sense. GCSE league table scores are widely publicised, and it is reasonable to assume that parents are broadly familiar with them. Value-added is not (until very recently) published, and the argument would have to be that parents know through informal means whether a school provides good teaching; this may not be a very accurate process. So the mechanism linking parental preference for schools to school quality is different.

We re-run the core specifications from Tables 2b and 6 using these alternative definitions of good schools. The results are in Table 7, along with the original results from Tables 2b and 6. In both, we see that the impact of FSM status is stronger at LEA than national level, and that the coefficient is lower using value-added than GCSE. Focussing on panel (b) of the Table, comparing column 5 (postcode fixed effects) and column 2 (just individual characteristics) gives us the relative importance of location and other factors. Varying the definition makes little difference: the column 5 coefficient as a fraction of that in column 2 is 22.0% for our basic definition, 21.8% using GCSE and the LEA distribution, 22.2% using value-added and the national distribution, and 21.8% using value-added and LEA distribution. We would argue therefore that the main results are robust to different definitions of the key variable of a good school.

d) Role of Choice

One of the factors that relates to a location is the feasibility of school choice. In this sub-section we examine whether the penalty to FSM eligible children varies with the degree of school choice. The arguments suggesting that the penalty may be more or less in areas with high degrees of choice were set out above.

We approach this in three different ways. First, we re-run the overall national regressions in Table 2, interacting FSM status with our measure of the degree of choice: whether the pupil has three schools within 2km of her home. Second, we run these regressions separately for pupils in London, in other urban areas, and in rural areas[19]. This split is highly correlated with the feasibility of choice. Third, we re-run the within-postcode regressions separately for ten groups of postcodes, split by deciles of choice.

Table 8 presents the results for the first two of these analyses. The interaction between FSM status and pupil choice is negative and significant. The interpretation of this is that conditional on the other observable characteristics, FSM-eligible pupils have a lower chance of attending a good school in an area where choice is high than in an

[19] A school is defined as urban/rural if it is in an urban/rural Local Authority District, and in London if it is in a London LEA.

area where choice is low. Columns 2 to 4 of the table confirm this finding. The FSM difference is higher in London than other urban areas and higher again than in rural areas.

However, the characteristics of the FSM population may well differ in relevant ways across these areas, affecting the results. While the inclusion of the full set of Mosaic dummies, broader IMD and their interactions along with LEA dummies should absorb a lot of that heterogeneity, the best way to control for location is to use the within-postcode variation. We re-run the postcode fixed effects regressions on the data split up by choice feasibility in Table 9. The feasibility of choice is measured by the minimum distance to reach three schools, with the lowest radius indicating the greatest feasibility of choice (see Burgess *et al*, 2006, for more details). The pattern of FSM coefficients is flat until the last two deciles of choice, when it moves towards zero (though not significantly so). Controlling for location, the lower school quality assigned to poor children does not vary much in association with the feasibility of choice.

e) Going to the Local School

We have argued that the pattern of children being assigned to schools is complex (recall Figure 1). The regression evidence above summarises outcomes from that process. In this section, we specialise the analysis to a particular question: what influences the probability that a child attends her nearest school? The relationship of going to the local school, the quality of that school ($q_{n(i,\ t),\ t-6}$) and the pupil's characteristics are shown in Figure 5. The quality variable on the horizontal axis is split into ventiles, and the vertical axis marks the fraction of pupils going to their local school for that quality of school. Panel A plots the relationship separately by FSM status. A number of points are clear. For non-FSM eligible pupils, there is a strong upward trend: the better the local school, the more likely are such pupils to attend. The probability ranges from less than 0.3 for the lowest score schools, to around 0.6 at the top. The final data point at the highest quality school relates almost entirely to schools with additional admissions criteria, principally faith schools. They draw pupils from a much wider catchment area. There is a strong contrast with FSM eligible pupils: over the top 75% of the distribution, the chance of a poor child attending the local school declines slightly for increasing school quality. In Panel B, we split pupils in a different way. We rank the 61 Mosaic postcode types by their national % FSM, and plot the relationship for those living in poor neighbourhoods (bottom third of Mosaic codes) and affluent neighbourhoods (top third). The same picture is clear. Affluent households living near a poor school, with high probability commute their child to another school. Over bulk of the range, poor households do not vary much in the rate of going local, even declining in the top third of quality.

But these graphs are unconditional, and it may be that some version of the “dynasties” argument means that persistence in school assignment produces a correlation between school quality and pupils’ FSM status. We report the results of estimating equation (4) in Table 10. In addition to the usual pupil characteristics, column (1) includes the quality of the local school (in ventiles), and this interacted with FSM status. Column (2) adds the full set of local area and neighbourhood controls from Table 2 column 3. Column (3) further adds: a dummy for whether the pupil’s nearest school is located in an urban area; a dummy for whether the pupil’s nearest school is located in a London LEA; interactions of school quality group with the pupil’s KS2 mean score; the distance of the pupil to her nearest school, distance squared, distance cubed and all these distances interacted with FSM status.

The key results are the interaction of local school quality and the pupil’s FSM status. The influence of quality is clear: higher quality is associated with a higher probability of attending the school. The interaction of quality and FSM shows a declining pattern, mirroring the graphs in Figure 5. Moving across the columns, this pattern is marginally flatter in column 3 with all the controls, but remains strong. The fitted pattern derived from column 3 is plotted in Figure 5, panel C.

We repeat that we do not interpret location here as exogenous. Location is chosen by some parents to influence school assignment, and “living near a good school” is an objective of that decision. Note that this will likely interact with the varying feasibility of choice around England. We control for that with the distance and other neighbourhood controls in Table 10.

5. Conclusions

We study the assignment of children to schools. The particular focus is the assignment of children from poor families to schools featuring high up in the published league tables. The primitives of the problem are parents’ demand for school places, and an assignment rule to resolve allocation at over-subscribed schools. Parents then formulate their response strategy to the assignment rule in the light of their preferences and characteristics. The most relevant part of the explicit assignment rules is the role of location. Other factors that may matter to schools include a pupil’s ability and background. Parents will attempt to choose location and to make any implicit advantages of their children visible to the admissions authorities. Children from poor families may stand a lower chance of attending a good school for a number of reasons. First, where they live; second, because over-subscribed schools find ways of choosing pupils according to their incentives and this correlates with FSM status; third, because middle class parents are better at working the system of school admissions; or fourth, the costs of exercising choice (acquiring the information, transport costs) may be prohibitive. Our strategy is to isolate how much of the difference in outcomes works through location, and how much through other channels, controlling for location.

In a typical LEA in England a child from a poor family is half as likely to attend a good secondary school as a non-poor child. Much of this is due to where they live within the LEA. But location is not all: comparing children living as neighbours, and controlling for observable differences, the poorer neighbour is less likely to go to a good school. This gap is 2 percentage points, compared to an overall gap of 14 percentage points. Thus location accounts for most of the gap, but not all.

We can compare across areas of England with differing degrees of choice. Controlling completely for location using within-postcode variation, the FSM-differential is relatively flat once we control fully for location.

How are we to interpret the findings, and how to relate them to the school choice debate? The promise of a well-functioning school choice system is that it reduces or eliminates the role of location, thereby enabling children from poor families to access good schools. The countervailing view is that a choice system without fully flexible school size will increase the role of choice by schools, and the scope for the middle class to beat the system. Our findings cast some light on this debate. The results show that location is associated with most but not all of the differential school quality. We find consistent significant differences in school quality even among next-door neighbours, but the magnitude of these are low, relative to the raw differences. The importance of location suggests that school choice may have an important part to play in narrowing the gap in admission to good schools. By contrast, the roles of choice by schools and middle class strategising operate given location and account for a smaller part of the gap. A policy which reduces the factor contributing to the greater part of the gap, at the potential expense of widening the smaller part, might have some attractions[20].

[20] Whether the policy just approved by the UK parliament will make such a difference is waiting to be seen.

References

Abdulkadiroðlu, A., Pathak, P. A., and Roth, A. E. (2005), "The New York City High School Match", *American Economic Review*, 95(2): 364-367.

Abdulkadiroðlu, A., Pathak, P. A., Roth, A. E. and Sönmez, T. (2005), "The Boston Public School Match", *American Economic Review*, 95(2): 368-371.

Bayer, P. and McMillan, R. (2005), "Choice and Competition in Local Education Markets", NBER Working Paper 11802.

Blanden, J., Goodman, A., Gregg, P. and Machin, S. (2004), "Changes in Intergenerational Mobility in Britain" in M. Corak ed., Generational Income Mobility in North America and Europe, Cambridge University Press.

Bradley, S., Crouchley, R., Millington, J. and Taylor, J. (2000), "Testing for Quasi-market Forces in Secondary Education", *The Oxford Bulletin of Economics and Statistics*, 62(3): 357-390.

Burgess, S., McConnell, B., Propper, C. and Wilson, D. (2006), "The Impact of School Choice on Sorting by Ability and Socio-economic Factors in English Secondary Education", in L. Woessmann and P. Peterson eds., Schools and the Equal Opportunity Problem, Cambridge, MA: MIT Press.

Burgess, S. and Briggs, A. (2006), "School Assignment, School Choice and Social Mobility", CMPO Working Paper 157, CMPO, University of Bristol.

Burgess, S., Briggs, A., McConnell, B. and Slater, H. (2006), "School Choice in England: Background Facts", CMPO Working Paper 159, CMPO, University of Bristol.

Burgess, S. and Slater, H. (2006), "Using boundary changes to estimate the impact of school competition on test scores", CMPO Working Paper 158, CMPO, University of Bristol.

Chubb, J. and Moe, T. (1990), "Politics, Markets and America's Schools", Washington, WA: Brookings Institution Press.

Clark, D. (2004), "Politics, markets and schools: quasi-experimental evidence on the impact of autonomy and competition from a truly revolutionary UK reform", mimeo, University of California, Berkeley.

Department for Education and Skills (2003), "Code of Practice on school admissions". Available at: http://www.dfes.gov.uk/sacode

Epple, D. and Romano, R. (1998), "Competition between private and public schools, vouchers and peer group effects", *American Economic Review*, 88(1): 33-62.

Epple, D. and Romano, R. (2003), "Neighborhood Schools, Choice, and the

Distribution of Educational Benefits”, in Hoxby, C. ed., The Economics of School Choice, Chicago, IL: U of Chicago Press.

Friedman, M. (1955), “The Role of Government in Education”, in Solo, R. A. ed., Economics and the Public Interest, Piscataway, NJ: Rutgers University Press.

Galindo-Rueda, F., Marcenaro, O. and Vignoles, A. (2004), “The Widening Socio-Economic Gap in UK Higher Education”, CEE Discussion Paper 44, LSE.

Galindo-Rueda, F. and Vignoles, A. (2005), “The Heterogeneous Effect of Selection in Secondary Schools: Understanding the Changing Role of Ability”, CEE Discussion Paper 52, LSE.

Gibbons, S., Machin, S. and Silva, O. (2006), “Competition, Choice and Pupil Achievement”, CEE Discussion Paper 56, LSE.

Howell, W. G. (2004), “Dynamic Selection Effects in Means-Tested, Urban School Voucher Programs”, *Journal of Policy Analysis and Management*, 23(2): 225-250.

Howell, W. G. and Peterson, P.E. (2002), “The Education Gap: Vouchers and Urban Schools”, Washington, WA: Brookings Institution Press.

Hoxby, C. (2000), “Does competition among public schools benefit students and taxpayers?”, *American Economic Review*, 90(5): 1209-38.

Hoxby, C. (2003a), “The Economics of School Choice” (ed.), Chicago, IL: U of Chicago Press.

Hoxby, C. (2003b), “School choice and school productivity: could school choice be a tide that lifts all boats?”, in Hoxby, C. ed., The Economics of School Choice, Chicago, IL: U of Chicago Press.

Hoxby, C. (2005), “Competition among public schools: a reply to Rothstein”, NBER Working Paper 11216.

Jencks, C. (1970), “Education Vouchers: Giving Parents money to pay for schooling”, *The New Republic,* July 4th, 1970.

Ladd, H. (2002), “School Vouchers: A Critical View”, *Journal of Economic Perspectives*, 16(4): 3-24.

Neal, D. (2002), “How Vouchers Could Change the Market for Education”, *Journal of Economic Perspectives*, 16(4): 25-44.

Nechyba, T. (2003), “Introducing School Choice into Multidistrict Public School Systems”, in Hoxby, C. ed., The Economics of School Choice, Chicago, IL: U of Chicago Press.

Nechyba, T. (2004), "Income and Peer Quality Sorting in Public and Private Schools", in Hanushek, E. and Welch, F. eds., Handbook of Economics of Education, North Holland: Elsevier.

Rothstein, J. (2005), "Does competition among public schools benefit students and taxpayers? A comment on Hoxby (2000)", NBER Working Paper 11215.

Söderström, M. and Uusitalo, R. (2005), "School choice and segregation: evidence from an admission reform", IFAU Working paper 2005:7, Uppsala, Institute for Labour Market Policy Evaluation.

Vignoles, A. and Machin, S. (2004), "Educational inequality: the widening socio-economic gap", *Fiscal Studies*, 25(2): 107-128.

West, A. and Hind, A. (2003), "Secondary school admissions in England: Exploring the extent of overt and covert selection", Centre for Educational Research Report: http://www.risetrust.org.uk/admissions.pdf.

West, A. (2003), "Secondary School Admissions in England", memorandum submitted to the House of Commons Select Committee on Education and Skills: http://www.parliament.the-stationery-office.co.uk/pa/cm200203/cmselect/cmeduski/uc1075-i/1075m04.htm.

West, A., Hind, A., Pennell, H. (2004), "School admissions and `selection' in comprehensive schools: policy and practice", *Oxford Review of Education*, 30(3), 347-369

Wilson (2003) "Which Ranking? The Use of Alternative Performance Indicators in the English Secondary Education Market. ", CMPO Working Paper 58, CMPO, University of Bristol.

Table 1: Summary statistics for full sample

Panel A: *1,354,985 observations for all variables (except *could only be calculated for 1,321,676 pupils and ***could only be merged for 1,306,007 pupils)*

Variable	Mean	Standard Deviation	P25	P75
Dichotomous				
Good school (=1 if school is in top third of school %5 A-C distribution)	0.324	0.47	0	1
Pupil has 3 schools within 2KM of home postcode*	0.349	0.48	0	1
Pupil FSM eligibility	0.176	0.38	0	0
Male	0.508	0.50	0	1
White	0.851	0.36	1	1
Black Caribbean	0.016	0.13	0	0
Black African	0.014	0.12	0	0
Black Other	0.009	0.09	0	0
Indian	0.022	0.15	0	0
Pakistani	0.026	0.16	0	0
Bangladeshi	0.010	0.10	0	0
Chinese	0.003	0.05	0	0
Other ethnicity	0.022	0.15	0	0
Ethnicity not known**	0.027	0.16	0	0
School located in an urban area	0.706	0.46	0	1
School located in London	0.126	0.33	0	0
School located in rural area	0.168	0.37	0	0
SEN without statement	0.168	0.37	0	0
SEN with statement	0.023	0.15	0	0
English as first language	0.914	0.28	1	1
Continuous				
School %5 A-C	0.482	0.17	0.35	0.61
Pupil KS2 mean score	27.299	4.20	25	31
IMD score***	28.556	18.24	13.39	40.96

Panel B: *Only 116 LEAs due to 5 LEAs having Good/Total places = 0.*

Variable	Mean	Standard deviation	P25	P75
Good Places	3,606.62	4,215.2	930	4,547
Total Places	11,165.69	8474.3	6,093	12,464
Good/Total places	0.293	0.18	0.149	0.418
Good Places – FSM pupils	270.16	247.6	104	367
Total Places – FSM pupils	1,969.86	1,470.1	1,233	2,205
Good/Total places – FSM pupils	0.169	0.13	0.070	0.0251
Good Places – Non-FSM pupils	3,336.46	3,997.43	830	4,191
Total Places – Non-FSM pupils	9,195.84	7,600.0	4,899	10,438
Good/Total places – Non-FSM pupils	0.320	0.19	0.180	0.463
G^f/G^{nf}	0.503	0.25	0.370	0.595

Notes:
***Old ethnicity codes from PLASC 2002 were used for all pupils. This group includes pupils age 16 in PLASC 2003 and 2004 who did not have an ethnicity record in PLASC 2002.*

Table 2: Probit of whether pupil goes to a good school

	1.	Marg eff[1]	2.	Marg eff[1]	3.	Marg eff[1]
	FSM only		*Personal characteristics*		*Area characteristics*	
Pupil FSM eligibility	-0.594 (29.75)**	-0.173	-0.443 (27.86)**	-0.130	-0.178 (13.11)**	-0.048
Pupil KS2 mean score			0.061 (22.77)**	0.019	0.038 (14.13)**	0.012
Male			-0.079 (4.38)**	-0.024	-0.098 (5.01)**	-0.027
Black Caribbean			-0.179 (3.71)**	-0.053	0.025 (0.56)	0.006
Black African			0.090 (1.32)	0.028	0.270 (4.34)**	0.077
Black Other			-0.057 (1.24)	-0.017	0.089 (1.94)	0.025
Indian			0.105 (1.68)	0.033	0.192 (3.36)**	0.054
Pakistani			-0.177 (2.69)**	-0.053	0.038 (0.62)	0.011
Bangladeshi			-0.277 (4.02)**	-0.081	-0.117 (2.00)*	-0.032
Chinese			0.237 (5.56)**	0.075	0.288 (6.33)**	0.083
Other ethnicity			0.071 (2.29)*	0.022	0.158 (5.14)**	0.044
Ethnicity not known			0.195 (3.87)**	0.062	0.076 (1.31)	0.021
Index of Multiple Deprivation Score					-0.009 (1.51)	-0.003
LEA dummies	Yes		Yes		Yes	
Cohort dummies	Yes		Yes		Yes	
Mosaic dummies	No		No		Yes	
IMD*Mosaic dummies	No		No		Yes	
Constant	-0.500 (33.72)**		-2.316 (29.81)**		-1.204 (4.85)**	
Observations[2]	1321591		1321591		1263256	

Notes:
z-scores in parentheses; standard errors in all specifications clustered on LEA.
*Significant at 5% **Significant at 1%
Also included in both specifications 2 and 3 but not reported are: month of birth, English as first language, SEN with and without statement dummies. Specification 3 also includes mosaic dummies and these interacted with IMD.

[1] Marginal effects are average marginal effect over the sample. For binary variables, reported is the average difference (over the sample) in probabilities between the dummy taking value 1 and taking value 0. For continuous variables, this is the sample average of $f(X'ß)*ß_j$, where f(.) is the standard normal pdf.

[2] Observations for five LEAs with no 'good' schools as defined here are necessarily dropped from the probit. These LEAs are Barnsley, Islington, Knowsley, Luton and Rochdale.

Table 2b: Probit of whether pupil goes to a good school

	1.	Marg eff[1]	2.	Marg eff[1]	3.	Marg eff[1]
	FSM only		*Personal characteristics*		*Area characteristics*	
Pupil FSM eligibility	-0.613	-0.165	-0.456	-0.122	-0.190	-0.047
	(33.55)**		(31.33)**		(14.05)**	
Pupil KS2 mean score			0.063	0.018	0.040	0.011
			(23.34)**		(14.42)**	
Male			-0.122	-0.034	-0.145	-0.036
			(5.57)**		(6.17)**	
Black Caribbean			-0.190	-0.051	0.005	0.001
			(4.17)**		(0.11)	
Black African			0.100	0.028	0.261	0.067
			(1.43)		(4.06)**	
Black Other			-0.077	-0.021	0.066	0.017
			(1.70)		(1.47)	
Indian			0.056	0.016	0.156	0.040
			(0.83)		(2.60)**	
Pakistani			-0.211	-0.057	0.013	0.003
			(3.05)**		(0.20)	
Bangladeshi			-0.332	-0.087	-0.162	-0.040
			(4.59)**		(2.69)**	
Chinese			0.190	0.054	0.257	0.066
			(4.07)**		(5.22)**	
Other ethnicity			0.045	0.012	0.135	0.034
			(1.47)		(4.32)**	
Ethnicity not known			0.186	0.053	0.056	0.014
			(3.42)**		(0.96)	
Distance to nearest good school (km)	-0.144	-0.041	-0.146	-0.041	-0.153	-0.042
	(11.30)**		(11.29)**		(11.24)**	
Index of Multiple Deprivation Score					-0.006	-0.001
					(1.03)	
LEA dummies						
Cohort dummies						
Mosaic dummies						
IMD*Mosaic dummies						
Constant	0.009		-1.853		-0.834	
	(0.21)		(21.00)**		(3.44)**	
Observations[2]	1298097		1298097		1248597	

Notes:
z-scores in parentheses; standard errors in all specifications clustered on LEA.
*Significant at 5% **Significant at 1%
Also included in both specifications 2 and 3 but not reported are: month of birth, English as first language, SEN with and without statement dummies. Specification 3 also includes mosaic dummies and these interacted with IMD.

[1] Marginal effects are average marginal effect over the sample. For binary variables, reported is the average difference (over the sample) in probabilities between the dummy taking value 1 and taking value 0. For continuous variables, this is the sample average of $f(X'ß)*ß_j$, where f(.) is the standard normal pdf.

[2] Observations for five LEAs with no 'good' schools as defined here are necessarily dropped from the probit. These LEAs are Barnsley, Islington, Knowsley, Luton and Rochdale.

Table 3: Overall sample compared to same postcode sample

	Full sample	Pupils who share the same postcode with at least one other pupil
School Quality	0.482	0.469
	(0.17)	(0.17)
FSM eligibility	0.176	0.180
	(0.38)	(0.38)
KS2 mean score	27.299	27.167
	(4.20)	(4.20)
Observations	1,354,985	1,028,889

Table 4: Numbers of pupils per postcode

Numbers of pupils within postcode	Observations	Percentage of postcodes where pupils all have the same FSM eligibility status	Percentage of postcodes where pupils all attend the same school
1	205,700	100%	100%
2	222,714	82.87%	61.65%
3	197,430	71.04%	45.70%
4	159,848	62.39%	37.27%
5	125,640	54.78%	32.33%
6	91,632	48.76%	28.42%
7	67,900	43.03%	25.67%
8	49,080	38.03%	23.11%
9	34,290	32.44%	21.52%
10	23,840	31.75%	20.51%
11	17,204	26.92%	16.43%
12	11,316	24.81%	16.22%
13	8,372	22.36%	11.96%
14	6,020	20.93%	14.42%
15	4,035	21.19%	14.50%
>15	9,578	13.88%	9.32%
Total	**1,234,599**	**66.67%**	**49.29%**

Table 5: Postcode-cohort FE regressions of school quality on personal characteristics on non-selective LEAs

	1.	2.
	Full sample	*Full sample*
Pupil Free School Meal eligibility	-0.019	-0.015
	(25.49)**	(20.06)**
Pupil KS2 Mean score		0.003
		(38.92)**
Male dummy		-0.001
		(1.70)
Black Caribbean		0.006
		(2.13)*
Black African		0.017
		(4.63)**
Black Other		0.009
		(2.36)*
Indian		0.011
		(4.05)**
Pakistani		-0.001
		(0.39)
Bangladeshi		-0.005
		(1.64)
Chinese		0.019
		(2.86)**
Other		0.008
		(3.87)**
Ethnicity Unknown		0.013
		(6.46)**
Constant	0.473	0.382
	(3605.67)**	(130.88)**
Observations	1028899	1028899
R-squared	0.86	0.86

Notes:
t-stats in parentheses *Significant at 5% **Significant at 1%
Standard errors clustered by postcode-cohort. Regressions also include month of birth, SEN with and without statement and English as first language dummies.

Table 6: Whether a pupil attends a good school, comparing LEA and postcode fixed effects

	Overall (LEA-cohort FE)			Postcode-cohort FE	
	1. FSM only	*2. Personal characteristics*	*3. Area characteristics*	*4. FSM only*	*5. Personal characteristics*
Pupil FSM eligibility	-0.140	-0.100	-0.027	-0.029	-0.022
	(27.64)**	(24.70)**	(13.45)**	(17.50)**	(13.14)**
Pupil KS2 mean score		0.016	0.007		0.005
		(31.55)**	(22.11)**		(27.98)**
Male		0.003	0.000		-0.001
		(2.82)**	(0.14)		(1.00)
Black Caribbean		-0.041	0.010		0.006
		(5.26)**	(1.48)		(0.94)
Black African		-0.001	0.039		0.027
		(0.13)	(5.36)**		(3.34)**
Black Other		-0.023	0.017		0.007
		(2.58)*	(2.19)*		(0.91)
Indian		0.015	0.032		0.005
		(1.25)	(3.18)**		(0.71)
Pakistani		-0.035	0.013		-0.017
		(3.22)**	(1.26)		(2.72)**
Bangladeshi		-0.021	0.015		-0.016
		(1.72)	(1.54)		(2.35)*
Chinese		0.053	0.058		0.028
		(4.53)**	(5.26)**		(1.76)
Other ethnicity		0.012	0.030		0.016
		(1.63)	(5.11)**		(3.10)**
Ethnicity not known		0.049	0.024		0.036
		(3.34)**	(1.40)		(6.78)**
Index of Multiple Deprivation Score			-0.007		
			(2.52)*		
Constant	0.322	-0.133	0.325	0.302	0.140
	(352.74)**	(9.10)**	(3.40)**	(1005.46)**	(19.39)**
Observations	1028899	1028899	993704	1028899	1028899
R-squared	0.17	0.18	0.28	0.87	0.87
Adjusted R-squared	0.17	0.18	0.28	0.68	0.68

Notes:
t-stats in parentheses *Significant at 5% **Significant at 1%

[1] Overall FE includes LEA-cohort fixed effects and reports SEs clustered by LEA-cohort. Postcode FE includes postcode-cohort fixed effects and reports SEs clustered by postcode-cohort. Dependent whether is 0,1, whether pupil attends a good school. This uses OLS-FE

Table 7: Different Definitions of a 'good' school

(a) Specification from Table 2b

Coefficient on Pupil FSM eligibility:

Definition of 'good' school		**1**		**2**		**3**	
Metric:	**Top third of:**	*FSM only*	**ME**	*Personal characteristics*	**ME**	*Area characteristics*	**ME**
GCSE	National	-0.613 (33.55)**	-0.165	-0.456 (31.33)**	-0.122	-0.190 (14.05)**	-0.047
GCSE	LEA	-0.551 (27.12)**	-0.171	-0.414 (26.40)**	-0.129	-0.175 (14.68)**	-0.052
VA	National	-0.340 (14.87)**	-0.105	-0.273 (15.52)**	-0.084	-0.115 (8.73)**	-0.034
VA	LEA	-0.344 (17.07)**	-0.117	-0.275 (18.33)**	-0.093	-0.116 (10.52)**	-0.038
Obs		1330477		1330477		1280188	

ME = marginal effect

For additional variables and notes, see notes to Table 2b

(b) Specification from Table 6

Coefficient on Pupil FSM eligibility:

Definition of 'good' school		**Overall (LEA-cohort FE)**			**Postcode-cohort FE**	
		1	***2***	***3***	**4**	***5***
Metric:	**Top third of:**	*FSM only*	*Personal characteristics*	*Area characteristics*	*FSM only*	*Personal characteristics*
GCSE	National	-0.140 (27.64)**	-0.100 (24.70)**	-0.027 (13.45)**	-0.029 (17.50)**	-0.022 (13.14)**
GCSE	LEA	-0.162 (40.00)**	-0.119 (36.04)**	-0.037 (16.46)**	-0.034 (27.98)**	-0.026 (21.82)**
VA	National	-0.092 (20.25)**	-0.072 (18.99)**	-0.022 (9.69)**	-0.020 (16.82)**	-0.016 (13.38)**
VA	LEA	-0.110 (23.93)**	-0.087 (23.36)**	-0.028 (11.74)**	-0.024 (18.03)**	-0.019 (14.39)**
Obs		1028899	1028899	993704	1028899	1028899

For additional variables and notes, see notes to Table 7

Table 8: Probit of whether pupil goes to a good school

	1.	Marg eff[1]	2.	Marg eff[1]	3.	Marg eff[1]	4.	Marg eff[1]
			London		*Urban*		*Rural*	
Pupil FSM eligibility	-0.135	-0.037	-0.333	-0.087	-0.149	-0.038	-0.062	-0.020
	(9.97)**		(10.99)**		(11.60)**		(4.11)**	
Pupil KS2 mean score	0.038	0.010	0.060	0.016	0.039	0.010	0.020	0.006
	(14.69)**		(10.78)**		(11.69)**		(7.83)**	
FSM*pupil has 3	-0.109	-0.030						
schools in 2KM	(4.10)**							
KS2 score*pupil has 3	0.001	0.000						
schools in 2KM	(0.64)							
Index of Multiple	-0.009	-0.002	-0.007	-0.002	-0.008	-0.002	0.021	0.007
Deprivation Score	(1.50)		(1.06)		(0.87)		(1.26)	
LEA dummies	Yes		Yes		Yes		Yes	
Cohort dummies	Yes		Yes		Yes		Yes	
Mosaic dummies	Yes		Yes		Yes		Yes	
IMD*Mosaic dummies	Yes		Yes		Yes		Yes	
Observations	1239888		159047		884816		206339	

Notes:
z-scores in parentheses *Significant at 5% **Significant at 1%

[1] Marginal effects are average marginal effect over the sample. For binary variables, reported is the average difference (over the sample) in probabilities between the dummy taking value 1 and taking value 0. For continuous variables, this is the sample average of $f(X'ß)*ß_j$, where f(.) is the standard normal pdf.

[2] Observations for five LEAs with no good schools as defined here are necessarily dropped from the probit. These LEAs are Barnsley, Islington, Knowsley, Luton and Rochdale.

[3] Also included but not reported are gender and ethnicity dummies, month of birth, English as first language, SEN with and without statement dummies and a constant in all specifications.

Table 9: Postcode-cohort FE on School Quality by deciles of School choice feasibility

Decile	Distance (Metres)	FSM Coefficient (unconditional)[2]	FSM Coefficient (conditional)[3]	Observations
1	1144.034	-0.032 (7.21)**	-0.023 (5.14)**	101,426
2	1539.576	-0.031 (6.58)**	-0.022 (4.59)**	101,426
3	1808.860	-0.032 (6.60)**	-0.024 (4.98)**	101,423
4	2063.148	-0.030 (6.58)**	-0.022 (4.79)**	101,414
5	2338.386	-0.028 (5.37)**	-0.021 (4.00)**	101,422
6	2687.280	-0.032 (6.12)**	-0.026 (4.84)**	101,423
7	3204.911	-0.025 (4.00)**	-0.017 (2.78)**	101,427
8	4182.528	-0.030 (4.65)**	-0.022 (3.47)**	101,418
9	6438.545	-0.021 (3.20)**	-0.017 (2.59)**	101,422
10	12,776.570	-0.023 (3.85)**	-0.018 (3.10)**	101,421

Notes:
t-stats in parentheses *Significant at 5% **Significant at 1%

[1] Postcode-cohort FE regressions as per table 6 with standard errors clustered by postcode-cohort.
[2] FSM is the only explanatory variable, as per column 1 of table 6.
[3] Includes all other personal characteristics, as per column 2 of table 6.
[4] School choice feasibility is measured by the minimum distance to three schools for each pupil

Table 10: Estimating the probability that a pupil attends their nearest school

		1.	Marg eff[1]	2.	Marg eff[1]	3.	Marg eff[1]
Pupil FSM eligibility		0.442	0.157	0.438	0.153	0.373	0.126
		(13.15)**		(13.95)**		(12.55)**	
Pupil KS2 mean score		-0.015	-0.005	-0.014	-0.005	-0.047	-0.016
		(12.54)**		(14.38)**		(17.62)**	
Index of Multiple Deprivation Score				-0.010	-0.004	-0.012	-0.004
				(1.92)		(2.45)*	
School quality group (based on prior school %5 A-C)	**2**	0.107	0.039	0.106	0.038	0.096	0.033
		(1.92)		(1.94)		(1.76)	
	3	0.194	0.070	0.190	0.068	0.199	0.068
		(3.35)**		(3.32)**		(3.37)**	
	4	0.281	0.101	0.281	0.100	0.289	0.098
		(4.11)**		(4.22)**		(4.21)**	
	5	0.323	0.116	0.331	0.117	0.362	0.123
		(4.66)**		(4.97)**		(5.46)**	
	6	0.430	0.153	0.428	0.150	0.437	0.147
		(6.55)**		(6.69)**		(6.86)**	
	7	0.464	0.165	0.477	0.167	0.493	0.165
		(6.99)**		(7.51)**		(7.72)**	
	8	0.503	0.178	0.508	0.176	0.502	0.168
		(7.39)**		(7.67)**		(7.82)**	
	9	0.558	0.196	0.557	0.193	0.586	0.194
		(7.16)**		(7.28)**		(7.87)**	
	10	0.571	0.201	0.586	0.202	0.603	0.199
		(8.51)**		(8.81)**		(9.48)**	
	11	0.589	0.207	0.618	0.212	0.611	0.202
		(7.40)**		(7.93)**		(8.21)**	
	12	0.642	0.224	0.658	0.225	0.653	0.215
		(8.20)**		(8.59)**		(8.78)**	
	13	0.689	0.239	0.694	0.235	0.683	0.224
		(9.82)**		(9.78)**		(9.63)**	
	14	0.707	0.245	0.742	0.250	0.719	0.235
		(9.61)**		(10.19)**		(10.16)**	
	15	0.757	0.260	0.772	0.259	0.735	0.239
		(10.43)**		(10.15)**		(10.20)**	
	16	0.777	0.266	0.788	0.264	0.752	0.244
		(10.44)**		(10.19)**		(9.99)**	
	17	0.844	0.286	0.866	0.286	0.814	0.262
		(10.71)**		(10.62)**		(10.27)**	
	18	0.850	0.287	0.884	0.291	0.798	0.257
		(10.06)**		(10.07)**		(9.35)**	
	19	0.698	0.241	0.720	0.243	0.625	0.206
		(7.37)**		(7.23)**		(6.35)**	
	20	-0.250	-0.090	-0.240	-0.086	-0.467	-0.157
		(2.02)*		(1.91)		(3.31)**	
FSM* School quality group	**2**	-0.128	-0.047	-0.122	-0.044	-0.114	-0.039
		(2.86)**		(2.84)**		(3.01)**	
	3	-0.094	-0.034	-0.092	-0.033	-0.076	-0.026
		(2.25)*		(2.30)*		(1.99)*	
	4	-0.137	-0.050	-0.145	-0.052	-0.123	-0.042
		(3.27)**		(3.56)**		(3.33)**	

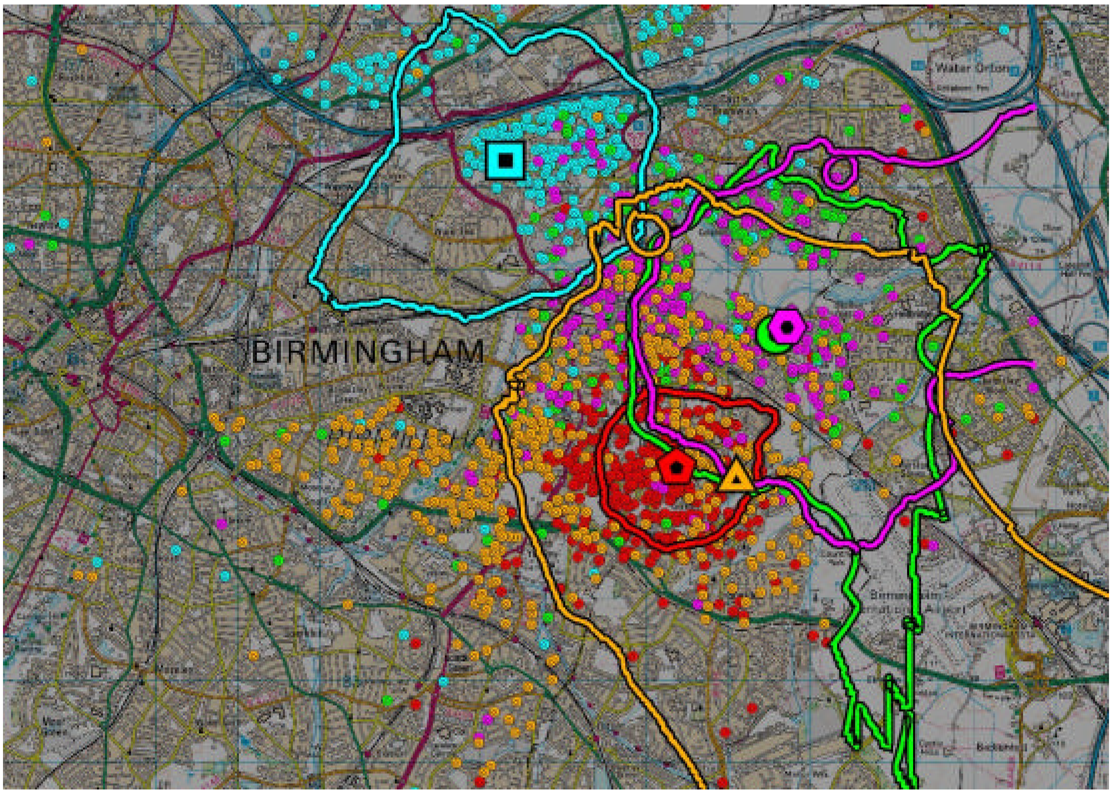
	5	-0.204	-0.074	-0.211	-0.075	-0.188	-0.064
		(4.57)**		(4.85)**		(4.80)**	
	6	-0.208	-0.075	-0.208	-0.074	-0.172	-0.059
		(4.64)**		(4.72)**		(4.09)**	
	7	-0.269	-0.097	-0.281	-0.100	-0.235	-0.081
		(6.42)**		(6.80)**		(6.31)**	
	8	-0.338	-0.121	-0.342	-0.121	-0.286	-0.098
		(8.19)**		(8.41)**		(7.52)**	
	9	-0.328	-0.117	-0.340	-0.121	-0.306	-0.104
		(7.23)**		(7.61)**		(7.23)**	
	10	-0.414	-0.146	-0.432	-0.152	-0.376	-0.127
		(8.51)**		(8.88)**		(8.68)**	
	11	-0.448	-0.158	-0.470	-0.164	-0.389	-0.132
		(7.97)**		(8.26)**		(7.67)**	
	12	-0.528	-0.183	-0.555	-0.192	-0.477	-0.160
		(9.65)**		(10.49)**		(9.86)**	
	13	-0.601	-0.207	-0.629	-0.215	-0.536	-0.178
		(13.01)**		(13.91)**		(13.47)**	
	14	-0.633	-0.216	-0.669	-0.227	-0.560	-0.186
		(11.94)**		(12.75)**		(11.64)**	
	15	-0.736	-0.247	-0.740	-0.248	-0.630	-0.207
		(12.65)**		(15.22)**		(14.18)**	
	16	-0.739	-0.248	-0.784	-0.260	-0.667	-0.217
		(13.85)**		(15.62)**		(14.42)**	
	17	-0.841	-0.275	-0.891	-0.289	-0.757	-0.243
		(16.00)**		(16.90)**		(15.60)**	
	18	-0.977	-0.309	-1.037	-0.325	-0.853	-0.269
		(16.93)**		(19.32)**		(17.23)**	
	19	-1.174	-0.353	-1.190	-0.359	-0.964	-0.297
		(18.79)**		(21.07)**		(18.42)**	
	20	-1.194	-0.356	-1.210	-0.362	-0.861	-0.271
		(14.48)**		(14.64)**		(10.45)**	
LEA dummies		Yes		Yes		Yes	
Cohort dummies		Yes		Yes		Yes	
Mosaic dummies		No		Yes		Yes	
IMD*Mosaic dummies		No		Yes		Yes	
Observations		1349117		1269180		1269180	

Notes:
z-scores in parentheses *Significant at 5% **Significant at 1%

[1] Marginal effects are average marginal effect over the sample. For binary variables, reported is the average difference (over the sample) in probabilities between the dummy taking value 1 and taking value 0. For continuous variables, this is the sample average of f(X'ß)*$ß_j$, where f(.) is the standard normal pdf.

[2] Standards errors are clustered by LEA in all specifications. All specifications include gender and ethnicity dummies, dummies for month of birth, SEN (with and without statement) and English as a first language. Specification 2 adds area characteristics (IMD and mosaic category). Specification 3 is the same as specification 2 but also includes dummies for whether pupils' nearest school are located in urban areas and a London LEA, interactions of school quality group with KS2 mean score, distance to nearest school, distance squared and distance cubed and these distances interacted with FSM, urban and rural dummies.

Figure 1: School Distance Contours in Birmingham



The large shapes represent the locations of 5 schools, and the dots represent a sample of pupils attending those schools. The lines are 50% distance contours round each school. Thanks to Rich Harris for producing this analysis.

Figure 2: Good to total places ratio for FSM pupils against good to total places ratio for FSM for Non-FSM pupils

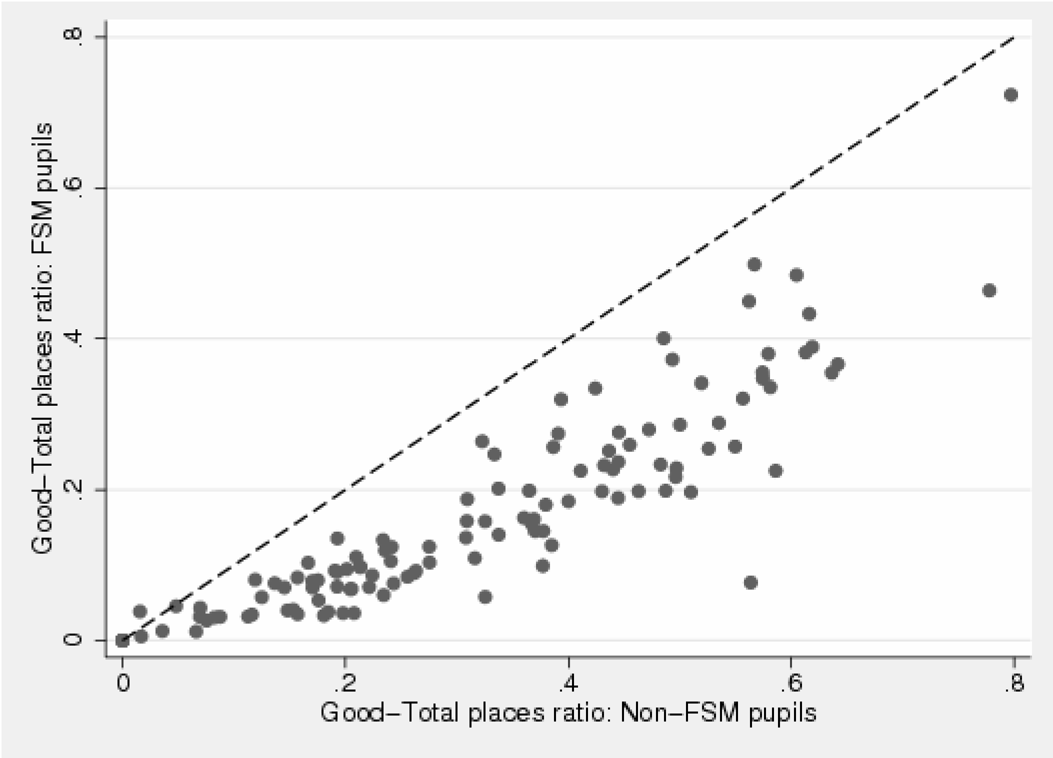


Figure 3: FSM vs Non-FSM gaps in mean school percentage 5 A-C for 2, 3, 4 and 5 pupil postcodes

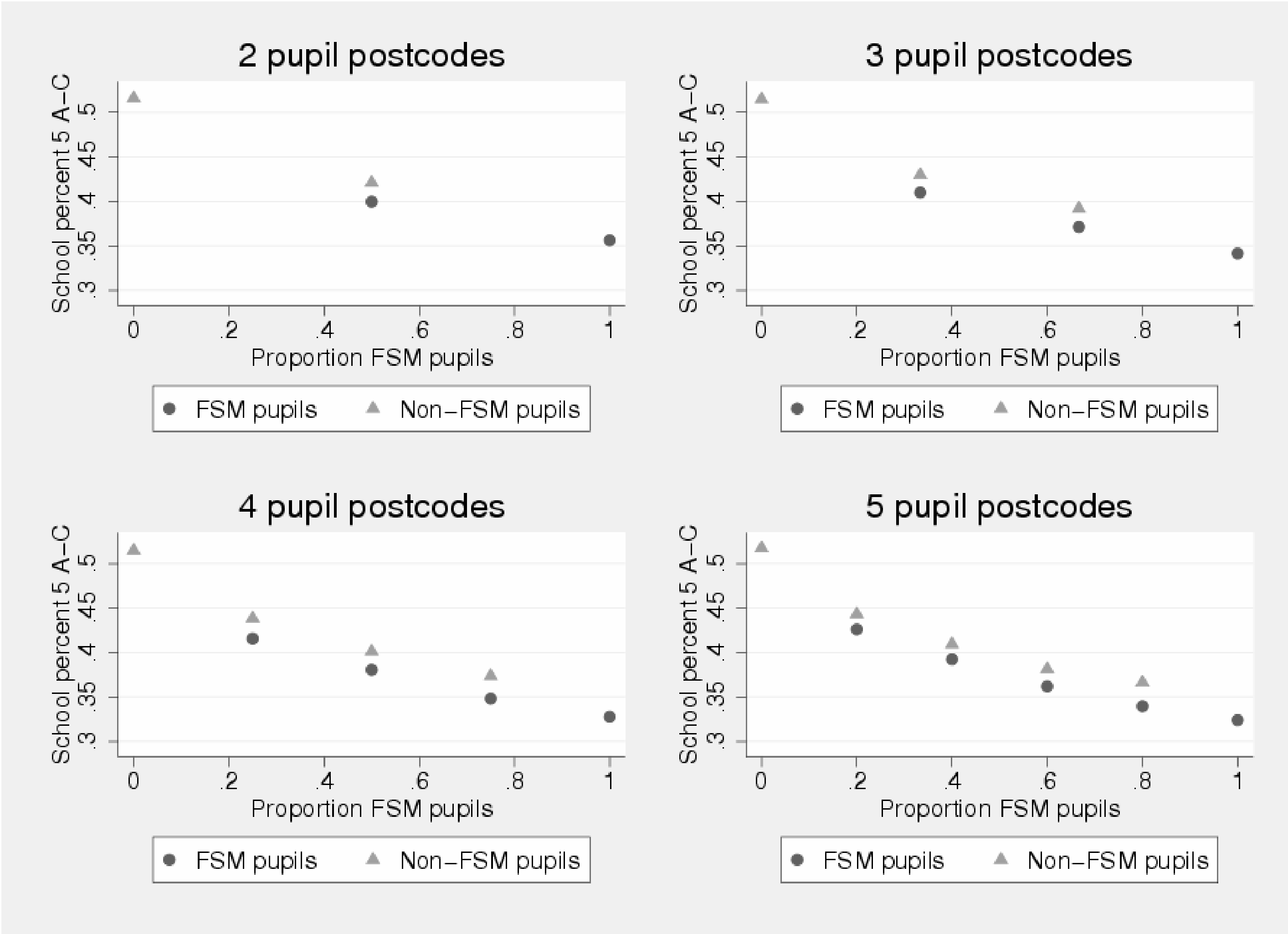


Figure 4: Differences in school quality by differences in FSM status by postcode

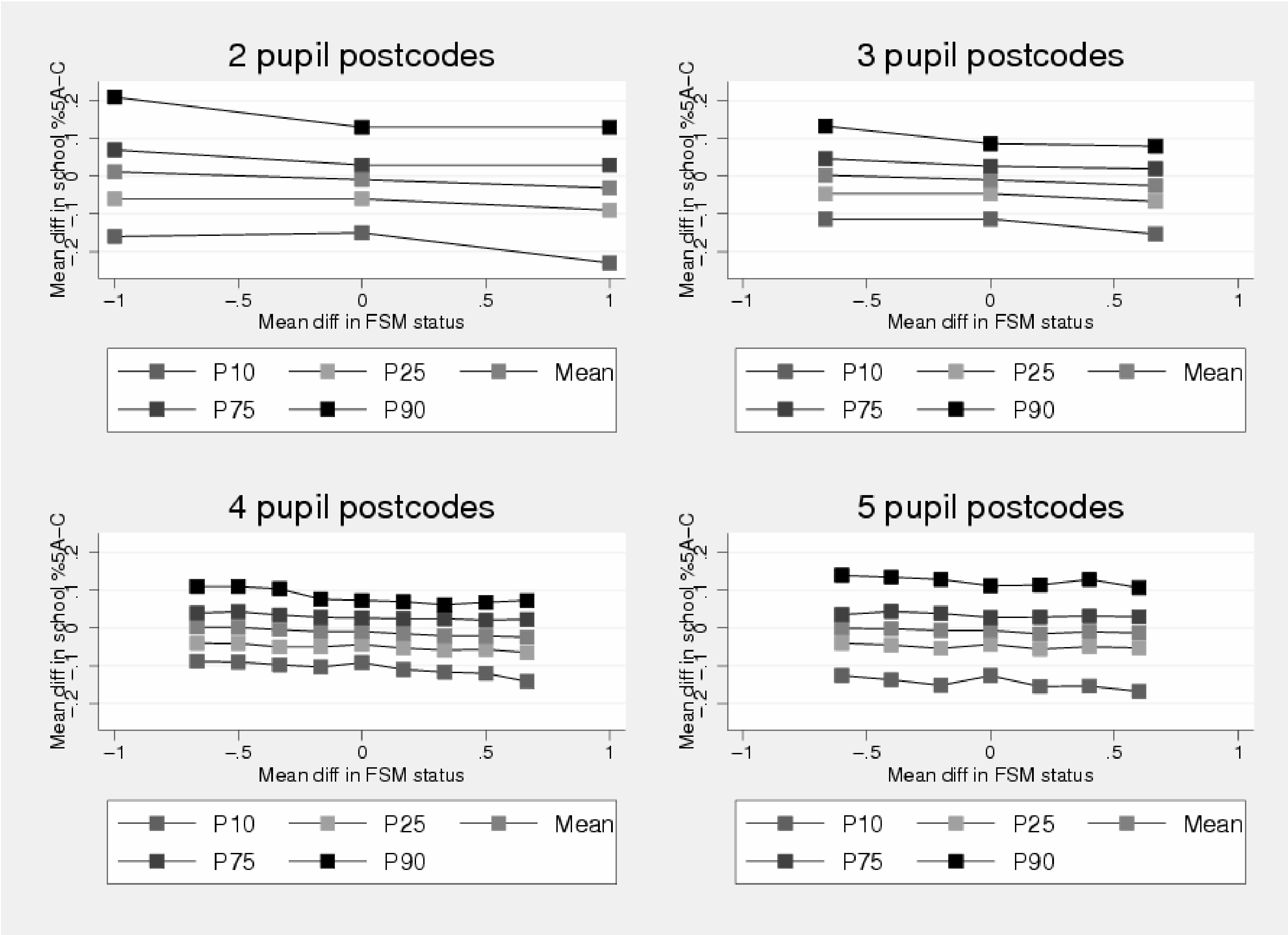


Figure 5a: Proportion of pupils attending their nearest school by FSM and ventiles of school quality (Raw data)

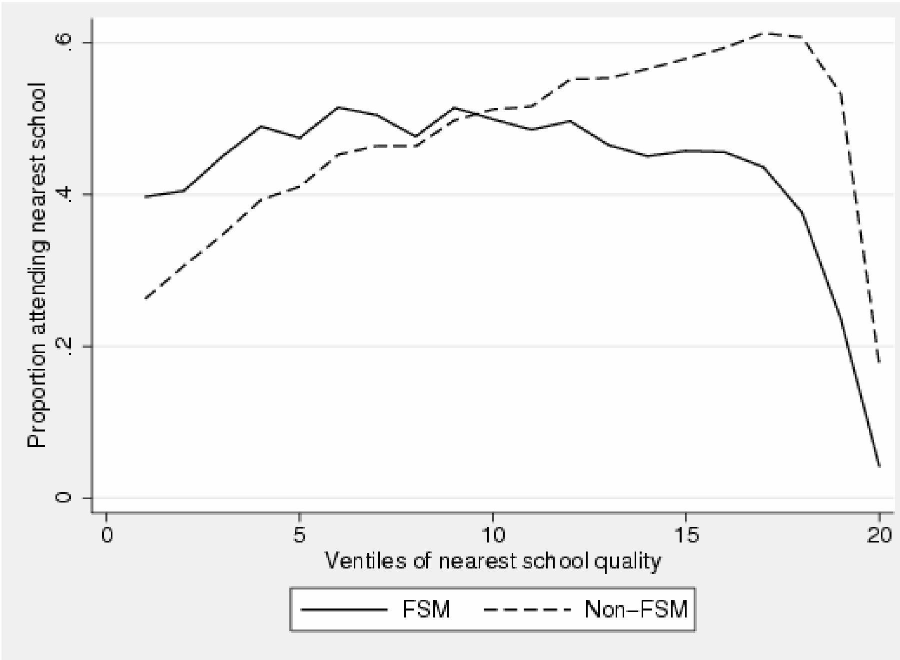
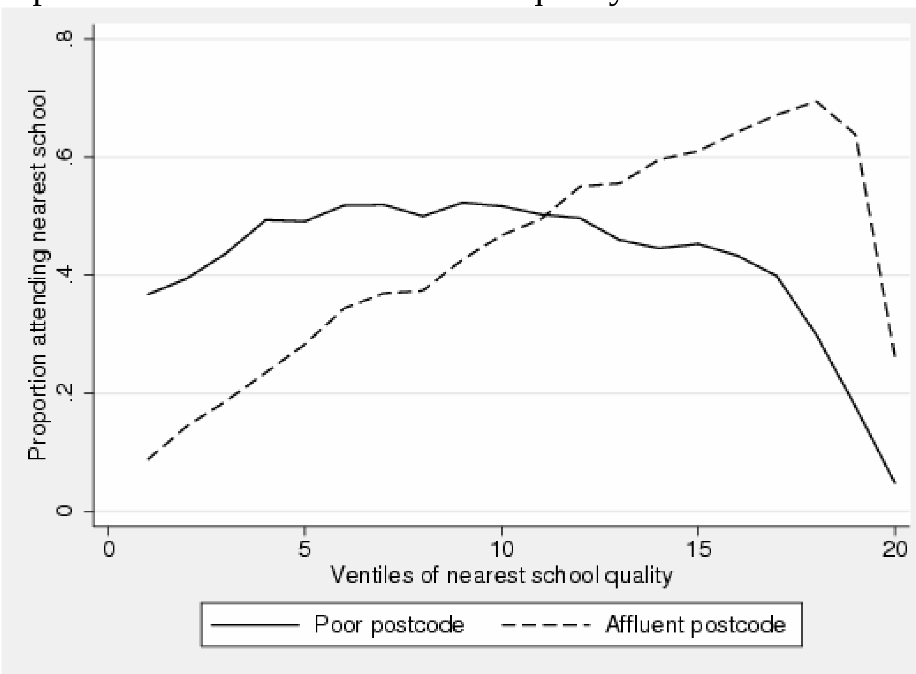


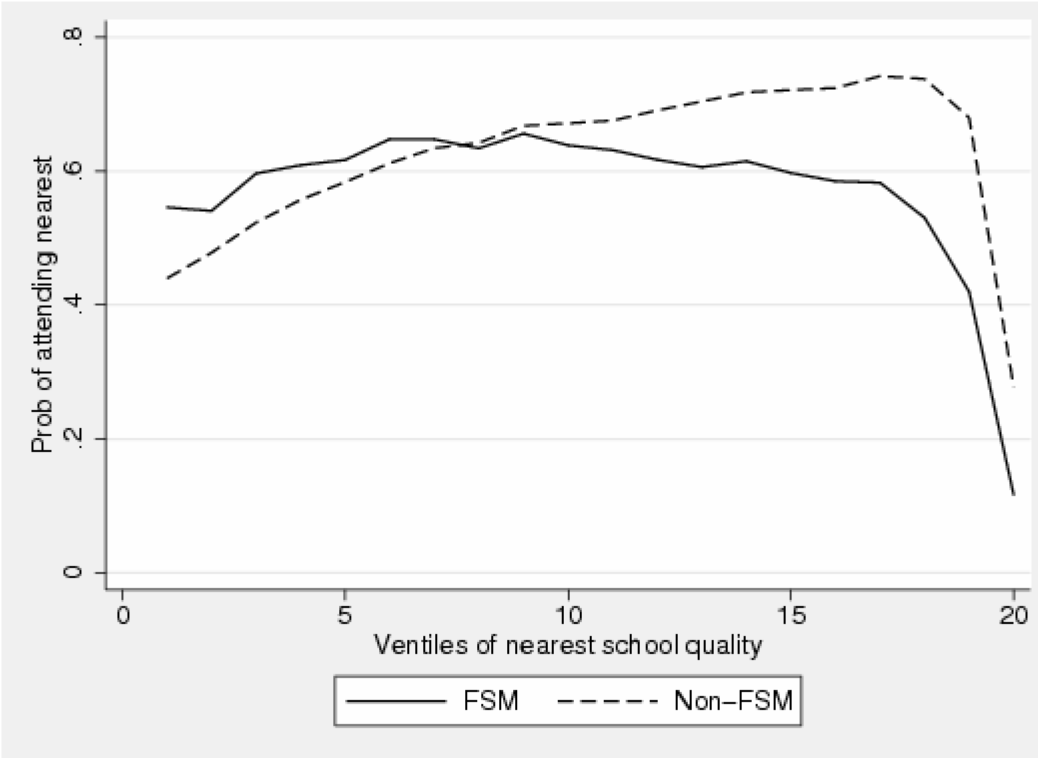
Figure 5b: Probability of pupils attending their nearest school by poor and affluent postcodes and ventiles of school quality



Notes:

Poor postcodes are those with mosaic categories in the top third of categories as ranked by proportion FSM and affluent postcodes are those in the bottom third.

Figure 5c: Fitted probability of pupils attending their nearest school by FSM status and ventiles of school quality



[1]Based on col 3 of table 10 for a white, female pupil born in September with average KS2 mean, English as first language, no SEN, attending a school in an urban area and with the mean distance to nearest good school