Department of Quantitative Social Science

# Choosing secondary school by moving house: school quality and the formation of neighbourhoods

Rebecca Allen
Simon Burgess
Tomas Key

DoQSS Working Paper No. 10-21
June 2012

Leading education and social research
Institute of Education
University of London

**DISCLAIMER**

**Any opinions expressed here are those of the author(s) and not those of the Institute of Education. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.**

**DoQSS Workings Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.**

Department of Quantitative Social Science. Institute of Education, University of London. 20 Bedford way, London WC1H 0AL, UK.

# Choosing secondary school by moving house: school quality and the formation of neighbourhoods

**Rebecca Allen**[*], **Simon Burgess**[†], **Tomas Key**[‡§]

**Abstract.** This paper uses the pupil census in England to explore how family house moves contribute to school and residential segregation. We track the moves of a single cohort as it approaches the secondary school admission age. We also combine a number of cohorts and estimate a dynamic nonlinear model for house moving with unobserved effects. These approaches yield the same result: moving is significantly negatively correlated with school quality, and segregation does increase as a cohort reaches age 11. However, this relationship is weak: the increase in segregation is slight and quantitative significance of the estimated relationship is low.

**JEL classification:** I20, R23.

**Keywords:** school quality, moving, segregation, neighbourhoods.

[*]Institute of Education, University of London. E-mail: r.allen@ioe.ac.uk.

[†]Centre for Market and Public Organisation, University of Bristol. E-mail: simon.burgess@bristol.ac.uk.

[‡]University of Oxford. E-mail: tomas.key@economics.ox.ac.uk.

[§]Thanks to the Department for Children, Schools and Families for providing NPD data, and to ESRC for funding Tomas Key through CMPO, grant number: RES-343-28-3001.

# 1. Introduction

There remain inequalities in the extent to which poor families are able to access high quality schools in England, despite several reforms to the school admissions code over the past ten years. These inequities potentially come about through informational and budget constraint disadvantages in the school choice process, given their household location at age 10, or they may derive from a constrained ability to choose a school by living in the 'right' neighbourhood. The former has been extensively studied in the UK (e.g. Coldron et al., 2008; Burgess et al., 2009); this paper studies the latter phenomenon. In this paper we present data from the National Pupil Database on house moves that take place by families who have a child in primary school. Since house location remains the most important determinant of the secondary school choice process in England, many of these house moves will be motivated by a desire to locate close to a good school. Analysing the nature of house moves in the early years of a child's life allows us to explore how schools and neighbourhoods are formed, and the relationship between the two.

Our analysis begins by quantifying the extent to which family house moves contribute to neighbourhood, and thus school, segregation by tracking a single cohort of pupils from the start of primary school through to the start of secondary school. Our expectation is that we would see increasing segregation in this cohort as the wealthier parents cluster ever closer to the popular schools. This is indeed what the data reveal, but the effect is very modest, a change of 0.02 in the segregation index, relative to a total of 0.38. We then focus specifically on modelling house move events directly to understand how this behaviour drives changes in sorting. Unfortunately the complexity of the secondary school admissions process, with proximity to school and a wide variety of other criteria being important, makes it impossible to replicate empirical strategies such as regression discontinuity designs used in areas with strict residence requirements for school access. Instead, our approach analyses the propensity to move house in relationship to the social background of the family and nature of their initial household location. Using this approach we are able to show that households with high income are not more likely to move house overall, but they are significantly more likely to respond to lower quality local schools in their initial house location. Again, however, whilst the effect is in the expected direction, the effect is small – a unit standard deviation in school quality induces a mean change in moving probability of 0.006, relative to a mean of 0.10. So although there are very large numbers of house moves between the ages of 5 and 11, these do not appear to make a large contribution to residential sorting overall.

The context for the study is compulsory education in England, which lasts for 11 years, covering the primary (age 5 to 11) and secondary stages (age 11 to 16). Most pupils transfer from primary to secondary school at age 11, although there are a few areas where this transfer is slightly different due to the presence of middle schools. Secondary school allocation takes place via a system of constrained choice whereby parents are able to express ordered preferences for three to six schools anywhere in England and are offered places on the basis of published admissions criteria that must adhere to a national Admissions Code[1] . First priority is usually given to pupils with a sibling already at the school, pupils with statements of special educational needs and to children in public care. However, the largest proportion of places is allocated giving priority to children living within a designated area or on the basis of proximity to school. This means that house location continues to be critical to securing a school place in England, despite the choice system that exists. There are a significant number of schools who do not give priority to local communities: at voluntary-aided religious schools (17 percent of secondary pupils), priority is usually given on the basis of religious affiliation or adherence; there are also state schools that offer a proportion of places on the basis of ability or aptitude for a particular subject, and this includes 164 entirely selective grammars schools.

The following section of the paper provides a review of the related literature, section 3 sets out the empirical framework, section 4 describes the data, section 5 presents the results and some conclusions are offered in section 6.

## 2. Background literature

The formation of school peer groups has been extensively studied in theoretical models and empirical data. A central feature of the models is that peer groups, and thus school quality and the value of housing, are endogenously determined, with most models directly building on hedonic pricing models such as those first developed by Tinbergen (1959) and Sattinger (1980). The models match consumers to locations and find prices that separate people based on willingness to pay for locational quality, of which local school quality is one dimension.

The general equilibrium models of Epple and Romano (1998) and Nechyba (2000) show location choice where school assignment is decided strictly via a residence requirement restriction and thus households purchase homes and public school access as bundles. This is the traditional means by

---

[1] The Admissions Code has changed twice over the period under analysis in this paper and this means that schools have needed to adapt their own admissions policies as certain criterion have been disallowed (see West et al., 2004; Coldron et al., 2008; West et al., 2009, for surveys of admissions policies in use over this period).

which parents 'choose' schools in the United States and is similar in nature to catchment areas or proximity oversubscription criterion in the English context. The elements and shape of the household utility function are the key determinants of the nature of stratification and house price differentials in equilibrium. Given basic assumptions about household attributes and preferences, Epple and Romano show there exists a strict hierarchy of house prices with boundary indifference and strict household preference for their own neighbourhood. In their models, the elasticity of the price premium with respect to neighbourhood quality is a function of the elasticity of the child's schooling attainment with respect to school quality (with school quality being determined by the average child's prior ability in the school) and of the correlation between willingness to pay for school quality and child's prior ability. If the first elasticity is large, then large proportionate differences in price are needed to segregate people. If the second correlation is high then a small difference in price is needed to segregate people.

Econometric studies have consistently shown that parents are willing to pay for school quality through the housing market (Black, 1999; Bogart and Cromwell, 1997, 2000; Goodman and Thibodeau, 1998; Sieg et al., 1999; Gibbons and Machin, 2003; Cheshire and Sheppard, 1995; Leech and Campos, 2003; Rosenthal, 2003). For example, Black (1999) examines the differential house prices for those residences either side of elementary school attendance zone boundaries, in an attempt to resolve the problem of unobservable neighbourhood characteristics. She estimates that families are willing to pay 2.5% higher house prices for a school quality increase of 5%. Bayer et al. (2007) estimate a mean marginal willingness to pay for a standard deviation increase in average test scores of approximately 2 percent of house value in San Francisco, with strong heterogeneity around this mean. Gibbons and Machin (2003) find similar results for English primary schools. The analysis of Rothstein (2006) suggests that this willingness to pay is related to peer, rather than school, characteristics.

Empirical studies suggest that the correlation between a family's willingness-to-pay for school quality and their child's initial ability is likely to be quite large, especially given the known relationship between the social background of the family and child's expected ability. For example, Bayer and McMillan find significant evidence for differences in preferences with more educated parents willing, *ceteris paribus*, to spend more for high quality schools (Bayer and McMillan, 2007). This finding is confirmed by Nesheim (2002) who uses a similar approach to estimate a correlation between parental education and willingness to pay for school quality (given income) of 0.59.

On the face of it, the theoretical general equilibrium models and the empirical studies using regression discontinuities appear to disagree because the former suggests large long-run differences in neighbourhood quality and house price stratification and the latter suggests only quite modest differences in house prices at discontinuities. Bayer et al. (2007) attempt to reconcile the theoretical predictions and the empirical findings. They note that the empirical literature outlined above has only considered the partial equilibrium effect of school quality on demand, and that the general equilibrium effect might be larger. This is because certain households place a large weight on the characteristics of their neighbours when choosing where to reside and where to send their children to school. Using simulations, they conclude that school-related sorting explains 25% of neighbourhood stratification on the basis of income and 30% on the basis of education. These general equilibrium effects are on average 4 times greater than the partial equilibrium results for education stratification, and twice the size for income.

These models of school and neighbourhood formation are only of limited relevance in a country such as England with complex, choice-based school admissions procedures and highly constrained housing stock. Despite school choice, the characteristics of individuals living in an area are almost always an important determinant of the formation of school peer groups. Choices about where families locate within an area in turn depend on the geographic distribution of housing stock and jobs, levels of inequality and other social differences, access to transport, and so on. The admissions policies of schools and their relative qualities both impact on the need to choose a house close to a particular school and the chances of being accepted at a school of choice given house location. The one simplifying factor in modelling household location in England is that there is essentially no tax-related process of Tiebout sorting (Tiebout, 1956): voters can, and do, also influence the level of education funding in the local authority, but the magnitude of this funding difference is very small compared to the variation in per pupil funding distributed to schools to compensate for educational and social disadvantage (Levačić et al., 2005).

There are developments on the basic general equilibrium models of household location that explore the potential impact of different types of school choice programs. Some models adapt scenarios whereby public schools make places available for non-neighbourhood families and these are allocated by lottery, with transportation costs providing an important disutility to travelling outside the neighbourhood (Epple and Romano, 2003; Brunner and Imazeki, 2008). The winners and losers from this type of choice system (relative to strict neighbourhood schooling) are complex. Other choice models introduce a voucher program for private schools into a neighbourhood school system (Epple and Romano, 1998; Nechyba, 1999; Ferreyra, 2007).

In England there has been little change in levels of free school meal segregation over the past two decades, despite policies to facilitate greater parental choice and several major changes in the school admissions code (Gorard et al., 2003; Allen and Vignoles, 2007; Gibbons and Telhaj., 2007). This persistent segregation makes it clear that there are unequal opportunities for poor and rich families to access high quality schools. Burgess et al. (2008) look at primary-to-secondary transitions, demonstrating that pupils who are eligible for free school meals attend lower quality secondary schools than the contemporaries who attended the same primary school but are not eligible for free school meals. Burgess and Briggs (2010) show that there are inequalities in the probability of getting a place at the good school for low and higher income pupils who are living on the same street. Allen (2007) shows that where free school meals pupils attend a non-nearest school, either because they choose to or because their local school will not accept them, they tend to see a deterioration in their school peer group.

Past studies have shown that levels of school segregation are consistently higher than levels of residential segregation. This gap between school and residential sorting – i.e. post-residential sorting – is greatest in areas with more 'choice' schools and where population density is greater (Allen, 2007; Burgess et al., 2007). However, these studies cannot distinguish between whether school choice policies lower residential sorting by reducing the need for some parents to choose a school by moving house or they raise school sorting through the use of social selective oversubscription criteria.

This paper addresses this research area by establishing the importance of school quality in household location.

## 3. Modelling framework

We set out a simple modelling framework in this section to form the basis for the empirical analysis below. We assume that families start without children and initially chose where to live on that basis. They then acquire children and consider relocating before the key date for school assignment. Obviously, there are other (random) reasons for moving too, unrelated to school quality. We think of a world in overall equilibrium so that the distributions of income, tastes, labour market states, amenities and school qualities are fixed. The last of these is obviously an important assumption, but we are modelling choice of neighbourhood and school for given school quality. Within that overall equilibrium, families may move house within a cohort as it ages. House prices are fixed; people move between locations, but in equilibrium, prices remain constant. School quality and neighbourhood quality are exogenous, unaffected by the people learning or living there; modelling the joint evolution of school and neighbourhood quality alongside endogenous mobility is beyond the scope of this paper.

Each family *i* chooses location *L* to maximise utility:

$$U = U\left(\left(y_i - p(L)\right), \lambda_i k_i q(L), m(L), \mu_i \varepsilon(L)\right) \tag{1}$$

where *y* is income, *p(L)* is the cost of housing in *L*, *q(L)* is the quality of schooling, *k* is an indicator of whether the family has a child or not, $\lambda$ is a taste for school parameter, *m(L)* represents labour market opportunities in *L*, $\varepsilon$ represents local amenities and $\mu$ idiosyncratic tastes for this. With a given and fixed supply of housing, house prices adjust to equate demand for housing in each location to its supply

$$p^*(L) = f\left(q(L), m(L), \varepsilon(L), \{y_i\}, \{k_i\}, \{\lambda_i\}, \{\mu_i\}\right) \tag{2}$$

where the *{X}* denotes the distribution of *X*. We are not interested in solving such a model explicitly here, not least because expected local school quality is a very complex notion in the English system where guaranteed local school places rarely exist and the chances of being admitted to many schools is highly uncertain. We simply set out the model here as the basis for our empirical approach.

Before they acquire a child, each family chooses $L^*$ with *k* = 0, so *q(L)* is irrelevant. Given that *q(L)* is factored into the price, they can therefore choose to live somewhere cheaper or with better amenities $\varepsilon$ or labour market prospects *m*. When a child arrives, exogenously, the family will re-compute the optimal *L* given *k* = 1, with *q(L)* now becoming relevant. The new optimum will therefore have a higher *q* as long as any weight at all is placed on school quality ($\lambda > 0$). The decision whether to move or not is based on a comparison of the utility gain with the cost of moving. Given a simple linear additive utility function, the former is simply

$$\Delta U_i = \left(y_i - p_1^*\right)^\delta - \left(y_i - p_0^*\right)^\delta + \lambda_i \Delta q + \gamma \Delta m + \mu_i \Delta \varepsilon \tag{3}$$

where $\Delta$ denotes the change between *k = 0* and *k = 1.* To make things operational in terms of our data, we can approximate the first two terms as

$$\left(y_i - p_1^*\right)^\delta - \left(y_i - p_0^*\right)^\delta \approx \alpha . I\left(y_i < \overline{y}\right) - \beta . \Delta p^* \tag{4}$$

We focus for now on a potential move within a labour market area, so $\Delta m = 0$. We substitute out for $\Delta p^*$ using $\Delta q$, $\Delta\varepsilon$ and location. We end up with a likelihood of moving for each family of

$$pr(\text{moving}) = f(\Delta U_i) = f\left(I\left(y_i < \bar{y}\right), \Delta q, \Delta\varepsilon\right) \tag{5}$$

In the empirical work reported below, we allow for heterogeneity in response to school quality, $q$, specifically in relation to income level. We also use the initial level of school quality, $q_0$, rather than the change in school quality achieved from the move; the actual change can be thought of as the realised value of an expectation conditional on $q_0$.

Understanding when and how households incorporate information about school quality is crucial to the validity of the analysis we undertake. It could be argued that families would foresee having to move later and therefore locate initially near a good school. However, the elapsed time between choosing the first dwelling and the relevant age to be in the "right" house may be a decade and moving costs are not so large as to outweigh the utility loss of living in the wrong place for so long. In practice, secondary school quality enters the household utility function at different times for different households: it could be as early as a first house purchase prior to having children for some couples and as late as the point where choice of school must be expressed on an official admissions form for others.

# 4. Data

## *a) Data on Pupils*

We draw pupil-level data from the National Pupil Database (NPD), an administrative dataset of all pupils in the state-maintained system that is made available to researchers by the Department for Children, Schools and Families. It provides an annual census of pupils, taken each year in January, from 2002 onwards (with termly collections since 2006). This census of personal characteristics can be linked to each pupil's test score history.

The first part of the analysis uses a single cohort of pupils and tracks them from age 5 (2001/2) into the first year of secondary school at age 11 (2007/8). This linked data is only recently available and has not yet been used to analyse the role of neighbourhood formation in secondary school choice. Our analysis uses the subset of pupils that are identifiable within the state-system throughout this period, which amounts to 90% of the cohort (about half a million pupils). Those that are excluded may have attended a private school for a period, may have spent time abroad (including Wales or

Scotland), or may have been entirely educated in the English state system but their Unique Pupil Number was lost during a school transfer.

The second part of the analysis uses a larger set of primary school pupils who are present in the NPD dataset for at least two years between 2002 and 2006. In order to create a group where their secondary school transitions are 'typical' we exclude those living in selective areas where at least 10 per cent of pupils transition to grammar schools. We also exclude areas that still operate a middle school system (with a first transition around the age of 8 or 9 and a second transition at the age of 12 or 13). Also, pupils from some special schools are omitted, as are the small number of pupils with missing values for critical data. The sample for the initial regressions in Table 5 is 7,355,284.

NPD provides data on gender, within-year age, ethnicity, and an indicator of Special Educational Needs (SEN, which measures learning or behavioural difficulties). A key variable for our purposes is eligibility for Free School Meals (FSM), an indicator of family poverty that is dependent on receipt of state welfare benefits (such as Income Support or Unemployment Benefit). Clearly while our FSM variable is probably a very good measure of the FSM status of 19 per cent children who have it, it has been shown by Hobbs and Vignoles (2009) to be a crude measure of household income or poverty status. This means that we should be cautious about simply comparing pupils with FSM across different areas of the country and different micro neighbourhoods.

We can use linked test score data to measure the academic attainment of children in Key Stage two tests at the end of primary school. This serves as a useful proxy of academic success to date in English, maths and science. We take the underlying test score data in each of these subjects and aggregate it, then create indicators for the pupils who scored in the top quartile and lowest quartile in these tests.

### *b) Data on Schools*

We characterise secondary schools in this paper according to their contemporaneous peer intake and their quality. The social and ability profile of the school is measured by aggregated pupil-level data for a particular cohort. We calculate the proportion of pupils at the secondary school who are eligible for free school meals to reflect the social deprivation of the cohort. We also use test score data taken from the end of primary school to measure the proportion of pupils who scored in the top and bottom quartile in these tests.

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 (%5A*-C). 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, 2004). A parent who chooses a school based on the %5A*-C is incorporating weighted information on both the teaching at the school (the value added) and the composition of the school (peer group).

We generate our measures of school quality in two ways. In the first we simply use the contemporaneous GCSE results for the school. Alternatively, we take the average of the %5A*-C measure over our observation window. There is considerable variation in this measure for individual schools across time, and so averaging may help to capture the underlying quality of schools.

### *c) Location and Distance*

Crucially for this analysis, we have access to each pupil's full postcode that we use to (i) describe the local characteristics of the neighbourhood; (ii) identify whether a house move takes place; and (iii) identify their nearest secondary school. This locates them quite precisely, to within 100m (Harland and Stillwell, 2007). We also have the coordinates of schools, which locates them exactly. We rely on the postal geography of the UK for this analysis[2]. Overall, there are about 1.78m unit postcodes covering 27.5m addresses[3]. On average, there are 15 addresses in a unit postcode. A subset of these addresses will house families with children attending state secondary schools.

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 capturing the individual's own household context. We have data at two different scales. Firstly, we have matched pupil's postcodes to the Mosaic classification[4] 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,

---

[2] For further details see: http://www.statistics.gov.uk/geography/postal_geog.asp (accessed 17/05/10).

[3] As of May 2005, in the UK as a whole.

[4] This is commercial geo-demographic data, kindly provided to us by Experian (http://www.experian.co.uk/ accessed 17/05/10).

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.[5] This ranks every super output area (SOA)[6] on a range of criteria (income, employment, health, education and skills, housing, and geographical access to services); we use the overall weighted index.

We have a postcode for every year the child is at school. It is the postcode of the child in year six that is critical to determining secondary schools they will have priority to attend at the point where school allocation is decided. However, we suspect that parents do not always keep schools fully informed of house moves as they happen and so the year seven postcode at their new school may better reflect their location at the time of school allocation.

We identify whether a pupil moves house through any variation of their postcode over time. Given the number of addresses in a postcode mentioned above, the vast majority of house moves will lead to a change in postcode. Of course, house moves occur for many reasons, some of which are related to schooling (for example, moving into the catchment area of a good school) and some of which are not (for example, changes in family circumstances). Certain groups may move more frequently than others, for example, families of those in the armed forces, those in social housing, refugee families and travellers (Dobson et al., 2000). Postcode changes may be caused by factors other than pupils moving house. There may be miscoding errors caused by either change in the postal system or by data-entry mistakes. We attempt to eliminate some of these miscoding errors by using Royal Mail information about postcode terminations/redistricting to try and pick out changes in a pupil's postcode that are not due to moving house:

- We do not count as moves postcode changes that leave the first and last two characters of the postcode unchanged: these are likely to be due to changing postcode classifications;
- We do not count as a move cases where either of the first or last two characters of the postcode only change;

[5] For more information see http://www.communities.gov.uk/documents/communities/pdf/131206.pdf (accessed 17/05/10).

[6] A SOA is a small geographical unit, containing a minimum population of 1000 and a mean of 1500.

- We do not count as a move cases where there are changes in the postcode length by one character, e.g. AB1 becoming AB12 or B12 becoming CB12, with all remaining characters unchanged;
- If all in former postcode moved, and all includes more than 8 pupils (our cut-off for a family), then we do not count this as a move: it is likely to have been a redistricting we have missed;
- We do not count moves of less than 100m: these are most likely due to redistricting that we have not been able to pick up. Even if it is a real move, it is unlikely (although not impossible) that this would change catchment areas;
- We do not count as a move cases where the first or last two characters only are coded in reverse compared with the postcode for the other academic year.

For each pupil postcode location and primary school attended we identify a nearest secondary school and a modal secondary school. The nearest secondary school is computed as that closest using straight-line distances and postcode coordinates (with single-sex school adjustments). 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[7]. The modal secondary school is the school most attended by pupils with the same primary school, sex and FSM status. Where a primary school is split into two institutions – infants (ages 5 to 7/8) and juniors (ages 7/8 to 11) – we link these institutions using transfers information to create pseudo primary schools.

## 5. Results

In this section we first provide descriptions of the extent of house moves by different groups of families using multi-cohort data on all pupils in primary schools. We then move to the smaller dataset of pupils for whom we can track their location from year 1 through to year 7 and document the total impact of changes in residential sorting over this period of time. Finally, we return to the large dataset and estimate models that predict the chances of moving house, given the individual and local characteristics of the current location.

[7] 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.

### *a) Description of extent of house moving*

Data from the pooled cohorts of pupils in primary schools between 2002 and 2006 show that in any particular year, 10 per cent move house. This is a high figure and reflects the changing household composition and wealth of families with young children, as much as the desire to locate close to a good neighbourhood or school. There are important differences within this average: 14% of FSM pupils move house during primary school, compared to 8.5% of non-FSM pupils. This may reflect the very strong correlation of FSM eligibility and use of social housing and other rented accommodation. The average distanced moved is over 15 kilometres, as shown in Table 1, but there is clearly a great deal of variation in the distance of these house moves. The table also shows that moves within a travel-to-work-area (TTWA) are far more frequent than cross-TTWA moves: 8% relative to 1%. The last three rows in the table show that the three different methods we have used to infer the quality of the likely secondary school for a pupil all have a similar distribution.

---------------------Table 1 here---------------------

Analysis of the single cohort for which we have linked data from year 1 to year 7 indicates that 44 per cent of the cohort moves house at some point during primary school, with many moving more than once. There is little regional variation in this figure. Many of these house moves are most likely not motivated by changing secondary school quality since 38% of the movers in this cohort retain the same nearest secondary school following their move.

Table 2 shows the mean changes in available school quality consequent upon moving house. Note that the table averages over all those moving postcode and so will include a substantial fraction of cases where there is no change in the nearest school. It shows that across all pupils, the average house move achieves a small positive improvement of 1.16 percentage points in local school quality, relative to a mean of 46 per cent. There is considerable heterogeneity within this average, however. For non-FSM pupils, the average gain of 1.50 percentage points is six times larger than for FSM pupils, 0.24 percentage points. There are also important differences within region: in the North East, East of England, London and the South East the typical move by a FSM family results in a deterioration in nearest secondary school quality. The disparity in the experiences between FSM and non-FSM pupils is largest in London. Similar results apply if we use either of the other two measures of available school quality.

-------------------------Table 2 here-------------------------

The table tells a similar story for neighbourhood change consequent upon moving, showing how the level of deprivation in the household's local neighbourhood changes for those families that move house during primary school. The figure of -2.18 indicates that the average family sees a modest improvement in neighbourhood quality (IMD mean is 26). This improvement is greater for non-FSM pupils than it is for FSM pupils (-2.67 relative to -0.86) and the disparity in the experiences between these groups is again largest in London. So, there are high levels of house moving taking place by families while their children are in primary school and that this movement produces some advantages for non-FSM households compared to FSM households in nearest school quality and neighbourhood quality.

### *b) Contribution of changes in residential sorting to secondary school segregation*

We can track a single cohort of pupils from year 1 to year 7 to show how substantively important this house moving activity is as a contributor to overall school segregation. This is done by comparing the residential sorting of this cohort at ages 5 and 11 to the secondary school segregation that they actually experience. Rather than describe changes in school quality, Table 3 reports the proportion of pupils who would be FSM at a school based on a nearest secondary school allocation at the age of 5 and the age of 11. This is compared to the actual secondary school FSM proportion they experience. The first row of Table 3 shows that there is little change in the FSM proportion living locally for the FSM pupils who move house between the ages of 5 and 11. The FSM movers attend secondary schools that are more deprived than they would be if this cohort had been forced to attend their nearest secondary schools based on their age 11 postcodes. For FSM pupils who do not move, their local neighbourhood cohort becomes a little more deprived, and again they attend schools that are more deprived still. For non-FSM movers the statistics are exactly reversed with house-moving improving the neighbourhood demographics and the actual FSM proportion at the school they finally attend being a little less deprived still.

---------------------Table 3 here----------------------

We calculate segregation using the Index of Dissimilarity,[8] the most commonly used segregation index, separately for binary indicators for whether the child is (i) eligible for FSM; (ii) scored in the

[8] The formula for the Index of Dissimilarity in an area with *FSM* free school meals pupils and *NONFSM* non FSM pupils and with j schools, each with $fsm_j$ pupils eligible for FSM and $nonfsm_j$ pupils not eligible for FSM is

$$D = \frac{1}{2}\sum_j \left| \frac{fsm_j}{FSM} - \frac{nonfsm_j}{NONFSM} \right|$$

top quartile in Key Stage 2 (KS2) tests; and (iii) scored in the lowest quartile in KS2 tests. A value of one indicates complete separation of two groups and a value of zero indicates no unevenness in the distribution of one particular group compared to another. Rather than calculate residential segregation using an areal sub-unit such as a SOA or a ward, pupils are re-allocated into their nearest secondary school using the age 5 or the age 11 postcode. This is done in order to retain the same sub-unit size for the residential segregation as for the school segregation, which is important since measured segregation for these unit sizes is susceptible to an upward small unit bias (Allen et al., 2009).

The right-hand column in Table 4 shows the level of segregation for this cohort in secondary school in year 7 and the values of 0.380, 0.265, 0.237 for FSM, top quartile KS2 and lowest quartile KS2 respectively indicate moderate levels of school segregation that are consistent with those reported in other studies (e.g. Gibbons and Telhaj, 2007; Allen and Vignoles, 2007). The middle column shows the amount of residential sorting on these pupil's 2008 postcodes by re-allocating the pupils to their nearest schools and re-calculating segregation. The amount of residential segregation is lower, especially for ability, and this is true across all regions and local authorities for this cohort, as found in Allen (2007). The level of residential sorting amongst these pupils at the age of 5 is again lower, but only marginally so. This is consistent with the descriptive analysis that shows that house moving that takes place during primary school advantages non-FSM pupils, but only marginally so. This analysis makes it clear that the primary sources of secondary school segregation in England result from residential sorting that exists before children even begin primary school and from admissions policies that break the link between residence and school attended (producing post-residential sorting).

---------------------------Table 4 here--------------------------

In Figure 1 we show that there is some regional variation in the degree to which house moving contributes to segregation, being greatest in London, and least in the more rural areas. Simple regression analysis (not reported here) shows that areas with greater income inequality tend to see more stratifying house moving.

---------------------------Figure 1 here--------------------------

### *c) Modelling house moves in pooled cohort data*

This section returns to the large dataset of pupils across several cohorts in primary school to explore how the chances of moving relate to household and neighbourhood characteristics. The dataset is very large at 7.36 million observations as this tracks 5 cohorts each of up to 0.5 million pupils for up to 4 years each.   We begin with a simple model for moving house, shown in Table 5. This displays a series of N*T probit analyses, reporting sample average marginal effects, and robust z statistics clustered on SOA. The probit models also include age, gender and ethnicity dummies, but coefficients are not reported.

The top row confirms that families eligible for FSM are more likely to move house. In the final two columns this is of the order of 1.5 percentage points per year. Focussing first on the quality of the nearest secondary school, we see a significantly negative effect on the moving probability. The alternative measure of modal secondary school quality confirms this. These effects are statistically well determined but are quantitatively rather small. We return to discuss the quantitative significance of the effects below for the panel model. The interaction between school quality and FSM status shows a strong differential impact of school quality on moving. The interaction is of the opposite sign and about equal magnitude, implying that available secondary school quality has no effect on the moving behaviour of poor families. If anything, the sum of the two terms is positive suggesting that these families tend to move away from high-scoring schools. This differential behaviour would be potentially significant in driving the formation of communities around high and low quality schools, were it not for the fact that the quantitative impact on non-FSM families is so low (see below). Again, using the alternative measure of school quality yields the same result.

---------------------------Table 5 here--------------------------

Before moving on to the panel analysis of moving, it is useful to show the relationship of strategic moving to the child's age, shown in table 6. This table displays the results of a separate regression for each age group; again the results are sample average marginal effects and other variables as above are included in the regression but not reported. The central finding is that the effect of school quality is essentially constant from age 5 to 10, rising slightly in the final two years of primary school, but then falls to about a third of its value in the first year of secondary school. This strongly suggests that we are indeed picking up school-related moves in the analysis rather than moves based more generally around neighbourhood characteristics. The impact of FSM eligibility is stable over this age range (apart from age 9-10). The interaction term between FSM status and school quality declines along with the direct school quality effect, roughly halving between primary and secondary school ages. Again this suggests that non-FSM-eligible families abruptly cease strategic moving after the pupil is in secondary school, but that FSM-eligible families continue moving and in fact typically move away from higher scoring schools.

-------------------------- Table 6 here --------------------------

### d) Modelling house moves in the panel

In this section we exploit the panel nature of our dataset, tracking families over four years during their primary school years in a dynamic nonlinear model. We do this so that we can attempt to account for some of the unobserved heterogeneity across our families. Including lagged dependent variables in panel models causes the estimators to be inconsistent and possibly badly biased. In the linear case this problem can be solved by differencing to eliminate the unobserved effect, then IV and GMM methods can be used to produce consistent estimators. This method is not possible in nonlinear models. Wooldridge (2005) proposes a parametric method by specifying a distribution for the unobserved heterogeneity conditional on the initial value and the exogenous explanatory variables. This approach essentially implements a random effects probit model with controls for initial state. The Wooldridge model requires the inclusion of the value of each right hand side variable in each year in order to obtain a consistent estimate of the lagged dependent variable (whether the family moved in the previous period). We also include a measure of the cumulative stock of observed moves, an indicator of whether the family moved in the first year of data, and the quality of the nearest school in the initial period.

The results are in Table 7 and the key findings from the previous section are confirmed. Having a low-scoring secondary school as the default school raises a family's probability of moving house. This is a precisely determined effect. However, it is not present for families eligible for FSM: the interaction of FSM status and school quality is of the opposite sign to the main effect of school quality and of greater magnitude.

-------------------------- Table 7 here --------------------------

We include the full set of Wooldridge controls from column 2 onwards. Note that once these are included, FSM status becomes negative: families who are typically poor (and live in social housing) move more often, but for any one such family, they are somewhat less likely to move if they cease to be eligible for FSM. We also include a group of variables to absorb heterogeneity in families' propensity to move and to address the initial conditions problem. These variables are: whether the family moved last year, whether they moved in the first period, the quality of the initial default school, and the stock of moves completed at each date. In particular, conditioning on the quality of the initial default school deals in part with the problem that we do not see families until the child is aged 5, and they may have moved to be near their preferred school before that age.

Whilst the analysis produces statistically significant effects, the quantitative effect is low. The table gives sample average marginal effects, and a unit standard deviation fall in school quality induces an increase in the probability of moving of less than one percentage point (0.659), against a mean annual moving probability of 10 percentage points (table 7, column 6).

We conclude this section by considering the main weaknesses of this model. Given that this is an administrative dataset focussed on education with little information on the family, it is unreasonable to expect to the model to explain a lot of moving behaviour, and indeed the pseudo R2s in Table 5 are low, but the key question is whether any of the omitted variables are biasing the results. We have included measures of the local neighbourhood and have experimented with a variety of other measures not reported here. We are using a single variable to summarise the desirability of schools, because that is the focus of this analysis; many other characteristics (such as social composition) are correlated with this measure. We have repeated observations on families over four years, which we deal with by using a random effects approach, absorbing heterogeneity using Wooldridge's (2005) technique. Reverse causation (moving causing school quality) is very unlikely to be an issue: the dating of the quality measure is before the family moves, and the movers are small fraction of any school. The main problem that we face is an initial conditions issue. If a family's location before they have a child is chosen independent of school quality, then there is only the practical problem that our data starts at age 5. But if some families move to live near a good default school either before having a child, or before we see them in the data from age 5, then the quality of the initial school is not random as it results from previous choices. The non-random group of families moving early are likely to be ones that have a high preference for schooling. We would therefore expect to see a greater effect if we could capture more and earlier parts of families' lives. We can in fact show this within our own data window: the size of the coefficient of initial location school quality is larger for non-poor families when their child is younger. The implication of this is that we may underestimate the effect of school quality on moving and so on neighbourhood formation, but if so this additional effect occurs before the child is aged 5 or even before the child is born.

## 6. Conclusions

The idea of families moving house to secure access to high-scoring schools is strongly embedded in the conventional wisdom about school admissions in England. The consequent "school selection by mortgage" is considered to be one of the inequities in the way that children are allocated to schools. Indeed, we expected our research in this paper to document and further explore this phenomenon. To a degree, we have confirmed these relationships; but we have also shown them to be quantitatively very weak.

We use a large dataset taken from the pupil census in England. First, we follow a single cohort of some half a million children as they approach age 11, the key secondary school admission age. We measure changes in this cohort's degree of social segregation through this process. We expect to see increasing segregation as richer households tend to move to cluster around the popular schools. Segregation does increase overall and in most regions, but the increase is very slight, and only a small fraction of the average level of segregation. For example, across England as a whole the change in residential segregation measured using a free school meals dissimilarity index is just 0.02 between the ages of 5 and 11, contributing only a little to overall school segregation of 0.38. Second, we combine a number of cohorts and model moving directly by estimating a dynamic nonlinear model with unobserved effects, following Wooldridge's (2005) approach. We specifically allow for heterogeneity by income in the response of the moving propensity to local school quality. We find that for the non-poor families, moving is significantly negatively correlated with initial location school quality; for the poor however, there is no relationship. This differential response is interesting and would lead to increasing social segregation as it plays out. However, quantitatively, the response is very minor: a unit standard deviation change in local school quality for non-poor families changes the probability of moving by 0.6 of a percentage point (against a mean of ten percentage points).

We have tried alternative ways of defining a family's "default" local school, and attempted to deal with the unobserved neighbourhood and family effects. Nevertheless, there may remain biases due to un-captured differences in preferences for schooling, preferences for very local neighbourhood amenities and so on. There is another caveat that needs to be borne in mind. We take school performance as given in this estimation. A general equilibrium model of school performance and neighbourhood formation would be valuable but is beyond the scope of this paper.

Our results have implications for the policy debate about school admissions and access, and implications for the formation of neighbourhoods. Significant inequalities in school access have been demonstrated for England and a major part of this inequality is accounted for by differences in the places that rich and poor families live. Proximity to a good school is the key determinant, so an important question is how those advantageous locations arise. Our results here suggest that the house-moves of families during school years only make a small contribution to this difference. Most of the rest of it derives from families being in the right place before the relevant child is aged 5. This is a rather different picture of the evolution of socioeconomic inequality in schools. Policies to reform school admissions and the role of proximity as a tie-breaker need to be based on a deeper understanding of this very early segregation.

We know little about why families are already moderately highly segregated by the time their children are 5 years old and, importantly from the perspective from school admissions reforms, we do not know how many families move house to access their choice of secondary schools in the very early years of their child's life. It seems unlikely that large numbers of parents are precisely 'choosing' secondary schools at such an early age, but they may be choosing to move to more affluent or rural neighbourhoods on the approximate assumption that the local secondary school will be acceptable.

Our findings show that there exists some inertia in the system because families do not appear to use house moves to optimise their school choice decision to the extent we might expect. Unfortunately we cannot use this analysis to determine whether this is because parents do not place a high value on school quality, relative to other factors including neighbourhood affiliation or instead whether the transaction costs of moving house are too high to compensate the improvement in school quality. Obviously neighbourhoods are subject to a variety of factors influencing the evolution of their composition. Local school quality is often thought of as being an important factor, and as a mutually reinforcing factor as high income neighbourhoods will tend to produce high scoring students, thus maintaining the school's public performance measure. Our results suggest that this school quality mechanism for creating communities is rather weak; that whilst families do leave neighbourhoods with poor schools, this is only a relatively minor factor in determining moving probabilities.

## References

Allen, R. and Vignoles, A. (2007) What should an index of school segregation measure?, *Oxford Review of Education* 33(5)643-668.

Allen, R. (2007) Allocating pupils to their nearest school: the consequences for ability and social stratification, *Urban Studies*, 44(4)751-770.

Allen, R., Burgess, S. and Windmeijer, F. (2009) *More reliable inference for segregation indices*, CMPO discussion paper 09/216.

Black, S. (1999) Do better schools matter? Parental evaluation of elementary education, *Quarterly Journal of Economics*, 114(2), pp. 577-599.

Bayer, P., Ferreira, F. and McMillan, R. (2007) A Unified Framework for Measuring Preferences for Schools and Neighbourhoods, Journal of Political Economy, 115, (4), 588-638.

Bogart, W. T. and Cromwell, B. A. (1997) How much more is a good school district worth? *National Tax Journal*, 50:215–232.

Bogart, W. T. and Cromwell, B. A. (2000) How much is a neighbourhood school worth? *Journal of Urban Economics*, 47:280–305.

Brunner, E. J. and Imazeki, J. (2008) Tiebout choice and universal school vouchers, *Journal of Urban Economics*, 63(1)253-279.

Burgess, S., Johnston, R., Key, T., Propper, C. and Wilson, D. (2008) The transition of pupils from primary to secondary school in England, *Transactions of the Institute of British Geographers,* 33(3)388-403.

Burgess, S. and Briggs, A. (2010). *School assignment, school choice and social mobility*. Forthcoming, *Economics of Education Review*.

Burgess, S., Greaves, E., Vignoles, A. and Wilson, D. (2009) *Parental choice of primary school in England: what 'type' of school do parents choose?*, CMPO discussion paper 09/224.

Cheshire, P. C. and Sheppard, S. C. (1995) On the price of land and the value of amenities, *Economica*, 62:247–267.

Coldron, J., Tanner, E., Finch, S., Shipton, L., Wolstenholme, C., Willis, B., Demack, S., and Stiell, B. (2008) *Secondary school admissions,* DCSF Research Report, RR020.

Dobson, J., Henthorne, K. and Lynas, Z. (2000) *Pupil Mobility in Schools: Final Report*, London: UCL Migration Research Unit.

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

Epple, D. and Romano, R. E. (2003), Neighbourhood schools, choice, and the distribution of educational benefits, in Hoxby, C. M. (ed) *The economics of school choice*, pages 227–286, University of Chicago Press, Chicago.

Ferreyra, M. M. (2007) Estimating the effects of private school vouchers in multidistrict economies, *The American Economic Review,* 97(3) 789-817.

Gibbons, S. and Telhaj, S. (2007) Are Schools Drifting Apart? Intake Stratification in English Secondary Schools, *Urban Studies*, 44(7) pp. 1281-1305.

Gibbons, S. and Machin, S. (2003) Valuing English primary schools, *Journal of Urban Economics*, 53, pp. 197-219.

Goodman, A. C. and Thibodeau, T. G. (1998) Housing market segmentation, *Journal of Housing Economics*, 7:121–143.

Gorard, S., Taylor, C., and Fitz, J. (2003) *Schools, markets and choice policies,* RoutledgeFalmer, London.

Harland, K. and Stillwell, J. (2007) *Commuting to School in Leeds: How useful is the PLASC?,* School of Geography, University of Leeds.

Hobbs, G. and Vignoles, A. (2009) Is children's free school meal 'eligibility' a good proxy for family income? *British Educational Research Journal,* forthcoming.

Leech, D. and Campos, E. (2003) Is comprehensive education really free? A case study in the effect of secondary school admissions policies on house prices in one local area, *Journal of the Royal Statistical Society Series A*, 166:135–154.

Levačić, R., Jenkins, A., Vignoles, A., Steele, F., and Allen, R. (2005) *Estimating the relationship between school resources and pupil attainment at key stage 3*, DfES Research Report 679.

Nechyba, T.J. (1999) School Finance-Induced Migration and Stratification Patterns: the Impact of Private School Vouchers, *Journal of Public Economic Theory*, Vol. 1.

Nechyba, T.J. (2000) Mobility, Targeting and Private School Vouchers, *American Economic Review,* 90(1), 130-46.

Nesheim, L. (2002) *Equilibrium sorting of heterogeneous consumers across locations: Theory and empirical implications,* CeMMAP Working Paper, 08/02.

Rosenthal, L. (2003) The value of secondary schooling, *Oxford Bulletin of Economics and Statistics*, 65:329–355.

Rothstein, J. (2003) Good Principals or Good Peers? Parental Valuation of School Characteristics, Tiebout Equilibrium, and the Incentive Effects of Competition Among Jurisdictions, *American Economic Review,* 96(4) 1333-1650.

Sattinger, M. (1980). *Capital and the distribution of labor earnings*. North-Holland, Amsterdam.

Sieg, H., Smith, V. K., Banzhof, H. S., and Walsh, R. (1999) The role of optimizing behavior in willingness-to-pay estimates for air quality, *American Journal of Agricultural Economics*, 81:1112–1117.

Tiebout, C.M. (1956) A Pure Theory of Local Expenditures, *Journal of Political Economy*, 64(5) pp. 416–24.

Tinbergen, J. (1959) On the theory of income distribution, In Klaassen, L. H., Koyck, L. M., and Witteveen, H. J., editors, *Selected papers of Jan Tinbergen*. North-Holland, Amsterdam.

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

West, A. and Barham, E. and Hind, A. (2009) *Secondary school admissions in England: policy and practice*, Research and Information on State Education Trust, London, UK.

Wilson (2004) Which ranking? The impact of a "value added" measure of secondary school performance, *Public Money and Management*, 24 (1) 37-45.

Wooldridge, J. (2005) Simple Solutions to the Initial Conditions Problem for Dynamic, Nonlinear Panel Data Models with Unobserved Heterogeneity, *Journal of Applied Econometrics* 20, 39-54.

**Table 1: Regression sample descriptives**

| Variable | Mean | Standard Deviation | Minimum | Maximum |
|---|---|---|---|---|
| FSM (%) | 19 | | | |
| Move (%) | 10 | | | |
| Move within TTWA (%) | 8 | | | |
| Move across TTWA (%) | 1 | | | |
| Distance Moved (Metres) | 15,284 | 49,525 | 100 | 601,643 |
| Index of Multiple Deprivation (IMD) | 25.94 | 17.72 | 0.59 | 86.36 |
| Nearest School's Quality | 45.88 | 18.31 | 0 | 100 |
| Average of 3 Nearest School's Quality | 46.63 | 12.27 | 8 | 92.67 |
| Modal School Quality | 46.33 | 16.74 | 4 | 100 |

Note: School Quality is the proportion of pupils obtaining 5+ A*-C at GCSE
Move within TTWA is a change of postcode, but both postcodes are in the same TTWA

**Table 2: Change in nearest school and neighbourhood quality for pupils who move**

| | Mean Change in Nearest School Quality | | | Mean Change in Neighbourhood Quality | | |
|---|---|---|---|---|---|---|
| **Region** | **Non-FSM Pupils** | **FSM Pupils** | **All Pupils** | **Non-FSM Pupils** | **FSM Pupils** | **All Pupils** |
| North East | 0.72 | -0.14 | 0.44 | -3.30 | -0.95 | -2.52 |
| North West | 1.16 | 0.41 | 0.95 | -3.20 | -1.12 | -2.61 |
| Yorkshire | 1.61 | 0.50 | 1.32 | -2.87 | -0.96 | -2.37 |
| East Midlands | 1.86 | 1.13 | 1.67 | -2.51 | -0.44 | -1.96 |
| West Midlands | 1.69 | 0.37 | 1.29 | -3.59 | -1.56 | -2.97 |
| East of England | 1.22 | -0.12 | 0.99 | -1.06 | 0.28 | -0.82 |
| London | 2.51 | -0.07 | 1.57 | -4.32 | -0.85 | -3.06 |
| South East | 1.01 | -0.21 | 0.82 | -0.67 | 0.30 | -0.52 |
| South West | 1.23 | 0.33 | 1.08 | -1.13 | -0.49 | -1.02 |
| **All** | **1.50** | **0.24** | **1.16** | **-2.67** | **-0.86** | **-2.18** |

Note: School Quality is the time-invariant proportion of pupils obtaining 5+ A*-C at GCSE
Neighbourhood quality is the Index of Multiple Deprivation score for the pupil's area of residence

**Table 3: FSM proportions for nearest school allocations and actual school**

| | 2002 residential (nearest school) FSM % | 2008 residential (nearest school) FSM % | Actual secondary school FSM % |
|---|---|---|---|
| FSM movers | 22.4% | 23.0% | 25.7% |
| FSM non-movers | 23.2% | 24.9% | 27.1% |
| Non-FSM movers | 15.1% | 13.7% | 13.5% |
| Non-FSM non-movers | 13.6% | 14.2% | 13.6% |

**Table 4: Levels of residential and school segregation for single linked cohort**

| | 2002 nearest school segregation | 2008 nearest school segregation | 2008 secondary school segregation |
|---|---:|---:|---:|
| FSM | 0.322 | 0.340 | 0.380 |
| Top KS2 ability | 0.177 | 0.182 | 0.265 |
| Lowest KS2 ability | 0.176 | 0.181 | 0.237 |

## Table 5: Initial regressions of the probability of moving

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| FSM Status | 0.046 | 0.030 | 0.029 | 0.019 | 0.015 |
| | (81.91)** | (59.84)** | (58.38)** | (18.04)** | (13.09)** |
| Nearest Secondary School's Quality | | -0.011 | | -0.016 | |
| | | (9.80)** | | (14.29)** | |
| Modal Secondary School's Quality | | | -0.017 | | -0.024 |
| | | | (14.01)** | | (18.85)** |
| Nearest Secondary School's Quality Interacted with FSM | | | | 0.022 | |
| | | | | (10.23)** | |
| Modal Secondary School's Quality Interacted with FSM | | | | | 0.034 |
| | | | | | (13.88)** |
| IMD Score of LLSOA | | 0.003 | 0.003 | 0.004 | 0.003 |
| | | (16.67)** | (15.51)** | (16.74)** | (15.97)** |
| Mosaic Postcode Classification Dummies? | N | Y | Y | Y | Y |
| Pseudo R2 | 0.01 | 0.02 | 0.02 | 0.02 | 0.02 |
| Obs. | 7,355,284 | 7,355,284 | 7,355,284 | 7,355,284 | 7,355,284 |

Notes:
a) Coefficients reported are sample average marginal effects. Robust z statistics, clustered on LLSOA, in parentheses.
b) *Significant at 5%. **Significant at 1%.
c) Dependent variable equals 1 if the pupil changes postcode, 0 otherwise. Estimation method is univariate probit.
d) Age dummies, ethnicity dummies and a gender dummy included but not reported.
e) School Quality is defined as the proportion of pupils obtaining 5 GCSEs at grade C or above.

## Table 6: Regressions of the Probability of Moving by age

| Age of child: | Nearest School's Quality | | Interaction of FSM Status and Nearest School's Quality | | FSM Status | |
|---|---|---|---|---|---|---|
| 5-6 | -0.017 | (9.2) | 0.029 | (8.6) | 0.021 | (12.1) |
| 6-7 | -0.017 | (8.8) | 0.024 | (6. 9) | 0.021 | (12.2) |
| 7-8 | -0.016 | (8.9) | 0.019 | (5.7) | 0.020 | (13.0) |
| 8-9 | -0.014 | (8.2) | 0.014 | (4.3) | 0.022 | (13.4) |
| 9-10 | -0.021 | (11.7) | 0.028 | (8.4) | 0.011 | (6.9) |
| 10-11 | -0.019 | (9.1) | 0.013 | (3.6) | 0.017 | (9.4) |
| 11-12 | -0.004 | (2.8) | 0.012 | (4.0) | 0.020 | (12.4) |
| 12-13 | -0.007 | (4.5) | 0.016 | (5.2) | 0.017 | (10.9) |
| 13-14 | -0.007 | (4.5) | 0.013 | (4.2) | 0.017 | (10.6) |
| 14-15 | -0.005 | (3.7) | 0.012 | (4.1) | 0.015 | (9.8) |

Notes:
Each row reports the results of a separate regression. Coefficients reported are sample average marginal effects.
Dependent variable equals 1 if the pupil changes postcode, 0 otherwise. Estimation method is univariate probit.
Age dummies, gender, ethnicity, and mosaic postcode classification dummies are included but not reported.
School Quality is defined as the proportion of pupils obtaining 5 GCSEs at grade C or above.

## Table 7: Panel Regressions (Sample of pupils with 4 observations): Probability of moving

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| FSM Status | 0.007 | -0.027 | -0.027 | -0.029 | -0.029 | -0.029 |
| | (6.18)** | (22.27)** | (22.24)** | (22.53)** | (15.66)** | (15.65)** |
| Nearest Secondary School's Quality | -0.016 | -0.031 | -0.031 | -0.035 | -0.035 | -0.036 |
| | (14.11)** | (15.01)** | (14.94)** | (15.94)** | (11.51)** | (11.59)** |
| Quality*FSM Status | 0.031 | 0.045 | 0.044 | 0.049 | 0.048 | 0.048 |
| | (13.96)** | (13.72)** | (13.69)** | (14.09)** | (9.92)** | (9.93)** |
| Lag of Move Variable ($y_{t-1}$) | | 0.018 | 0.018 | 0.060 | | |
| | | (24.22)** | (24.21)** | (62.58)** | | |
| Move in First Period ($y_0$) | | 0.050 | 0.050 | | 0.005 | |
| | | (71.95)** | (71.88)** | | (4.05)** | |
| Quality of Initial Nearest School | | | -0.006 | -0.009 | | |
| | | | (4.33)** | (5.85)** | | |
| Stock of Moves = 1 | | | | | 0.054 | 0.059 |
| | | | | | (36.38)** | (79.62)** |
| Stock of Moves = 2 | | | | | 0.115 | 0.125 |
| | | | | | (31.25)** | (52.71)** |
| Stock of Moves = 3 | | | | | 0.179 | 0.195 |
| | | | | | (18.57)** | (22.38)** |
| IMD Score of LLSOA | 0.001 | 0.002 | 0.002 | 0.003 | 0.003 | 0.003 |
| | (59.35)** | (76.77)** | (76.78)** | (81.73)** | (61.83)** | (63.76)** |
| Values of RHS Variables in Each Time Period Included? | N | Y | Y | Y | Y | Y |
| Observations | 2595244 | 1946433 | 1946433 | 1946433 | 1946433 | 1946433 |
| Number of pupils | 648811 | 648811 | 648811 | 648811 | 648811 | 648811 |

*Notes:*

1. Coefficients reported are sample average marginal effects. *Z* statistics, in parentheses. *Significant at 5%. **Significant at 1%.
2. Dependent variable equals 1 if the pupil changes postcode, 0 otherwise. Estimation method is random-effects probit.
3. School Quality is defined as the proportion of pupils obtaining 5 GCSEs at grade C or above.
4. Stock variable is the sum of lagged moves. This has a maximum value of 3 in the fourth observation window, if a pupil has moved in the three previous observation windows.
5. Year dummies included but not reported.

**Figure 1: Regional variations in the contribution of moving**