Stephen Gibbons, Olmo Silva
Urban density and pupil attainment

Article (Accepted version)
(Refereed)

http://eprints.lse.ac.uk

URBAN DENSITY AND PUPIL ATTAINMENT

Stephen Gibbons*, Olmo Silva**

March 2007

Author for correspondence:
Stephen Gibbons,
Department of Geography and Environment
London School of Economics
Houghton Street
London
WC2A 2AE
+44 (0)20 7955 6245
Email s.gibbons@lse.ac.uk

Abstract

We explore the association between urban density and pupil attainment using three cohorts of pupils in schooling in England. Although – as widely recognised – attainment in dense urban places is low on average, this is not because urban environments disadvantage pupils, but because the most disadvantaged pupils with low average attainments attend the most urbanised schools. To control for this, we exploit changes in urban density faced by pupils during compulsory transition from Primary to Secondary school, and measure educational progress at the end of the Secondary phase, relative to attainment at the end of Primary schooling. Our results suggest that there are small but significant benefits from education in schools in more densely urbanised settings. We detect this density advantage even amongst pupils moving relatively short distances between Primary and Secondary schools within urban areas, so we cannot attribute it to broad urbanisation effects experienced by pupils making rural-urban school moves. A more likely explanation lies in greater school choice and competition between closely co-located educational providers.

1. Introduction

City schools in the UK, US and elsewhere are widely admonished by politicians, school inspectors, parents and academics for failing to provide pupils with high standards of education (Machin and Vignoles [33], Murnane [36], Neal [39], Ofsted [41,42]). With the aim of redressing this situation, governments in many countries have targeted schools in disadvantaged urban areas. Just a few recent policy interventions include the 'Excellence in Cities' program (Machin, McNally, and Meghir [32]), and the 'Aimhigher' initiative (Emmerson, Frayne, McNally, and Silva [10]) in the UK, and the STAR experiment (Hanushek [20], Krueger [28]) in the US.

In some ways, this critique comes as something of a surprise to economists interested in the benefits of urbanisation and agglomeration. City schools can potentially draw from a diverse pool of high-quality teachers and are geographically placed in settings which should provide incentives through competition with other schools and which facilitate sharing of teaching know-how through cooperation. Of course, given the concentration of poverty in cities, it is also recognised that urban schools are building human capital on a lower base in terms of pupil characteristics, in particular higher rates of hardship and lower initial ability. However, these facts rarely seem to be taken into account when drawing inferences about the effectiveness of urban schools. The simple observation of a high concentration of low achievers in dense urban places is taken as evidence that urban environments disadvantage pupils.

This paper uses a census of over 1.2 million pupils in England, matched to records on their academic progress, to assess whether pupils in city schools really show low-educational progress relative to students in schools in lower density suburban, semi-rural and rural areas. We make use of the fact that pupils in England have to switch schools at the end of the Primary phase when they start their Secondary education, and we look at how a pupil's educational progress after this compulsory transition depends on the change in contextual urban density experienced as a result of the change in school. From these density changes we infer small but significant benefits from education in schools

in more densely urbanised settings. Pupils in schools in relatively dense places – measured in terms of school density and other urban indicators – progress faster than others in their cohort, although the elasticity is low, at around 0.02. We find this association even amongst pupils who move relatively short distances between Primary and Secondary schools *within* urban areas. Therefore, we argue, our results do not emerge from broad agglomeration and urbanisation effects experienced by pupils making rural-urban school moves. Instead, we interpret our findings as providing evidence on the effects of greater school choice, inter-school competition or cooperation between closely co-located institutions in more urban settings.

The paper has the following structure. The next section outlines some relevant literature and sets the work in the context of the studies on agglomeration economies. Section 3 sets out our empirical approach, while Section 4 describes data and institutions. Section 5 presents and discusses our results, while some concluding remarks are provided in Section 6.

2. The literature

"Conceptually, a city is just a dense agglomeration of people and firms. All of the benefits of cities ultimately come from reduced transport costs for goods, people and ideas" (Glaeser [16]). This simple intuition, borrowed from the seminal writings of Marshall [34, 35], is nowadays at the core of most research on agglomeration and urbanisation processes. In a nutshell, the fundamental reason why firms and workers concentrate in geographically contained areas, giving rise to cities, is because spatial clustering generates some form of external economy of scale. As discussed in Glaeser [16] and Rosenthal and Strange [44], agglomeration benefits emerge from proximity in three factors: *People*, associated with labour market pooling and accessibility to a wider sets of customers; *Goods*, coupled with input sharing and specialization of services for producers; *Ideas*, linked to the

emergence of knowledge and technological spillovers across firms[1]. Technological spillovers are also often invoked as an explanation for the reason why cities not only exist, but prosper and grow too.

For these reasons, the idea that urban schools fail to provide pupils with a high standard environment for learning comes as something of a surprise to researchers interested in economies of urbanization and agglomeration. City schools have the potential to exploit dense labour markets, therefore attracting high-quality teachers. They are located in diversified areas where pupils and teachers can be more efficiently matched through the exercise of school choice. Pupils in urban schools may also benefit from learning spillovers associated with closer connections to a larger and more diverse group of students. Further, city schools share better common infrastructures, e.g. faster connections for information technologies or public transport. Finally, city schools are geographically placed in settings that should provide incentives for improvement or adoption of new teaching technologies through competition with other schools, and that facilitate sharing of teaching know-how through cooperation. What evidence is available?

A growing body of research on the functioning of the labour market for teachers has been assembled over the past years. Using different methods and data, Dolton and van der Klaauw [9], Hanushek and Rivkin [21], Hanushek, Kain, and Rivkin [22], Murnane and Olsen [37] and [38] show that individuals respond to (relative) wage incentives in their decision to start teaching or leave this occupation, and that teaching in urban areas might be a poor option compared to more remunerative alternatives available in these locations. Chevalier, Dolton, and McIntosh [5] and Hanushek, Kain, and Rivkin [22] also report that working conditions in urban schools are perceived to be worse than in provincial areas, and that teachers changing schools within urban districts might

---

[1] While proximity of goods and individuals is a natural requirement for of economies of scale to emerge in urban environments, knowledge spillovers might not constrained by distance or within the boundaries of urban areas. However, Jaffe, Trajtenberg, and Henderson [29] provide evidence on the importance of proximity for know-how diffusion too.

seek out schools with easier to teach pupils, such as private schools. Whether this sorting results in lower teacher quality in urban state schools is unclear. Hanushek, Kain, O'Brien, and Rivkin [23] suggest that the opposite is true: teachers moving from Texan urban state schools to suburban/rural schools, as well as teachers leaving the state education system, are on average less effective at raising pupil attainments than teachers who stay. Additionally, Clotfelter, Glennie, Ladd, and Vidgor [6] show that a pecuniary bonus granted to the most qualified teachers in North Carolina greatly reduced their hazard rate for exit from high-poverty urban schools; similar evidence is provided by Hanushek and Rivkin [21] and Lavy [30]. Overall, it seems that city schools are in a favourable position to exploit urban labour markets to hire and retain high-quality educators.

There also appears to be greater scope for choice and competition among schools in urban areas. A large number of closely located schools implies that parents have a wider set of schools to choose from within feasible travel-distance. If pupils with many available schools can be more efficiently matched to educational providers that suit their preferences and capabilities, average educational standards should increase. Moreover, in a school market where parents can exercise choice and funding follows pupils (as in the UK setting)[2], schools have to provide the 'quality' that parents demand or face falling enrolment, loss of money and ultimately closure. The empirical case on the benefits of choice and competition is however mixed. On the one hand, research by Cullen, Jacob, and Levitt [7] and [8] shows no performance gains associated with greater parental choice in the Chicago urban setting. On the other hand, Hoxby [25], [26] and [27] finds that competition in US metropolitan areas is beneficial to pupil achievements. In England, Gibbons, Machin, and Silva [13] provide evidence that Primary school competition is not generally associated with performance gains for state school pupils, except for a minority enrolled in specific types of school that enjoy more freedom in managing their governance and admission practices.

---

[2] There is no strict attendance zoning in most of England: pupils can in principle attend any non-oversubscribed school of their choice, without restrictions regarding place of residence and school proximity (more detail in Section 4).

While inner-city school competition is commonly credited with inducing teaching-related innovations and high school standards, know-how externalities may also emerge because closer proximity of schools facilitates cooperation and sharing of teaching practices and technology among neighbouring institutions. In fact, this is the rationale behind a recent policy initiative of the UK government – the Beacon school scheme – providing incentives for outstanding institutions to work in partnership with other neighbouring schools to help those achieving similar standards[3].

Why, then, are city schools so commonly associated with failure? On the one hand, this could just be because cities have a high concentration of children with fewer home resources and family background disadvantages who on average do worse at school (Glaeser, Kahn, and Rappaport [17]). Whilst this observation suggests that urban schools are building human capital on a lower base in terms of pupil characteristics, it does not have any bearing on the educational progress of students in city areas or on the effectiveness of urban schools in educating their pupils. On the other hand, there are reasons why pupils might indeed do less well if they attend an inner-city school rather than a suburban or rural one. Some of these are analogous to the dispersion forces that appear in standard agglomeration theories. Many congestion-related factors that accompany high urban density are likely to be detrimental to pupil learning: overcrowding in schools and supporting services (such as libraries), high levels of property crime, violence and other social/emotional problems that cause disruption directly and through peer group influence, and high pupil turnover because of demographic mobility. The reports by Ofsted [42] and Lupton [31] present a range of such features that are common to schools in urban areas in England, most of which could be broadly considered as negative peer-group effects.

---

[3] Under the Beacon program, schools that deliver outstanding teaching and are well managed are awarded a 'beacon' status (renewable every third year) to highlight examples of successful practice. Beacon schools are expected to organizing meetings, cross-institution working and pastoral support, as well as increasing teacher participation and retention, to help neighbouring schools improve their education standards. See GHK [12].

Ultimately, whether the combination of positive and negative factors characterizing city schools is beneficial or detrimental to pupil educational progress is an empirical question. Although levels of attainment in dense urban areas might be low because the most disadvantaged pupils attend the most urbanised schools, urban areas might provide better learning environments in terms of pupil academic progress, relative to lower density suburban, semi-rural and rural areas. This issue is rarely taken into account when assessing the performance of pupils in urban schools, and is the empirical problem which we tackle in the next sections.

3. Empirical methods: the value added model of attainment

Our aim is to study the influence of urban density on pupil attainment in schools in England. We will investigate this relationship in the context of compulsory-age Secondary schooling between the ages of 11 and 16. At the beginning of this period, nearly all pupils in the state sector in England switch schools as they move from the Primary to Secondary phase. Our identification strategy will exploit changes in school setting that occur on this transition. *Compulsory* school-transition offers an advantage over most empirical strategies that exploit *voluntary* changes initiated by movers. Everyone here is a 'mover', so that the problem of endogeneity of the choice and direction of move is less acute.

In our empirical model, attainment y_{ij} of individual i in school j depends on unobserved individual characteristics (a_i) that are constant within individuals across schools and with age, observed school characteristics (z_j) and random individual-school specific factors (ε_{ij}); we also allow for the possibility that an individual's trend in attainment with age t depends on observable personal characteristics x_i:

$$y_{ij} = \beta u_{ij} + a_i + \gamma x_i t + \lambda z_j + \varepsilon_{ij} \tag{1}$$

The key variable of interest is a measure or vector of measures u_j describing the urbanisation of the environment in which a school is located. The exact form of these proxies will be discussed later

(in Section 4), but the intention is to capture the general aspects of density, agglomeration and urbanisation that may influence pupil attainment. These forces may act through greater school accessibility and competition, a deeper pool of good teachers in urban labour market, better inter-school networking, more efficient matching of pupils and teachers to schools, broadly defined 'neighbourhood' effects, such as role models and expectations, and much else which might lead pupils to perform better (or worse) in places where there are more schools, more infrastructure and more people. In fact, we make no definite attempt at separating the impact of these indices on pupil educational achievement from other unobservable characteristics of urban schools, as we want to capture general efficiency/quality benefits associated with attending a school in dense urban areas. Our main concern, instead, is to take care of individual unobservable heterogeneity which may simultaneously drive attainment and school of choice, creating a spurious link between measures of urbanization and pupil educational performance. Notice that this model is an analogous (but more general) set up to that typically used to study agglomeration economies in firm or aggregate productivity, where the dependent variable is individual pupil attainment, rather than productivity.

Consistent estimation of β in (1) is not straightforward in the cross section because unobserved individual factors (such as family income and various forms of advantage/disadvantage) are highly correlated with choice of residential location, and hence choice of school and the urban density of its surroundings. However, since an pupil is observed in two or more schools, at different ages − namely Primary school j and Secondary school k − it is (at least) possible to difference out fixed individual factors using the familiar transformation:

$$y_{ik} - y_{ij} = \beta(u_{ik} - u_{ij}) + \gamma x_i \bar{t} + \lambda(z_k - z_j) + (\varepsilon_{ik} - \varepsilon_{ij}) \tag{2}$$
$$\Delta y_{i,jk} = \beta \Delta u_{i,jk} + (\gamma \bar{t}) x_i + \lambda \Delta z_{jk} + \Delta \varepsilon_{i,jk}$$

Equation (2) is a pupil educational production function in 'value-added' form, where the influence of education in an urban setting leads to a gain (or loss) in attainment when a pupil changes school on compulsory transition from the Primary to Secondary phase. Here, $\bar{t}$ is the number of periods

between our observation of the pupil in school j and k. In model (2), it is still possible that the changes in unobservable individual-specific factors $\Delta\varepsilon_{i,jk}$ are correlated with the change in urban density. However, we can partly control for this source of bias by estimating fixed effects models that allow for Primary or Secondary school influences on pupil progress that are common across pupils within a school. Identification is achieved using either the variation in the change of urban density for pupils going to the same Secondary but coming from different Primary schools, or the variation in the change of urban density for students attending the same Primary and moving to a different Secondary school. We therefore condition out school-specific, and school location-specific influences on attainment growth that are common to all individuals within a school, including unobservable individual school preferences shared by schoolmates (components of $\Delta\varepsilon_{i,jk}$). Additionally, given that few Primary to Secondary school transitions involve long-distance geographical mobility, school fixed effects also control for broader agglomeration effects which are common to both the Primary and Secondary phase. Note then, that we will mainly identify our effects off the change in the density between the urban environments in the immediate proximities of the two schools (Primary and Secondary) that each pupil attends.

To go one step further, we also estimate models controlling for Primary or Secondary school influences, and including residential neighbourhood fixed effects. This can be done using detailed information on individual's home postcodes (corresponding to 10-12 contiguous housing units) and allows us to control for unobservable characteristics, such as income or preferences over local amenities and schools, common to families sorting into the same small residential neighbourhood. The underlying assumption for obtaining an unbiased estimate of β after controlling for either Primary or Secondary school influences and home postcode fixed effects, is that the differences in urban density between the Primary and Secondary phase are not systematically correlated with a change in unobservable pupil characteristics that drives attainment growth between the two phases, but only reflect changes in school quality/effectiveness associated to more or less dense

environments. In the analysis that follows, we will spend considerable effort to assess the validity of our assumptions.

Another crucial assumption for our strategy to work is that pupils must have a range of schools that they can choose to attend, given where they live. This condition is essential if there is to be any variation in Secondary school urban density, conditional on Primary school and postcode of residence (or variation in Primary school urban density conditional on Secondary school and postcode of residence). In fact, in the UK, unlike the US, parents and pupils are free to choose from a number of schools and there is not anything like a one-to-one mapping from place of residence to school attended. We give more details in the next Section.

In conclusion, our main empirical results are based on estimates of Equation (2) in more or less restricted forms, allowing for postcode and Primary or Secondary fixed effects. The main challenge to estimating these two-way fixed effects models (once the necessary data has been assembled) is the large number of school (around 14,000 Primaries and 2,800 Secondaries) and postcode fixed effects (more than 500,000) that need to be estimated or partialled-out, especially when we have a large number of pupil observations. Direct estimation of the complete model using group dummy variables is infeasible on the full data. We therefore follow a step-by-step procedure inspired by a series of papers by Abowd and co-authors (Abowd and Kramarz [1], Abowd Kramarz, and Margolis [2] and Abowd, Creecy, and Kramarz [3]) for firm and individual effects, We describe this method in Appendix Section 11. Implementing this strategy requires very rich data, with information on pupil characteristics, schools attended and their exact location, attainment in at least two periods arising from education in at least two different school settings, and detailed pupil residential address. The next section describes how our data sources allow us to obtain all required information.

4. Data and institutional setting

4.1. Schools, pupils and tests

Compulsory education in England is organised into five stages referred to as Key Stages. In the Primary phase, pupils usually enter school at age 4-5 in the Foundation Stage and then move on to Key Stage 1, spanning ages 5-6 and 6-7. At age 7-8 pupils move to Key Stage 2, sometimes – but not usually – with a change of school. At the end of Key Stage 2, when pupils are 10-11, children leave the Primary phase and go on to Secondary school where they progress through Key Stage 3 to age 14. At the end of each Key Stage, prior to age-16, pupils are assessed on the basis of standard national tests (SATS). At age 16, at the end of the compulsory schooling, pupils sit GCSEs (academic) and/or NVQ (vocational) tests in a range of subjects.

The UK's Department of Education and Skills (DfES) collects a variety of data on school pupils centrally, because the pupil assessment system is used to publish school performance tables and because information on pupil numbers and characteristics is necessary for administrative purposes – in particular to determine funding. A database exists from 1996 holding information on each pupil's assessment record in the Key Stage SATS throughout their school career. For Key Stages 2 and 3 we have information on pupil test scores in Maths, Science and English, while for GCESs/NVQs, we make use of pupil 'Point Scores' in a range of subjects – indicators of total achievement devised by the Qualifications and Curriculum Authority (QCA) and used by the DfES in the performance tables[4]. These point scores are based on allocating points to different grades, and aggregating across types of qualification using appropriate weights (details available from the DfES or the QCA). To make age-16 scores comparable to earlier Key Stage grades and construct measures of educational

---

[4] Note that, although at age 16 pupils can choose to sit exams from a wide range of different subjects, Maths, Science and English are still compulsory.

progress during the Secondary phase, we add pupil test scores across all subjects taken and then assign pupils a level which is their percentile ranking in the distribution across pupils.

Since 2002, the DfES has also carried out a Pupil Level Annual Census (PLASC) which records information on pupil's school, gender, age, ethnicity, language skills, any special educational needs or disabilities, entitlement to free school meals and various other pieces of information including postcode of residence (a postcode is typically 10-12 neighbouring addresses)[5]. PLASC is integrated with the pupil's assessment record (described above) in the National Pupil Database (NPD), giving a large and detailed dataset on pupils along with their test histories. Unfortunately, the length of the time series in the data means that it is not, at present, possible to trace individuals through from their first tests (Key Stage 1) to their final tests (GCSE/NVQ). It is however, possible to follow the academic careers of three cohorts of children through from age-11 to age-16, and to join this information to PLASC data at age-16. We use information on these three cohorts – those aged 16 in 2002, 2003 and 2004 – as the core dataset in this study. Various other data sources can be merged in at school level, including institutional characteristics (from the DfES) and information on the geographical location of each school (down to postcode level). This allows us to geo-code the pupil data based on school attended, and to perform spatial data operations using a Geographical Information System (GIS).

From this large and complex combined data set we are able to construct a balanced panel providing information on three cohorts of over 400,000 pupils each, observed over three academic years, attending more than 14,000 Primary schools (when aged 11) and around 2,800 Secondary schools (when aged 16). We include only those pupils who are in schools that do not admit students on the basis of academic ability and we do not have data on pupils attending private schools[6]. We will use this information to estimate the influence of changes in urban density resulting from

---

[5] Prior to 2002 this information was collected only at school level.

[6] Private schools educate around 6-7% of pupils in England as a whole.

compulsory moves from the Primary to Secondary phase on pupils' subsequent movement through the attainment distribution[7]. In the next section we briefly discuss how school choice and primary-to-secondary school transition operate in England.

4.2. School choice and transition

Since the Education Reform Act of 1988, the 'choice' model of school provision has been progressively extended in the state-school system in England (Glennerster [18]). In this setting, pupils can attend any under-subscribed school regardless of where they live, and parental preference is the deciding factor, All Local Education Authorities (LEAs, now LAs) and schools must organise their admissions arrangements in accordance with the current statutory DfES Admissions Code of Practice. The guiding principle of this document is that parental choice should be the first consideration when ranking applications to schools. However, if the number of applicants exceeds the number of available places, other criteria which are not discriminatory, do not involve selection by ability and can be clearly assessed by parents, can be used to prioritise applicants[8].

How much school choice is there in practice? Gibbons, Machin, and Silva [13] report that, for the Greater London metropolitan area, less than one in four pupils have no Primary school – other than the one they actually attend – within convenient travel distance from home. The figures for Secondary schools are similar (Briggs, Burgess, McConnel, and Slater [4]). Generally, less than 50% of pupils attend their closest school, and even in rural areas the figure is only 60%. This evidence shows that there is not a one-to-one mapping between where a child lives and where he or she goes to school. Hence, and crucial to our analysis, pupils will have a range of Secondary schools from

---

[7] In the work that follows, pupils' attainment is always measured in terms of their percentile ranking within their cohort at each Key Stage, using the distribution within pupils in this balanced panel.

[8] These vary in detail, but preference is usually given first to children with special educational needs, next to children with siblings in the school and to those children who live closest. For Faith schools, regular attendance at local designated churches or other expressions of religious commitment is foremost.

which to choose when making the transition from the Primary to the Secondary phase and urban density can vary quite considerably amongst these schools.

Note however that, despite this flexibility in choice, most Primary to Secondary transitions involve relatively short distance moves. The typical distance between a pupil's Primary and Secondary school is quite low: the median is only 1.67km. The distribution is however skewed, with much higher mean (6.6km) and top decile (above 7km). This distance will play an important role when interpreting the link between pupil achievement and measures of urban density of the school setting. We describe these indices in the next section.

4.3. Indices of urban density

There are obviously innumerable ways to describe the extent to which a school's geographical setting can be characterised in terms of urban density. We pick three that we think capture key aspects relevant to our goals: the density of schools in the locality, the amount of local developed land and the residential population density. The first can be thought of as an 'educational' urban definition and mainly relates to school competition or cooperation, choice and accessibility. To construct this we use GIS to calculate the number of schools within predetermined distances of each school using the matrix of inter-school distances. The second is an environmental definition, and picks out schools in dense metropolitan areas through built environment and infrastructure. To derive this index we turn to a land cover dataset – Landcover Map 2000 – based on Landsat satellite imagery for the late 1990s[9]. Using this data, we compute for each school the proportion of land, within a predetermined radius, that is defined as continuous urban or suburban/semi rural according to the Landcover Map 2000 definitions. The third index identifies schools as urban if they are in places where there are high concentrations of people, and is derived from the 2001 Population

---

[9] This data set records land cover type in 27 categories for 25m square tiles covering the whole of Great Britain. Information is provided by the Centre for Ecology and Hydrology, Huntingdon, Cambridgeshire, England.

Census. Again, we estimate the population density within predefined radii of each school using population counts and land areas of the smallest Census unit (Output Areas). With this data in hand we next proceed to fit the models described in Section 3[10].

5. Results and discussion

5.1. Descriptive statistics

The descriptive statistics on the main variables we will use in the regression analysis are presented in Table 1. As described in Section 4.1, attainment at age 16 and 11 is based on percentiles derived from pupil test/point scores; by construction, descriptive statistics of these attainment percentiles are not very interesting. More informative are the statistics for the change in achievement percentiles between ages 11 and 16 (educational progress). As expected, this is close to zero on average. Moreover, the scores at age 11 and 16 are highly correlated, with a correlation coefficient of 0.69.

Our empirical goal is to see to what extent these movements in the achievement distribution are linked to transition-related changes in the urbanisation of the school environment. The next four rows of Table 1 summarise the urban indices we set out in Section 4.3. We can see that there is a lot of variation in school setting in England: The number of neighbouring schools within 2km varies substantially ranging from 0 to 63, with a mean of 10 and standard deviation of 7.4. The map in Figure 1, which illustrates the number school per square kilometre in and around the Greater London area, suggests that school density is picking out inner city locations in particular. Yet even within London, the number of schools per square kilometre varies widely over short distances. The proportion of neighbouring land that is defined as continuously developed varies from 0 up to almost

---

[10] Formally, these measures are all kernel estimates of school density, proportion of developed land and population density at the school site, using a uniform kernel. We investigated various bandwidths for estimating these indices – 2km, 5km and 10km – but present results only for the 2km radii. Experimentation indicated unambiguously that there is no additional information in the 5km and 10km-based estimates that is relevant for the attainment models we estimate below.

90%, though on average schools are in locations where only 17% of land is urban under this definition. It should be pointed out that this definition mainly captures development in inner city settings. Land that is a closely integrated combination of buildings and open spaces (e.g. gardens) is classified as suburban and rural-developed. Finally, it can be seen that neighbouring population densities vary widely too.

These findings are not wholly unexpected, since we know that England contains a wide mix of urban, suburban and rural schools. Perhaps more surprising is the range of variation in the change in school setting that occurs when pupils move from Primary to Secondary school[11]. The change in the number of local schools varies between a decrease of 58 and an increase of 59, almost the whole range possible. Similarly, pupils experiencing the biggest shift from high-density to low-density environments see an 84 percentage point fall in the proportion of developed land surrounding their school, a change that is mirrored by pupils moving from low-density to high density locations. However, this variation is exceptional, while standard deviations are much more modest – at 4.5 and 11 percentage points for the two variables, respectively. Importantly, most changes seem quite exaggerated considering that pupils typically move to Secondary schools that are fairly close to their Primary schools.

Indeed, this distance is going to be important when it comes to interpreting our results: are the changes in urban density predominantly the result of long-distance movements of pupils between rural and urban locations, or between inner and outer metropolitan locations; or are they the result of small shifts in urban context that occur within cities, towns and other localities? As we can see in the last two rows of Table 1, the typical distance a pupil moves between Primary and Secondary school is low, with a median value of 1.67km. However, the distribution is highly skewed with a mean of

---

[11] Though its average is close to zero, so apparently Secondary schools are not located in predominantly more urban areas. One caveat applies however: the results on urban development and population density might be slightly misleading since Secondary schools usually have much larger sports grounds, and appear to be in lower density and less developed locations than Primary schools.

nearly 6.6km, the top 10% moving over 7.6km and the most mobile 1% of pupils moving over 160km.

5.2. *Estimates of the attainment models*

We move now to regression estimates of the models described in Section 3. Table 2 presents the results from various specifications incorporating the urban indicators described in Section 4.3. Note that the rows report coefficients from three different regressions, one with school density (Row 1) as the urban indicator, one with both types of developed land cover entered together (Rows 3 and 4), and lastly one in which population density enters individually (Row 5).

Column (1) shows the raw Ordinary Least Squares (OLS) association between age-16 attainment percentile and the urban density of the school in which the pupil takes their GCSE/NVQ examinations. In this column there are no controls for pupil or school characteristics. It is immediately clear that attainment is, on average, worse in schools in denser urban areas – regardless of which measure of urban density we use. The coefficient in the first row shows that an additional school within 2km (i.e. an additional 0.08 schools per km^2) is associated with a drop in pupil attainment of just under 0.28 percentiles. This estimate implies that children in schools in the densest urban areas are around 8.4 percentiles below others in their cohort who are in the least dense areas (based on a four standard deviation difference of 30 schools within 2km). Similarly, if we consider land-cover or population density (Rows 2-4), we see that schools in dense or more developed areas have lower performing pupils. A four standard deviation increase in urban density is associated with a six to seven percentile attainment gap. Moving to Column (2), we add controls for basic pupil characteristics – ethnicity, gender, and indicator of entitlement to free school meals (poverty) and an indicator that English is the pupil's first language. This attenuates the coefficients slightly, but the overall picture is unchanged: pupils in dense areas are doing badly.

As discussed earlier, we conjecture that this association could be the result of sorting of more disadvantaged children into urban areas. It is obvious and well known that the poor live,

predominantly, in cities. Moreover, the demographics of school attendance are driven to a large extent by local housing costs (especially social housing) and community characteristics, which are in turn determined by factors other than expected school quality only. The value-added attainment model in Column (3) shows that this conjecture is, by and large, correct. This specification differences within pupils between Primary (age-11) and Secondary (age-16) schools, and removes any (unobservable) factors that are fixed for individuals over time or for the same pupil in the two different schools. These factors include innate abilities and family background, and, for the vast majority of pupils who do not move house between the two schooling phases, neighbourhood influences from residential location too[12]. The results are striking: pupils experiencing a relative increase in the density of their school location when they change from Primary school to Secondary school move up the attainment distribution relative to other pupils in their cohort. The expected attainment gap for someone experiencing a two standard deviation increase in urban density relative to another experiencing two standard deviation decrease is around 2.5 – 2.8 percentiles. This is a substantial difference, and is comparable to the gap for poor pupils who are eligible for free school meals, and only slightly less than the gender gap (4 percentiles), both of which have been the subject of much academic attention. Even so, the gains are not large enough to close the gap in the level of attainments of pupils in the most urban schools relative to students in more rural areas at the end of Secondary compulsory education (which we have shown remains as large as 8.5 percentiles). Note finally, that the suburban/rural developed land cover is not significant in the value-added models – it is the influence of central urban location (measured by continuous urban land cover) that is most closely related to attainment.

In terms of magnitude, there is not much to choose between the various urban indicators employed in these value-added specifications (except suburban/rural developed, which is always

---

[12] Based on information in our data, we can calculate that only 10% of pupils change their residential address between the end of Primary school (age 11) and the end of the Secondary phase (age 16).

non-significant), although interestingly the coefficient on our measure of school density is always more statistically significant than the others. In the results in the next section, we will try to disentangle the relative influence of these indices as best we can.

Note finally that most schools in England have been established for a long time and their location is driven predominantly by long-standing housing and population patterns, so it is very unlikely that school density is endogenous to pupil attainment through school location decisions. This is in contrast to the standard empirical models of agglomeration economies in production, where firms and employees choose where to locate, and hence employee and firm density is clearly endogenous to the productivity of those locations. Still, there are many potential explanations for a correlation link between density and attainment other than a causal one, and we delve further into this in the following sections.

5.3. *Estimates of one and two-way fixed effects models*

Table 3 presents results of more fully specified models of age-11 to age-16 pupil attainment gains, including school and, in some cases, residential fixed effects. Since it does not seem theoretically or statistically relevant (from Table 2), we drop the suburban land-cover indicator, but now include all the other urban indicators together in the regression. In addition to the controls described in Table 2, we include the mean age-11 attainment of the age-16 Secondary school pupils (unless the model has Secondary school fixed effects) and a dummy variable indicating if the pupil switched in or out of the standard secular Community school sector, for example into or out of a Faith state school[13].

---

[13] About 30% of Primary school pupils in England and 15% of Secondary school students are educated in religiously affiliated state-schools. These schools are funded by the central government, via Local Education Authorities, and have to comply with the national curriculum in terms of teaching and assessment of their students. Including these additional controls may be important because there are differences between the spatial distribution of Faith and secular schools and we want to rule out the possibility that the association of attainment with urbanisation is driven by covert selection of higher ability pupils into religious schools (West [46], Gibbons and Silva [14]). Moreover, other work on similar data

Looking at Column (1), with Primary school fixed effects, school density wins out convincingly in the 'fight' between our highly-correlated urban indices. The coefficient on population density is near zero and completely insignificant. The coefficient on urban land cover is sizeable, but also insignificant. Undoubtedly, there are many things that are correlated with school density that we have not taken into account here, and we make no claim at this stage that it is actually school density that has a causal impact on attainment. Yet our findings suggest that aspects of the urban environment associated with high school density – such as accessibility and choice, competition or cooperation – are conducive to higher attainment, a conjecture to which we will return later in Section 5.6. Column (2) repeats this analysis with Secondary school fixed effects, which makes little difference to interpretation of the school density coefficient, though we have stronger evidence of a more general association with urbanisation reflected in land cover. This finding might suggest that urban density in Primary schools has a larger effect on earlier age-11 attainments, than urbanization in Secondary schools has on age-16 achievements. On the other hand, we find a small but weak negative association between educational progress and population density.

Appendix Section 11 outlines how we can estimate urban density influences whilst controlling for Primary or Secondary influences *and* residential postcode fixed effects, and discusses methods of implementation. This strategy controls for trends in educational achievements shared by individuals within the same Primary or Secondary schools, which may arise from school institutional features or unobservable school preferences shared by schoolmates and their parents. The strategy also controls for unobservable characteristics common to families sorting into the same residential neighbourhoods, and other local unobserved amenities which may affect pupil value added.

Columns (3) and (4) of Table 3 present the results of this exercise with home postcode *and* either Primary school or Secondary school fixed effects. Even with these very stringent

---

has shown that mean peer-group attainment at the start of Secondary school has a small but significant impact on pupil progress through Secondary school (Gibbons and Telhaj [15]).

specifications, we find a robust link between urban density and pupil attainment. The impact of school density is almost unaffected in both size and significance relative to the earlier results, while proportion of urban land and population density still do not play a significant role. Overall, it seems that the effect we are capturing in terms of density influences on attainment is not related to place of residence (including residential neighbourhood effects and unobserved family background influences), but more genuinely to a change in school setting.

Given the results so far, we conclude that a relative increase in urban density on transition between Primary and Secondary school is linked to a small but significant improvement in expected pupil attainment. The results and discussion in the next sections try to shed some light on the scope and sources of these gains from urban density.

5.4. Geographical scope: who benefits?

In this section we uncover possible sources of the density attainment advantage by considering which groups of pupils gain and which groups lose after the transition between Primary and Secondary school. Table 4 shows the impact of urban density on attainment for various sub-groups of the population. All models are age-11 to age-16 value-added models, with the same controls as in Table 3, Column (1). Results for this baseline specification are repeated in the first column of the new table for comparison.

A first question to ask is whether the density effect is driven predominantly by pupils moving from high-density (e.g. urban) locations or by those moving from low-density (e.g. rural) locations. We consider this by splitting the sample into pupils originating in Primary schools in the bottom (Column (1)) and top (Column (2)) density quartile. Looking at the coefficients in the table we see that the magnitude of the relationship between school density and attainment is similar for both groups, though the point estimate is larger and more significant for pupils starting out in high-density Primary school locations. This results is perhaps surprising: the gains arise not just from pupils moving from low-density areas, but amongst pupils *already* schooled in high-density locations who

either: a) Lose out through moving to lower density Secondary school locations; or b) Gain by moving into higher density locations within the top quartile[14].

Apparently, what we are seeing is not easily explained by simple agglomeration or urbanisation stories in which rural or semi-rural pupils gain from moving into urban locations for their Secondary schooling. We come to similar conclusions when we consider the distance between the Primary and Secondary schools that pupils attend, as in Columns (3) and (4). The first of these columns reports the results for the sub-group of pupils whose Primary-Secondary distance is in the lowest quartile (<900m) and the second for the upper quartile (>3.6km). Surprisingly, there are positive attainment impacts from school density for both groups, though as we would expect these are much larger and much more statistically significant for the long-distance movers. Even so, many of the moves in this upper quartile are still quite short and it is doubtful that many of the density changes in our data are really the result of major transitions between rural and urban locations; they mainly capture more marginal changes amongst lower and higher density places in urban settings.

This pattern is revealed even more explicitly when we look at the distinction between pupils moving schools within London and those moving within the rest of England in Columns (7) and (8). We see here that school density is strongly linked to attainment even for pupils moving within the metropolitan London region, moves that are clearly not in any way rural-urban. Finally, the same argument carries over to Columns (7) and (8) where we split the sample by pupils making the primary-secondary transition within Local Education Authorities (by far the majority) and those moving to secondary schools in a different LEA. Although the magnitude of the link between school density and achievement is bigger for pupils making between LEA-moves, the difference between

---

[14] Coming at this from a different angle, we also split the sample into two other groups: those who make a school transition that results in a top-quartile change (rise) in school location density, and those whose transition results in a bottom-quartile change (fall) in density. Again, we found that, although both groups seem to experience relative gains from density increases (or relative losses if density decreases), the effect is strongest amongst those whose Secondary school is in a less dense location than their Primary school.

these groups is not enough to suggest broad agglomeration forces and local labour markets as the driving factors.

5.5. Robustness checks: strategic behaviour, pre-transition trends, mean reversion, greater variance in cities?

One possible explanation for what we have found is that pupil attainment is measured inappropriately. Our measures at age 11 and at age 16 may not really be directly comparable, because pupil percentile score at age 11 is based on tests in Maths, Science and English only, whilst at age 16 it is based on a wide range of different tests in different subjects (though Maths, Science and English are compulsory). The exact mix of subjects taken by pupils at age 16 will depend to some extent on school expertise and strategic behaviour. Schools tend to be evaluated on the basis of the proportion of students passing at least 5 GCSE/NVQ grade C exams (graded between G and A*), so there are incentives for schools to act strategically (encouraging pupils to qualify in many exams at moderate grade or take 'softer' options) to maximise the number of students reaching this target. Clearly there would be a problem in terms of the interpretation of our results if they were driven by a link between urbanisation and the incentives for schools (or pupils) to act deliberately in this way. We go some way to allaying these fears in Columns (1) and (2) of Table 5, which present results for alternative measures of attainment[15]. In the first column, we look at pupil percentile based on *mean* grade across all GCSE/NVQ subjects, rather than on *total* points added up across subjects. In Column 2 we instead use the age-16 percentile based on total points obtained in Maths, English and Science only, rather than added up across all subjects. The results here are almost identical to before, so it does not seem that strategic behaviour in terms of 'numbers' versus 'grades', or 'soft' versus 'tough' exams can explain our findings.

---

[15] To keep things simple we just consider school density as the urban index here, where we observe most of the action.

Another argument against a causal interpretation of our urban density impact is that we cannot properly account for pre-existing trends in pupil performance. Perhaps, what is happening is that pupils on more rapidly rising attainment trajectories are also those who experience the biggest change in urban density, through their choice of Primary and Secondary schooling. In fact, our previous estimates imply that the pupils showing the fastest progress must both: a) Prefer lower density Primary schools, conditional on choice of Secondary school and place of residence; and b) Prefer higher density Secondary schools, conditional on choice of Primary school and residential neighbourhood. There is no obvious way to reconcile these patterns with any consistent preferences that might be related to pupil attainment trends. However, we would clearly like to rule this selection process out as a candidate explanation for our findings.

As explained above, we do not have information on pupil attainment trends prior to the Primary-Secondary school transition for these cohorts. We can however, provide a weaker test based on the deviation of the age-11 to age-16 attainment gain from a linear trend, using an intermediate measure of attainment at age 14. This robustness check is presented in Column (3), where the dependent variable is the difference between age-16 and age-14 percentile, minus the difference between the age-14 and age-11 percentile – i.e. the acceleration in attainment. The coefficient in Column (3) suggests that attainment does rise at a faster rate in more densely located schools, so our main results are not solely due to sorting of pupils with heterogeneous linear attainment trends.

As one further step, we devise an Instrumental Variable (IV) strategy based on differences in travel costs at different home locations, which change behaviour over school choice in ways that are not directly related to pupils' expected growth in achievement. Our argument is that individuals having easy access to public transport in metropolitan areas are likely to travel longer distances to Secondary school and experience the biggest changes in urban density on transition from Primary to Secondary school. So, our strategy is to predict the change in urban density experienced by pupils on transition from Primary to Secondary school using proximity between pupil homes and underground

or railway stations within a metropolitan area. To implement this approach, we confine our attention to pupils living in London and attending Secondary schools in London when they are aged 16. Although movements in either direction seem possible – i.e., either towards the centre or the periphery of the metropolitan area – we argue that living near a station in London makes it more likely that a pupil travels towards the centre of town; this is because it is only in this direction that the rail network brings pupils anywhere close to Secondary schools[16]. So, we construct the following instruments: two measures for (straight line) distance between individuals' residential postcode and the closest train and underground stations, within 2km of home and two dummies indicating whether there are no railway or underground stations within 2km of residential location[17].

Results from using these instruments for the change in school density are reported in Columns (4) and (5) of Table 5. Column (4) includes the set of controls detailed before, while in Column (5) we add Postcode District fixed effects to control for broad geographical differences[18]. First stage statistics are reported at the bottom of the table. For both the specifications, these test statistics show that our instruments are powerful predictors of changes in urban density and not significantly correlated with unobservable characteristics which may affect educational progress. The second-stage IV estimates confirm that pupils tend to progress faster if they move to more densely located schools: the coefficient on the change in the number of schools within 2km is larger than before[19],

---

[16] This is simply a feature of the geographical density of schools and stations, which both increase towards the centre of London. This means that it is much more likely that a Secondary school can be reached by walking from a station in inner London, than by walking from a station in the outskirts. This intuition is borne out by our results.

[17] Critical to this instrumental variable approach is the assumption that residential choice in relation to transport access is not linked to unobserved individual and family characteristics or other local amenities that may affect pupil educational performance. In our defence we should emphasize that our value-added models already deal with pupil unobservable attributes (including family background and preferences) that are fixed over time.

[18] Full UK postcodes are typically of the form AB# #CD, where # is numeric. Deleting the last three characters generates a Postcode District code. There are more than 400 Postcode Districts in the extended London region.

[19] Although it is comparable to the estimates for the number of schools only (Column 5, Table 2) and for the London area only (Column 8, Table 4).

though its statistical significance is reduced. This finding lends further credence to the idea that a relative increase in urban density on transition between Primary and Secondary school is linked to small but significant improvement in pupil attainment.

Mean reversion is another concern. As we have noted, pupils enrolling in urban schools tend to be those with lowest attainments and from the most disadvantaged backgrounds. Pupils in urban Primary schools may then experience the fastest value-added simply through reversion towards mean attainment at later stages in their education. Could this give rise to a spurious correlation between density and attainment in our value-added models? Since we consider how *changes* in attainment are correlated with *changes* in school density, our positive coefficients could only arise through mean reversion if pupils experiencing the biggest increase in density when they move school are also those starting from the lowest base in terms of their initial conditions, which would require that urban pupils experience the biggest density increases. A priori, this does not seem likely because if mean-reversion is a feature of our value-added measure then it is just as likely to be a feature of the school density index. So, a pupil with a very low test score in a very high density Primary is most likely to end up next time with a somewhat higher test score in a lower density Secondary, leading to a negative correlation between the change in performance and change in density[20]. Nonetheless, we investigate this issue directly by repeating our empirical analysis separately for high and low age-11 achievers in high density Primary schools, and high and low age-11 achievers in low density Primary schools. Appendix Table A1 reports the findings from this exercise. For all four groups we find that an increase in school density is associated to higher value-added during Secondary school, which makes it unlikely that mean reversion in pupil attainment accounts for our main findings.

Finally, we consider whether our regressions – based on what happens to pupils on average – mask heterogeneity in school quality that belies our general point that urban schools are not systematically failing. Perhaps there really are many 'very bad' schools in dense places, but these are

---

[20] This intuition is borne out by our results. See Appendix Table A1.

more than compensated by many 'very good' schools in similar areas. This argument implies that the variance in school quality is higher amongst schools in dense urban settings, a point which we can easily test by regressing the square of our regression residuals (from the specification in Column (1) of Table 3) against the urban indicators and other regressors (i.e. an *à la* White's test for heteroscedasticty). The results of this exercise suggest that there is no statistically significant link between urban density and the variance in attainment; the F-test on our three urban indicators yields an *F(3,2816)* statistic of 0.18, with a p-value of 0.9085. In other words, higher urban density seems to be associated with a rightwards-shift throughout the distribution of school quality[21].

5.6. Looking for an explanation: school choice and competition, resources or policy?

In the discussion so far, we have shown that pupils progress faster between ages 11 and 16 when they move school from low to high density locations, and that they progress more slowly when the change is in the other direction. The results in Section 5.4 also seemed to rule out explanations based on better functioning urban teacher labour markets, deeper pools of competent teachers in urban settings, and other broad urbanisation, or agglomeration-based explanations. The distances between Primary and Secondary schools are simply not large enough for these explanations to make much sense; moreover we can detect a density impact on attainment amongst pupils moving short distances, within Local Education Authorities, and within metropolitan areas. We have also ruled out strategic course choice selection and selection of pupils with stronger attainment trends into schools in denser settings. Rather, we have emphasised that school density seems to be the driving factor behind our findings. In what follows, we try to unpick what aspects of school density matter and why, and to consider the role of local educational policies.

Although we argued that the inter-school distances over which we detect an effect are not large enough for these schools to be operating in distinct teacher labour markets, they can be operating in

---

[21] A better test would use quantile regressions, but this proved infeasible given our model specifications and sample size.

zones facing very different markets in terms of pupils and competing or cooperating schools. Indeed, transport costs dictate that actual enrolment patterns for schools are geographically based and quite localised, so two schools just a few kilometres apart in an urban environment may face very different sets of potential pupils and different schools with which they are effectively competing. Similarly, two such schools may be linked to entirely different networks of other schools with which they cooperate and share knowledge. Such networking and cooperation in professional development is a heavily promoted aspect of current Government educational policy, e.g. through the 'Beacon' schools initiative (described in Footnote 3) and so called Networked Learning Communities (NCSL [40]). The underlying story may thus be one in which school density – and the competition or networking opportunities that it engenders – is the driver behind the density-attainment link.

We provide some evidence that is supportive of this interpretation in the first column of Table 6, where we split up the school density index into two components: the number of Primary schools within 2km and the number of Secondary schools. What is apparent here is that all the impact of the change in density on the change in attainment between ages 11 and 16 arises through the number of neighbouring Secondary schools, not the number of local Primary schools. This is exactly what we would expect if it is choice, competition or cooperation among co-located schools that matters for our results, because, when we include Primary school fixed effects, it is only variation in the density of schools near the Secondary school of destination that identifies the coefficient on school density. If school markets matter, only the density of Secondary schools should be relevant at the Secondary phase. Primary schools do not provide competition, and are unlikely to be closely networked with Secondary schools. This result is robust to the inclusion of postcode and Primary school fixed effects in the two-way fashion detailed before (not reported in Table 6).

Pupils in denser locations may also perform better because schools in these locations are better resourced and are part of Government initiatives to encourage collaboration and boost performance. Indeed, a recent raft of Government policies has been targeted specifically at 'failing' inner city

schools in disadvantaged areas in an effort to raise attainments and improve other school-related outcomes. Others have found evidence of benefits arising from these policies (Machin, McNally, and Meghir [32]). Perhaps then, what we observe is the benefit to pupils attending schools in locations subject to these policy initiatives? In Columns (2) and (3) of Table 6, we answer this question by adding some controls for basic school resources to our attainment models. These are average expenditure per pupil within the Local Education Authority that funds schools, the number of qualified teachers in the school, plus indicators of whether or not the Secondary school comes under the umbrella of various city-related educational policy programmes; these initiatives are: 'Excellence in Cities' (EiC), which puts additional money into areas facing disadvantages, and was rolled out over three phases in 1999, 2000 and 2001 followed by a 'Clusters' scheme focussed on group of schools from 2001 onwards; and 'Education Action Zones', which are groups of schools in disadvantaged areas which receive extra funding, face tougher targets and receive other assistance. Additionally, we control for the number of pupils enrolled at the school, mainly because school funding in England is linked to pupil numbers. Looking at Table 6, the signs of the coefficients on the resources variables in the attainment models are as expected and the impact of additional teachers is statistically significant, but the EiC and Education Action Zone dummies are insignificant and unsystematically signed (not reported). The specifications in Columns (2) and (3) are otherwise comparable to those in the first column. Moving right across these three, it can be seen that the impact of Secondary school density is unchanged after controlling for resource differences. Clearly, differences in resources matter, but are not responsible for the difference in pupil attainment between dense and less dense school locations.

Finally, in Column (4), we investigate whether co-operation among near-by schools and networks of local professionals lie behind our correlations. To do so, we include in our specification variables counting the number of 'Beacon' schools (separately for Primaries and Secondaries) within 2km of the school. We also include a term counting the number of Independent (private) schools

within a 2km range from the school. This variable should shed some light on whether the competitive threat posed by a large number of local private institutions raises educational standards in the public sector (Epple and Romano [11] and Hoxby [24]). It is still the number of Secondary schools that has the strongest association with pupil educational progress, though we also find some link between local Beacon schools and pupil achievement. However, this link is mainly through the number of Primary Beacons, which and a causal impact from networking with Beacon Primaries seems unlikely. Finally, the number of Independent schools does not enter our specification significantly. At least in this context, private school competition does not raise state-school average performance. The link between the change in urban density and pupil educational progress during Secondary school seems primarily related the competition forces that are in place among closely co-located state Secondary schools, or cooperation of a different sort from that set up through the Beacon scheme.

6. Conclusions

Although pupil attainment in dense urban places is low on average, this is not because urban environments disadvantage pupils but because disadvantaged pupils with low average attainments attend the most urbanised schools. Our results show that comparable pupils progress better in schools located in denser urban settings, where we measure urban density in terms of density of schools, continuous urban land cover and population density. Interestingly, amongst these three factors, school density generally dominates as an explanatory factor.

Our results additionally show that progress during Secondary school is linked explicitly to the density of local Secondary schools and conditionally unrelated to Primary school density. We argue this is indicative of localised educational effects related to inter-school competition or cooperation. We have also shown that the density impact on attainment can be measured amongst pupils moving between schools in quite close proximity, within Local Education Authorities and within urban areas. For this reason, it seems unlikely that these density economies can be attributed to teacher

labour markets or other sources of urbanisation and agglomeration advantages cited in the context of firm productivity.

From our estimates, we can derive the elasticity of pupil attainment with respect to school density. This is low – less than 0.02 – but it still implies that pupils educated in the most dense environments could gain around 2-3 percentiles in the national pupil attainment distribution relative to others in their cohort educated in the least dense settings. The urban educational advantage is quite substantial in relation to the total contribution of schools to pupil attainment, considering that pupils in the top 1-in-10 schools in terms of attainment are, on average, only 20 percentiles above pupils in the bottom 1-in-10 schools. In economic terms, a rough assessment based on the returns to additional years of education would put the value of this attainment gap at up to 2% on average earnings[22]. Even so, the gains are not large enough to close the attainment gap between pupils educated in urban schools and students in more rural areas at the end of the Secondary compulsory phase. This gap is due to city schools enrolling children with serious background disadvantages rather than urban school ineffectiveness, so sensible urban educational policy should aim to identify and tackle the disadvantages that urban children carry with them to school.

Interestingly, but perhaps just coincidentally, the elasticity of pupil achievement to school density is at the lower end of the range of estimates on the effects of urban size and density on firm productivity, in the 0.03 to 0.08 range[23].

---

[22] Estimates of the returns to an additional year of education in Britain are typically 7-10%. The point system used to measure pupil progress in our data allows us to deduce that 1 percentile in the pupil distribution of attainment is roughly equivalent to 0.067 of a year's educational progress during compulsory schooling. Assuming equal labour market returns at the intensive and extensive margins, this means that 1 percentile advantage is worth around 0.67% on earnings.

[23] See Rosenthal and Strange [45] for an international survey, and Rice and Venables [43] and Graham [19] for British evidence.

7. References

[1] J. Abowd, F. Kramarz, Econometric analyses of linked employer-employee data, Labour Economics 6 (1999) 53-74.

[2] J. Abowd, F. Kramarz, D. Margolis, High wage workers and high wage firms, Econometrica 67 (1999) 251-333.

[3] J. Abowd, R. Creecy, F. Kramarz, Computing person and firm effects using linked longitudinal employer employee data, mimeo INSEE, 2002.

[4] A. Briggs, S. Burgess, B. McConnel, H. Slater, School choice in England: Background facts, CMPO working paper, 06/159, 2006

[5] A. Chevalier, P. Dolton, S. McIntosh, Recruiting and retaining teachers in the UK: An analysis of graduate occupation choice from the 1960s to 1990s, CEE discussion paper 21, 2002.

[6] C. Clotfelter, E. Glennie, H. Ladd, J. Vidgor, Would Higher Salaries Keep Teachers in High-Poverty Schools? Evidence From a Policy Intervention in North Carolina, NBER working paper 12285, 2006.

[7] J. Cullen, B. Jacob, S. Levitt, The Effect of School Choice On Student Outcomes: Evidence From Randomized Lotteries, NBER working paper 10113, 2003.

[8] J. Cullen, B. Jacob, S. Levitt, The Impact of School Choice on Student Outcomes: An Analysis of the Chicago Public Schools, Journal of Public Economics 89 (2005) 729-60.

[9] P. Dolton., W. van der Klaauw, Leaving teaching in UK: A duration analysis, Economic Journal 105 (1995) 431-444.

[10] C. Emmerson, C. Frayne, S. McNally, O. Silva, Aimhigher: Excellence Challenge: A Policy Evaluation using the Labour Force Survey, mimeo CEP-LSE, 2006.

[11] D. Epple, R. Romano, Competition Between Public and Private Schools: Vouchers and Peer Effects, American Economic Review 88 (1998) 33-62.

[12] GHK, A Review of Beacon innovation activities, report for Department for Education and Skills, Birmingham, 2005.

[13] S. Gibbons, S. Machin, O. Silva, Choice, Competition and Pupil Achievement, IZA discussion paper 2214, 2006.

[14] S. Gibbons, O. Silva, Faith Schools: Better Schools or Better Pupils?, CEE discussion paper 72, 2006.

[15] S. Gibbons, S. Telhaj, Peer effects and pupil attainment: Evidence from Secondary school transition, CEE discussion paper 63, 2006.

[16] E. Glaeser, Are cities dying?, Journal of Economic Perspectives 12 (1998) 139-160.

[17] E. Glaeser, M. Kahn, J. Rappaport, Why do the poor live in cities?, NBER working paper 7636, 2000.

[18] Glennerster, H. (1991), "Quasi-Markets for Education", *The Economic Journal*, 101, 1268-76.

[19] D. Graham, Wider economic benefits of transport improvements: link between agglomeration and productivity, stage 2 report, Department of Transport, London, 2005.

[20] E. Hanushek, The Failure of Input-Based Schooling Policies, Economic Journal 113 (2004) 64-98.

[21] E. Hanushek, S.G. Rivkin, How to improve the supply of high quality teachers, in: D. Ravitch (Ed.), Brooking Papers on Education Policy, Brookings Institution Press, Washington, DC, 2004, pp. 7-25.

[22] E. Hanushek, J.F. Kain, S.G. Rivkin, Why public schools lose teachers, Journal of Human Resources 39 (2004) 326-354.

[23] E. Hanushek, J.F Kain, D.M. O'Brien, S.G. Rivkin, The market for teacher quality, NBER working paper 11154, 2005.

[24] C. Hoxby, Do Private Schools Provide Competition for Public Schools?, NBER working paper 4978, 1994.

[25] C. Hoxby, Does Competition Among Public Schools Benefit Students and Taxpayers?, American Economic Review 90 (2000) 1209-38.

[26] C. Hoxby, School Choice and School Productivity (Or, Could School Choice be a Rising Tide that Lifts All Boats?), in: C. Hoxby (Ed.), The Economics of School Choice, University of Chicago Press, Chicago, IL, 2003.

[27] C. Hoxby, School Choice and School Competition: Evidence from the United States, Swedish Economic Policy

Review 10 (2004) 11-67.

[28] A. Krueger, Economic Considerations and Class Size, Economic Journal 113 (2003) 1-34.

[29] A. Jaffe, M. Trajtenberg, R. Henderson, Geographical Localization of Knowledge Spillovers as Evidenced from Patent Citations, Quarterly Journal of Economics 108 (1993) 577-598.

[30] V. Lavy, Performance pay and teachers' effort, productivity and grading ethics, mimeo Hebrew University of Jerusalem, 2005.

[31] R. Lupton, Schools in disadvantaged areas: Recognizing context and raising quality, CASE discussion paper 76, 2004.

[32] S. Machin, S. McNally, C. Meghir, Resource and standards in urban schools, mimeo CEP-LSE, 2005.

[33] S. Machin, A. Vignoles, What's the good of education? The economics of education in the United Kingdom, Princeton University Press, Princeton, 2005.

[34] A. Marshall, Principles of Economics, MacMillan, London, 1890.

[35] A. Marshall, Principles of Economics, eighth edition, MacMillan, London, 1920.

[36] R. Murnane, Improving urban public schools: suggestions for teacher union leaders, conference proceedings, Federal Reserve Bank of Boston (2002) 297-303.

[37] R. Murnane, R. Olsen, The effects of salaries and opportunity costs on length of stay of teaching: Evidence from Michigan, Review of Economics and Statistics 71 (1989) 347-352.

[38] R. Murnane, R. Olsen, The effects of salaries and opportunity costs on length of stay of teaching: Evidence from North Carolina, Journal of Human Resources 25 (1990) 106-124.

[39] D. Neal, The effects of Catholic Secondary schooling on educational achievement, Journal of Labor Economics, 15 (1997) 98-123.

[40] NCSL, What makes a network a learning network?, National College for School Leadership and Department of Education and Skills, http://www.ncsl.org.uk/networked/networked-introduction.cfm accessed June 2006.

[41] Ofsted (Office for Standards in Education), Improving city schools, London, 2000.

[42] Ofsted (Office for Standards in Education) Access and achievement in urban education, London, 2003.

[43] P. Rice, T. Venables, Spatial determinants of productivity: Analysis of the regions of Great Britain, CEP discussion paper 642, 2004.

[44] S. Rosenthal, W. Strange, The determinants of agglomeration, Journal of Urban Economics 50 (2001) 191-229.

[45] S. Rosenthal, W. Strange, Evidence on the nature and sources of agglomeration economies, in: V. Henderson, J.F. Thisse (Eds.), Handbook of Regional and Urban Economics, Vol. 4, Elsevier, Amsterdam, 2004, pp. 2219-2171.

[46] A. West, School choice: The limits of quasi regulation, Consumer Policy Review, 15 (2005) 94-98.

8. Figures

Figure 1: Density of schools, illustrated for the London area

Notes: Figure shows Kernel density estimates for total school density in the Greater London region and environs. In our estimates we use information for the whole of England.

9. Tables

Table 1: Descriptive statistics for age-16 sample

Variable	Mean	Std.Dev.
Percentile attainment at age 16	49.939	28.842
Change in percentile age 11-16	-0.029	22.852
Number of schools within 2km	10.089	7.445
Proportion continuous urban within 2km	0.172	0.180
Proportion suburban/rural developed within 2km	0.300	0.139
Population density within 2km (per hectare)	26.554	19.328
Change in number of schools within 2km	0.014	4.590
Change in proportion continuous urban within 2km	-0.014	0.108
Change in proportion suburban/rural developed within 2km	0.013	0.109
Change in population density within 2km	-0.711	11.977
Number of Primary schools within 2km	7.836	6.136
Number of Secondary schools within 2km	2.343	1.593
Change in number of Primary schools within 2km	-0.322	3.850
Change in number of Secondary schools within 2km	0.352	1.316
Distance between Primary and Secondary school (km)	Mean = 6.594	s.d. = 28.984
	Median = 1.676	i.q.r. = 2.750

Note: Sample size 1,202,970 based on pupils with non-missing data from age 11 to age 16 in non-selective Secondary schools, excluding special needs schools. Population density is based on output areas of the 2001 Census and is in persons per hectare. Change in pupil percentile is non-zero because percentiles are calculated on sample of 1,202,566 pupils, with some missing data; pupil attainment percentile is based on total GCSE/NVQ points at age 16 and test scores in Maths, Science and English tests at age 11. Number of primary schools: approximately 14500. Number if secondary schools: approximately 2800. Number of LEAs: 147.

Table 2: Urbanisation and pupil attainment percentile age-16; various urban indicators

	Separate regressions for each urban indicator groups		
	(1)	(2)	(3)
	Level	Level	Age-11 to 16 Diff.
School density:			
Number of schools	-0.279	-0.210	0.139
within 2km	(0.029)	(0.029)	(0.017)
Proportion developed within 2km:			
Continuous urban	-9.069	-7.044	6.322
	(1.158)	(1.176)	(0.785)
Suburban/rural developed	-10.857	-9.668	0.494
	(1.466)	(1.336)	(0.728)
Population:			
Population density	-0.093	-0.071	0.050
within 2km	(0.012)	(0.011)	(0.007)
Pupil characteristics	No	Yes	Yes
Primary school effects	No	No	No
Secondary school effects	No	No	No

Note: Table reports regression coefficients from pupil level regressions. Standard errors in parentheses (clustered on Secondary school in Columns 1 to 4; clustered on Local Education Authority in Column 5). Underline significant at 1% or better. Dependent variable is pupil's percentile attainment at age-16, or percentile gain from age 11 to age 16. Pupil controls are ethnicity (8 categories), entitlement for free school meals, English as additional language, male. Models include year dummies. Level of urbanisation is proportion of 2km radius around school that is classified as continuous urban in 2000 (Landcover map 2000). Population density is persons per hectare based on population and land areas of 2001 Census Output Areas for which centroid is within 2km of school. Sample size approximately 1.2million. Number of primary schools: approximately 14500. Number if secondary schools: approximately 2800. Number of LEAs: 147.

Table 3: Urbanisation and percentile attainment gains; alternative one and two-way fixed-effects strategies

	(1)	(2)	(3)	(4)
	Primary fixed effects only	Secondary fixed effects only	Residential postcode and Primary	Residential postcode and Secondary
Number of schools within 2km	*0.087* (0.037)	<u>0.100</u> (0.028)	<u>0.083</u> (0.032)	<u>0.098</u> (0.025)
Proportion urban within 2km	1.103 (1.440)	<u>4.530</u> (0.795)	0.846 (1.289)	<u>3.056</u> (0.809)
Population density within 2km	0.004 (0.015)	-0.015 (0.010)	0.008 (0.013)	-0.016 (0.009)

Note: Table reports regression coefficients from pupil level regressions. Standard errors in parentheses (clustered on Secondary schools in Columns 1 and 3; clustered on LEAs in Columns 2 and 4). Underline significant at 1% or better. Italics significant at 5% or better. Dependent variable is pupil's percentile attainment gain from age 11 to age 16. Controls are pupil ethnicity (8 categories), pupil entitlement to free meals, English as additional language, male dummy, year dummies, Community school dummy, Secondary school average age-11 attainment (Columns 1 and 3 only). Proportion continuous urban is proportion of 2km radius around school that is classified as continuous urban in 2000 (Landcover map 2000). Population density is persons per hectare based on population and land areas of 2001 Census Output Areas for which centroid is within 2km of school. Sample size approximately 1.2million (about 900,000 in Columns 3 and 4, where a minimum of two observations per postcode is needed). Number of primary schools: approximately 14500 in one way fixed effects models and 13200 in two way fixed effects models. Number if secondary schools: approximately 2800. Number of LEAs: 147.

Table 4: Urbanisation and percentile attainment gains; geographical scope

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
-	Whole sample	Primary school low-density	Primary school high-density	Short distance move	Long distance move	Move not within London	Move within London	Move within LEA	Move outside LEA
Number of schools within 2km	*0.087* (0.037)	<u>0.075</u> (0.029)	<u>0.099</u> (0.029)	0.062 (0.035)	<u>0.088</u> (0.023)	<u>0.073</u> (0.019)	<u>0.092</u> (0.036)	<u>0.073</u> (0.019)	<u>0.110</u> (0.029)
Proportion urban within 2km	1.103 (1.440)	-1.183 (1.300)	1.076 (0.936)	0.288 (1.529)	0.732 (0.885)	*1.672* (0.723)	-2.945 (1.982)	*1.472* (0.712)	-0.294 (1.193)
Population density (per hectare)	0.004 (0.015)	-0.011 (0.010)	0.018 (0.011)	0.004 (0.011)	0.003 (0.009)	-0.004 (0.007)	<u>0.045</u> (0.016)	0.003 (0.007)	0.005 (0.011)
Sample	1201894	327949	274365	314728	302363	1049819	135130	1082839	119055

Note: Table reports regression coefficients from pupil level regressions. Standard errors in parentheses (clustered on Secondary school). Underline significant at 1% or better (Italics significant at 5%). Dependent variable is pupil's percentile attainment gain from age 11 to age 16. Controls are pupil ethnicity (8 categories), entitlement to free school meals, English as additional language, male dummy, Community school dummy, Secondary school average age-11 attainment, year dummies, Primary (age-11) school fixed effects. Proportion continuous urban is proportion of 2km radius around school that is classified as continuous urban in 2000 (Landcover map 2000). Population density is persons per hectare based on population and land areas of 2001 Census Output Areas for which centroid is within 2km of school. Low and high density refers to bottom (≤ 5) and top (≥ 13) quartile of schools within 2km. Long and short distance refers to bottom (<0.9km) and top (>3.6km) quartile of distance between Primary and Secondary school attended.

Table 5: Urbanisation and percentile attainment gains; robustness of results

	(1)	(2)	(3)	(4)	(5)
	Percentiles based on mean age-16 points	Percentiles based on English, Maths and Science only	Acceleration in attainment, ages 11-14-16	Instrumental variable estimates	Instrumental variable estimates
Number of schools within 2km	0 .082 (0.017)	0.057 (0.016)	0.067 (0.024)	0.205‡ (0.108) [0.122]	0.181 (0.113) [0.123]
First Stage Statistics					
F-Statistics	-	-	-	210.40 (0.0000)	198.41 (0.0000)
Hansen Statistics (p-value)	-	-	-	3.515 (0.319)	2.740 (0.4335)
Pupil characteristics	Yes	Yes	Yes	Yes	Yes
Primary school effects	Yes	Yes	Yes	Yes	Yes
Secondary school average peer achievement	Yes	Yes	Yes	Yes	Yes
Postcode Area effects	No	No	No	No	Yes

Note: Table reports regression coefficients from pupil level regressions. Standard errors in parentheses (clustered on Secondary school, expect Columns 5 and 6). Underline significant at 1% or better (‡ significant at 6%). Dependent variable is pupil's percentile gains from age 11. Controls are pupil ethnicity (8 categories), entitlement to free school meals, English as additional language, male dummy, Community school dummy, Secondary school average age-11 attainment, year dummies. Sample size approximately 1.2million (Columns 1 to 4). Instrumental variable in Columns 5 and 6 for London metropolitan area only; sample size about 140,000 observations. Instruments are: residential postcode distance from closest tube station and to closest rail station within 2km, plus dummies for no rail station and no train station within 2km of residential postcode. Column 6 additionally controls for postcode area fixed effects. First stage statistics from models with standard errors clustered at the postcode level (p-values in round brackets); second stage standard errors clustered at postcode level in round brackets and at the Secondary school level in square brackets.

Table 6: Urbanisation and percentile attainment gains; resources, policy interventions and school competition

	(1)	(2)	(3)	(4)
	All schools	All schools	All schools	All schools
Number of state Secondary schools within 2km	0.363 (0.101)	0.363 (0.101)	0.374 (0.102)	0.332 (0.109)
Number of state Primary schools within 2km	0.016 (0.044)	0.005 (0.044)	0.005 (0.045)	-0.026 (0.046)
Number of qualified teachers (FTE)	-	0.065 (0.017)	0.066 (0.017)	0.064 (0.017)
Number of pupils x100 (FTE)	-	-0.321 (0.104)	-0.324 (0.104)	-0.321 (0.103)
Expenditure per pupil in Local Education Authority (£00s)	-	*0.080* (0.034)	*0.093* (0.041)	0.074 (0.034)
Number of Beacon Secondary schools	-	-	-	0.039 (0.306)
Number of Beacon Primary schools	-	-	-	0.419 (0.160)
Number of Independent Secondary schools	-	-	-	-0.106 (0.083)
Number of Independent Primary schools	-	-	-	0.037 (0.120)
Urban land cover and population density	Yes	Yes	Yes	Yes
Urban school policies	No	No	Yes	No
Primary school fixed effects	Yes	Yes	Yes	Yes

Note: Table reports regression coefficients from pupil level regressions. Standard errors in parentheses (clustered on Secondary school). Underline significant at 1% or better (Italics significant at 5%). Dependent variable is pupil's percentile attainment gain from age 11 to age 16. Controls are pupil ethnicity (8 categories), entitlement to free school meals, English as additional language, male dummy, Community school dummy, Secondary school average age-11 attainment, year dummies, Primary (age-11) school fixed effects. FTE means full time equivalent. Column 4 also controls for whether current Secondary is a Beacon. Sample size approximately 1.2million.

10. Appendix: Additional Tables

TableA1: Urbanisation and percentile attainment gains; mean reversion

	(1)	(2)	(3)	(4)
	High Achievers, High Density	High Achievers, Low Density	Low Achievers, High Density	Low Achievers, Low Density
Number of schools within 2km	0.174 (0.016)	0.091 (0.019)	0.109 (0.0167)	0.054 (0.020)
Average density change	*-1.508*	*1.596*	*-1.429*	*1.465*
Pupil characteristics	Yes	Yes	Yes	Yes
Primary school effects	Yes	Yes	Yes	Yes
Secondary school peer achievement	Yes	Yes	Yes	Yes
Sample	287972	322916	321791	269215

Note: Table reports regression coefficients from pupil level regressions. Standard errors in parentheses (clustered on Secondary school). Underline significant at 1% or better (Italics significant at 5%). Dependent variable is pupil's percentile attainment gain from age 11 to age 16. Controls are pupil ethnicity (8 categories), entitlement to free school meals, English as additional language, male dummy, Community school dummy, Secondary school average age-11 attainment, year dummies, Primary (age-11) school fixed effects. High and low achievers refer to bottom (<50) and top (>=50) half of the KS2 percentile distribution. High and low density refers to bottom (<9) and top (>=9) half of school density distribution.

11. Appendix: Estimation of two-way fixed effects

The main challenge to estimating the two-way fixed effects models described by equation (3) in Section 3 is the large number of school and postcode fixed effects that need to be estimated or absorbed, especially when we have a large number of pupil observations.

In the data for England (described above) we have over 14,000 Primary schools, over 2,800 Secondary schools, around 500,000 postcodes, and data on over 1.2 million pupils. The simplest approach to estimating the full model would be to include either the 2,800+ Secondary school or the 14,000+ Primary school dummies as regressors, and de-mean the data using a within-groups transformation to eliminate home postcode fixed effects. Since this is infeasible on the full dataset of pupils, we follow a procedure inspired by a series of papers by Abowd and co-authors (Abowd and Kramarz [1], Abowd Kramarz, and Margolis [2], and Abowd, Creecy, and Kramarz [3]) for firm and individual effects. Our approach involves, as a first stage, using this estimator on sub-groups (A) of the full data set in order to estimate the postcode fixed effects for each sub-group. When this is complete, we use the full sample to re-estimate the model, using the either within-Primary or within-Secondary school transformation on a model that includes the postcode fixed effects estimated from the first stage as a regressor.

More formally, consider the case of Primary and residential postcode fixed effects; we estimate:

$$\Delta y_{i,jk} - \Delta \bar{y}^l_{i,jk} = \beta(\Delta u_{jk} - \Delta \bar{u}^l_{i,jk}) + (d'_{sk} - \bar{d}'^l_{sk})\delta + \upsilon_{i,jk}, i \in A \qquad (4)$$
$$\hat{n}_l = \Delta \bar{y}^l_{i,jk} - \hat{\beta}\Delta \bar{u}^l_{jk} + \bar{d}'^l_{sk}\hat{\delta}$$

for all subgroups A (defined below); then, on the full data set:

$$\Delta y_{i,jk} - \Delta \bar{y}^j_{i,jk} = \beta(\Delta u_{jk} - \Delta \bar{u}^j_{i,jk}) + \theta(\hat{n}_l - \bar{\hat{n}}^j_l) + \omega_{i,jk} \qquad (5)$$

Ideally, each subgroup A would be defined such that that it includes all pupils living in a specific set of postcodes R, who all attend Primary schools within a set P, to which no other pupils (other than those in R) are admitted. All the relevant postcode and Primary fixed effects could then be estimated consistently for each of these 'autarchic' systems using only data from pupils in each subgroup A.

In practice, school admissions systems are not so cleanly defined, but the process of identifying suitable groups is aided by the fact that admission in England is organised geographically by Local Education Authorities (LEAs). In most cases, pupils attending a school within a given LEA come from neighbourhoods within that LEA, or from a closely neighbouring LEA. So we form sub-groups A on an LEA-by-LEA basis and proceed using the following steps:

(a) For a given LEA L, draw all pupils who live in the set of postcodes R in that LEA, plus all pupils in a set Primary schools P which intersects the set R through pupil Primary school choice, i.e. pupils in a set of postcodes Q outside the LEA L, going to the same set of Primary schools P as pupils in R.

(b) Estimate (4) on this sub-sample of pupils (sub-group A), using dummy variables for the Primary schools P, allowing for postcode fixed effects for the postcode set R plus the set Q of postcodes that intersect with P through pupil Primary school choice, but are outside the LEA.

(c) Retrieve estimates for the postcode fixed effects for the set R, i.e. $\hat{n}_l$ in (4).

(d) Repeated for each LEA until all postcode fixed effects have been estimated.

(e) Estimate (5) using the full England sample and the estimated postcode effects from step (c).

Note that this is an approximate method, because at step (b) we have not used data on all pupils in the set of postcodes Q that are outside the LEA in order to compute the fixed effects Q. However, these postcode fixed effects will be estimated consistently if pupils in the set $Q \cap P$ are a random

sample of all pupils living in postcode Q. Moreover, we do not use the estimates of effects Q in the last stage (e), but re-estimate them when we reach the LEA in which they are located during step (d). So, the only concern is that not estimating Q correctly (i.e. on the full set of pupils living in those postcodes) may bias our estimates of the Primary school effects P, which in turn could affect our estimates of postcode effects R. We checked the sensitivity of our results to these sorts of problems by re-defining our group A in various ways, for example by excluding Q completely and using only the sample of pupils in R. A similar method can be used to control for residential postcode and Secondary school fixed effects.