OXFORD BULLETIN OF ECONOMICS AND STATISTICS, 0305–9049
doi: 10.1111/obes.12076

Using Compulsory Mobility to Identify School Quality and Peer Effects*

Francis Kramarz† and Stephen Machin‡ and Amine Ouazad§

†*CREST, CEPR, and IZA Boulevard Gabriel Péri, 92240 Malakoff, Cedex, France (e-mail: kramarz@ensae.fr)*
‡*University College London, Centre for Economic Performance, London School of Economics, and CEPR, Gower Street, London WC1E 6BT, UK (e-mail: s.machin@ucl.ac.uk)*
§*INSEAD, Boulevardde Constance, 77300 Fontainebleau, France (e-mail: amine.ouazad@insead.edu)*

Abstract

Education production functions that feature school and student fixed effects are identified using students' school mobility. However, student mobility is driven by factors like parents' labour market shocks and divorce. Movers experience large achievement drops, are more often minority and free meal students, and sort endogenously into peer groups and school types. We exploit an English institutional feature whereby some students must change schools between grades 2 and 3. We find no evidence of endogenous sorting of such compulsory movers across peer groups or school types. Non-compulsory movers bias school quality estimates downward by as much as 20% of a SD.

I. Introduction

Policymakers, parents and researchers alike have shown considerable interest in the measurement of school quality – that is, the unbiased measurement of a school's causal effect on student achievement.[1] Causal school quality estimates can help in the design of accountability systems,[2] funding mechanisms and also help parents choose schools.[3]

The main challenge faced in study of this question is to disentangle what part of schools' average test scores is due to the quality of the school from what is due to either time-varying

*We thank audiences of the COST meeting, the Paris School of Economics, CREST, the 2006 IZA prize in labor economics conference, Boston College, Cornell University, the US National Academy of Sciences and Sciences Po Paris. John Abowd, Sandra Black, Eric Maurin, Thomas Piketty and Jean-Marc Robin offered helpful suggestions for improving the manuscript. The authors acknowledge financial and computing support from INSEAD, the Centre for Economic Performance and CREST-INSEE.
JEL Classification numbers: I21, J00

[1]The literature dates back at least to Coleman (1966), Summers and Wolfe (1977), and Card and Krueger (1992). A recent contribution is Dustmann, Puhani and Schönberg (2012).

[2]Under the US No Child Left Behind Act, states are required to publish adequate yearly progress (AYP) data for schools by racial and gender group as well as for 'special-needs' groups. This is a rudimentary way of controlling for student observables. A failure to make AYP can lead to school closure, so it is implicitly assumed is that AYP measures partly reflect school quality.

[3]The English Department for Education publishes league tables with students' average progress by school that take account of observable contextual factors, so called contextual value added. This may be not free of potential time-varying confounders such as family events (Gibbons, 2007) and may not necessarily reflect school quality.

or non-time-varying student characteristics. Controlling for observable non-time-varying student characteristics is straightforward in a least-squares regression, and when such characteristics are unobservable, control is usually via the inclusion of student fixed effects.

It is noteworthy that identifying education production functions that feature school and student fixed effects relies on student mobility across schools. This is clearly illustrated by the inability of a single cross section of data to jointly estimate student and school fixed effects,[4] and by the impossibility – absent the mobility of students across schools in longitudinal data – of separately identifying student and school fixed effects.

Hence student mobility across schools is the main source of identification in regressions involving school quality and student fixed effects. Therefore, estimates of school effects can be biased if identification relies on the selected subsample of pupils who move because of family events that affect both their educational achievement and their mobility. In this paper, we consider a specific education production function, but this point applies also to a range of different education production functions that control for student and school effects.[5]

As it turns out, there is evidence that student mobility across schools may be correlated with other events that can affect test scores. Research – including Gibbons (2007) in the United Kingdom and Burkam, Lee and Dwyer (2009) in the United States – indicates that mobility is systematically associated with family events such as unemployment, family break-ups and labour market opportunities. There is also evidence that these family events have a negative impact on test scores (James-Burdumy, 2005; Stevens and Schaller, 2011). This is likely to cause bias in the estimation of school effects: if, for instance, students who experience negative family events move to worse (resp., better) schools, then the difference between good and bad schools is likely to be overestimated (resp., underestimated). Hence there will be an endogenous mobility bias if one cannot control for these time-varying shocks. Regrettably, large-scale administrative data with student test scores does not usually include such time-varying observables, which have the potential to confound the estimates of school quality.[6]

These data limitations offer a significant and difficult research challenge to researchers, and a warning to both parents and policymakers. To try and address this challenge in the context of this paper (i.e. primary school students in England), we note that children who start school at age 5 in an infant-only school must change schools between grades 2 and 3; these students are called compulsory movers. Other students start school at age 5 in an

[4] Unless one assumes that the effects are uncorrelated, in which case student effects are random and the model can be estimated as a hierarchical random effects model. Unfortunately, regression estimates suggest that the correlation between student fixed effects and school effects is significantly non-zero. Also, a random effects model does not allow for the correlation between observable covariates and the effects.

[5] Education production functions either regress student test scores (in levels) on school quality measures, student effects, peer effects and past test scores (Todd and Wolpin, 2003) or regress student progress on school quality measures, peer effects, student effects and other controls (Rivkin, Hanushek and Kain, 2005). Random-effects estimation of education production functions does not rely on student mobility because such estimates can rely on a cross section. However, random-effects estimation depends on orthogonality conditions, which are rejected in English education data (and most likely in other settings as well).

[6] This is the case not only in England but also the case in data from North Carolina (Rothstein, 2010), New York (Rockoff, 2004) and a number of other school districts in the United States. In Europe, Danish and Swedish data include a wealth of information on parents (unemployment, health, family status), but lack test scores before age 16.

infant and junior school, so can stay in the same school for grades 1 through 6; when these students change school they are called non-compulsory movers.

We show that non-compulsory movers are unlikely to be a credibly exogenous source of identification for the measurement of school quality. First, non-compulsory movers tend to have markedly different characteristics than the average student: lower achievement and a greater likelihood of being a free meal, minority or special-needs student. Second, they experience large achievement drops between grades 2 and 6. Third, they tend to move to schools that are far away from the previous school, which suggests that family events may be the main driving force behind their mobility. Fourth, non-compulsory movers sort endogenously into school types and peer groups.

In contrast, there is little evidence that compulsory movers have different characteristics than the average student or that there is sorting between infant-only and infant and junior schools within each neighbourhood. Compulsory movers also move to schools that could allow school mobility without residential mobility. There is little evidence that compulsory movers tend to sort endogenously into school types or peer groups. Finally, between grades 3 and 5, the mobility of compulsory movers is lower than that of non-compulsory movers.[7]

In this paper, we therefore estimate school quality and peer effects, while controlling for student effects, by focusing solely on the observations of compulsory movers – that is, on the set of students who start their education in an infant-only school. It is interesting that some school effects can be estimated either by focusing on the sample of compulsory movers or by including all students in the estimation; this is possible for grade 6 schools that welcome both compulsory and non-compulsory movers. We show that it is possible to estimate the endogenous mobility bias on school effects: schools that welcome a large number of non-compulsory movers exhibit school effects that are downward biased in previous studies that do not focus on compulsory movers. If all students who move to a given school are non-compulsory movers, then the school quality estimate (controlling for student effects, cohort dummies and peer composition) is downward biased by 4% a SD of test scores on average. This bias is substantial: it amounts to about 20% of the SD in school quality.

Our approach also allows us for the estimation of peer effects. Previous research achieves such identification in large administrative data sets by looking at year-to-year changes in grade composition (Hoxby, 2000).[8] In this paper we show that non-compulsory movers tend to sort endogenously into peer groups across grades 2 and 6. Changes in peer groups will therefore be correlated with students' time-varying events and will confound the estimates (Manski, 2000). Focusing on compulsory movers mitigates the endogenous mobility bias in the measurement of peer effects. Also, when we focus on compulsory movers, actual variations in grade composition are close to random demographic shocks.

[7] When the impact of mobility on test scores is homogeneous across students, our education production functions also provide unbiased estimates of school quality.

[8] Using these year-to-year random demographic shocks is a versatile technique since it can be readily applied to a large number of data sets. There are several other identification strategies, including the use of randomized assignments of students to schools (Hoxby and Weingarth, 2007), random changes in class size (Graham, 2008) and various instruments for peers' achievement and characteristics (Sacerdote, 2001; Gould, Lavy and Paserman, 2004; Sacerdote, 2010).

The rest of the paper proceeds as follows. Section II presents the endogenous mobility bias in education production functions, and indicates that non-compulsory movers are not likely to be a credible source of identification. Section III provides evidence that focusing on compulsory movers may reduce the endogenous mobility bias. Section IV presents the results of our estimation of school quality and peer effects and then estimates the magnitude of the endogenous mobility bias both for school quality estimates and for peer effects. Section V concludes by summing up results and formulating policy recommendations.

II. Endogenous mobility bias

Time-varying factors that can confound school quality estimates

To first make the point that endogenous mobility is a potential identification concern, we start with a simple education production function. The full education production function is presented in section IV and will feature peer effects, school quality, the effect of past inputs on current achievement. The more basic education production function of this section decomposes test scores into a school effect and a student effect, where the latter captures student-specific unobserved factors that determine test scores and may be correlated with school quality:

$$\text{Test Score}_{i,f,t} = \text{School}_{J(i,t)} + \text{Student}_i + \varepsilon_{i,f,t} \tag{1}$$

Here $\text{Test Score}_{i,f,t}$ is the test score of student i in field f in year t, and School_j is the school effect of school j ; $J(i,t)$ is the school of student i in year t, which captures the impact of school j on test scores. Finally, Student_i is the student effect of student i and $\varepsilon_{i,f,t}$ denotes the time-varying unobservables. This specification does not capture the impact of previous scores and/or previous school inputs and does not control for peer effects. However, it does allow us to illustrate the use of student mobility in the estimation of school effects. More complete specifications are introduced later in the paper.

School effects and student effects are estimated without bias if the residual $\varepsilon_{i,f,t}$ is uncorrelated with student mobility across schools. Indeed, take the first difference of specification (1), which shows that the progress of student i in field f between two assessment periods is equal to the difference between the quality of school of the second period and the quality of the school of the first assessment period:

$$\text{Test Score}_{i,f,2} - \text{Test Score}_{i,f,1} = \text{School}_{J(i,2)} - \text{School}_{J(i,1)} + \varepsilon_{i,f,2} - \varepsilon_{i,f,1}$$

Endogeneity biases appear if, for instance, students who experience parental divorce or unemployment (i) experience a drop in test scores and (ii) are more likely to move to either better or worse schools – for instance, if $\varepsilon_{i,f,t} = -\delta\text{Parental Shock}_{i,f,t} + \eta_{i,f,t}$. In this expression, δ is the size of the downward shock to student achievement, $\text{Parental Shock}_{i,f,t} = 1$ if the student experiences a parental shock, and $\eta_{i,f,t}$ is a residual. If students who experience a drop in test scores are more likely to move to better schools, then the correlation between the parental shock and the difference in school effects is negative and the difference between good and bad schools is underestimated; conversely, if students who experience a drop in test scores are more likely to move to worse schools, then the difference between good and bad schools is overestimated.

The impact of endogenous mobility on estimates of school quality can be expressed by a simple equation. We start by writing specification (1) in matrix form as follows.

$$\mathbf{TS} = D \cdot \mathbf{Student} + F \cdot \mathbf{School} + \varepsilon$$

where **TS** is the vector of test scores (of size 2NT, with two test scores per student per year, N students, T time periods and a balanced panel[9]), **Student** is the vector of student effects (of size N), **Student** is the vector of school effects (of size J), D is the design matrix of student effects, and F is the design matrix of school effects. If the residual ε includes a parental shock that affects test scores and that is correlated with mobility, then $\varepsilon = -\delta \cdot \mathbf{Parental\ Shock} + \eta$ where is **Parental Shock** the stacked vector of 0–1 indicator variables $\mathbf{Parental\ Shock}_{i,f,t}$ for $i = 1, 2, \ldots, N$ and $t = 1, 2, \ldots, T$. The vector of estimates of school effects will then be biased such that

$$\widehat{\mathbf{School}} = \mathbf{School} - \delta(F' M_{\mathrm{D}} F)^{-1} F' M_{\mathrm{D}} \cdot \mathbf{ParentalShock}$$

$\widehat{\mathbf{School}}$ is the vector of estimates of school quality from specification (1), and **School** is the vector of the 'true' values of school quality. The matrix M_{D} is defined as $Id - D(D'D)^{-1}D'$. Provided that students who experience a parental shock move to different schools, mobility is correlated with school effects; in that case, $F' M_{\mathrm{D}} \cdot \mathbf{ParentalShock}$ is non-zero,[10] and all estimates of school effects are biased. Such mobility bias affects not only the estimates of the effects for the pair of schools between which students move endogenously but also all other estimates of school quality.

In the following sections, we present the English education system and its interesting institutional features that will allow us to substantially alleviate the endogenous mobility bias.

Schools and national assessments in England

The educational system in England currently combines market mechanisms (many of which were introduced in the Education Act of 1988) with a centralized assessment operating through a national curriculum (Machin and Vignoles, 2005). National exams for all students are taken (and reported) at the end of four Key Stages throughout the years of compulsory schooling: in primary school, Key Stage 1 (from ages 5 to 7 in grades 1 and 2) and Key Stage 2 (from ages 7 to 11 in grades 3–6); and afterwards in secondary schools Key Stage 3 and 4. At the end of each Key Stage, pupils are assessed in the core disciplines: mathematics, English and science (though no science test is given for Key Stage 1). These tests are nationally set and are anonymously marked by external graders.[11] This paper's data set is the administrative data set that collects these test scores, the National Pupil Database.

Two primary school systems coexist in England. In some areas, schools cater only to Key Stage 1 students (infant schools) or to Key Stage 2 students (junior schools). In other

[9] Similar results obtain with an unbalanced panel data set.

[10] Each element of the vector $F' M_{\mathrm{D}} \cdot$ Parental Shock is the average time-varying shock for each school. Each element is 0 if, on average, each school has an equal number of students who experience an upward shock (parental divorce in the first period) and students who experience a downward shock (parental divorce in the next period).

[11] While Key Stage 1 tests were externally marked during the period we are considering, they are not externally marked anymore, as described in the Key Stage 1 assessment and reporting arrangements.

areas, schools cater to both types of students (so-called infant and junior schools). Students who start school in an infant and junior school but change schools between grades 2 and 6 are called non-compulsory movers, and students who start school in an infant- only school are called compulsory movers. The Data S1 suggests that junior-only schools are distinct geographic and administrative entities managed by a different headteachers in a majority of cases.

The primary schooling system in England is characterized by a variety of different organizational structures, funding sources, admissions procedures, and teacher contracts. These differences are likely to generate different educational outcomes for pupils. We group schools into three subsets: (i) schools that have some say in admissions and staff recruitment-voluntary aided schools and foundation schools – other schools' teacher recruitment and dismissal is mostly determined by the local education authority – community schools, voluntary controlled and other schools;[12] (ii) faith schools, who are more able to manage their intake and differ in some of their curriculum; (iii) foundation schools, whose board majority is controlled by a foundation. These three groups are not mutually exclusive: for instance, voluntary-aided schools employ and dismiss staff, and are mostly faith schools.

Correlation between mobility and achievement

Because a number of references suggest that time-varying shocks to students are correlated with mobility, we suspect that student time-varying factors are a confounding factor in the estimation of school quality. In this section we show that (i) non-compulsory movers are not representative of the overall student population, (ii) movers experience large drops in achievement, and (iii) that this is mostly accounted for by non-compulsory movers.

The National Pupil Database provides rich information on pupils' characteristics and location: gender, free school meal status, special educational needs, ethnicity group, school identifier and school postcode. Pupils who receive free meals are from families in the lowest quintile of income (17%). The ethnicity variables are described in Data S1. Three cohorts of pupils are followed in grades 1 and 2 (Key Stage 1) and in grades 3, 4, 5 and 6 (Key Stage 2): the 1998–2002 cohort, the 1999–2003 cohort and the 2000–04 cohort. The sample consists of 1,705,300 students and 21,360 schools.

Table 1 lists the demographic features of the overall student population in column (1) and those of non-compulsory movers in column (2).[13] It is remarkable that the test scores of non-compulsory movers are about 26% of a SD lower than the average. The table's third row shows that although the fraction of free meal students is 17% in the overall population, it is 25% in the subset of non-compulsory movers. The fraction of special-needs students is about 4 percentage points higher among non-compulsory movers. The parents of non-compulsory movers are also much less likely to speak English at home, are less likely to be White, and are more often Black or Asian. Of course, these demographic characteristics are non-time-varying and are captured by the student effect. However, the existence of sharp

[12] Code of Practice on LEA Schools Relationships (Department for Education and Skills, 2011).
[13] Table 1 of Data S1 gives the number of students in infant-only and infant-and-junior schools.

TABLE 1

Comparison of the characteristics of compulsory and non-compulsory movers

		Movers:	
	All Students *(1)*	*Non-compulsory* *(2)*	*Compulsory* *(3)*
Test score (all fields, standardized)	49.90	47.34	49.70
	(9.96)	(11.01)	(9.77)
Male	0.51	0.51	0.51
	(0.50)	(0.50)	(0.50)
Free school meal	0.17	0.25	0.16
	(0.38)	(0.43)	(0.37)
Special needs	0.23	0.27	0.21
	(0.42)	(0.44)	(0.41)
English spoken at home	0.91	0.84	0.93
	(0.28)	(0.37)	(0.26)
White	0.85	0.76	0.87
	(0.36)	(0.43)	(0.34)
Black	0.03	0.08	0.02
	(0.18)	(0.27)	(0.15)
Asian	0.06	0.07	0.05
	(0.23)	(0.26)	(0.23)
Observations	1,705,300	204,119	546,769

Notes: Test score: pooled grade 2 and grade 6 scores in English and mathematics. Standard deviations in parentheses. The number of observations is the number of students in each subsample.

demographic differences between compulsory and non-compulsory movers suggests that these movers may also differ in their time-varying characteristics.

To check for that possibility, we estimate the relative change in test scores for stayers, movers and non-compulsory movers. This is what is estimated by the regression of student test scores on an interaction between a mobility dummy and the grade 6 dummy. Formally,

$$\begin{aligned}\text{Test Score}_{i,f,t} = {} & \text{Grade6}_t + \delta \text{Mobility}_i \times \text{Grade6}_t + \text{School}_{j(i,t)} \\ & + \text{Controls}_{i,f,t}\delta + \gamma \text{Mobility}_i + \varepsilon_{i,f,t}\end{aligned} \tag{2}$$

Here $\text{Test Score}_{i,f,t}$ is the test score of student i in field f (English or Math) in assessment grade t. The term Grade6_t is a dummy set equal to 1 when t = grade 6, which captures the difference in the average test score in grade 6 and in grade 2. $\text{School}_{j(i,t)}$ is a school effect; and $\text{Controls}_{i,f,t}$ is a set of controls for field and cohort. The coefficient of interest is δ, the (non-causal) impact of mobility on test scores in grade 6 conditional on school effects, controls and grade 6 effects. Although δ does not have a causal interpretation, a strongly negative effect is a measure of the correlation between the magnitude of downward shocks on test scores and mobility.[14]

[14] Also, a school (and the teachers therein) that welcomes both grades 2 and 6 students may possibly have an incentive to understate grade 2 test performance to show maximum value added in grade 6. With different schools in grades 2 and 6 that may not be the case. Both schools may have an incentive to maximize children's performance. This might additionally bias any school quality effect observed that is based on infant/junior school combinations.

TABLE 2

Test scores of movers

	(1) Test Score	(2) Test Score
Mobility in Grade 6	−2.269*	
	(0.020)	
Non-compulsory mobility		−2.243*
		(0.019)
Observations	6,901,337	6,901,337
R^2	0.146	0.146
Year and english dummies	Yes	Yes
Mobility dummies	Yes	Yes
F	19,647	19,284

Note: Robust standard errors in parentheses. $^*P < 0.01$.

The results presented in Table 2 suggest that movers experience declines in test scores, of about 18% (see column 1; also see Table 3 for the achievement of non-compulsory movers relative to their school peers). Column 2 reports the correlation between test scores and mobility for non-compulsory movers, replacing the dummy Mobility$_{i,f,t}$ with the dummy Non-compulsory Mover$_{i,f,t}$. The effect observed in column 1 is mostly accounted for by non-compulsory movers. Indeed, the drop in achievement for non-compulsory movers is 29.8% of a SD – a full 1.6 times the original coefficient – even though non-compulsory movers account for 27% of movers. This finding suggests substantial endogeneity bias (due to the presence of endogenous mobility when we estimate the education production function, i.e. specification (1) and, later, specification (3) in section IV).

III. Compulsory movers

In England, infant-only schools serve children aged from 5 to 7, and junior-only schools from 7 to 11. Other children attend combined all-through primary schools that cover the infant and junior years. Thus, children who start their curriculum in an infant-only school cannot remain in the same school after grade 3. Repeating a grade seldom occurs in England, so there is a nearly perfect mapping of ages and grades. Children who must change schools between grades 2 and 3 are called compulsory movers

In principle, compulsory movers can provide us with a more credibly exogenous source of mobility than non-compulsory movers. However, at least five important conditions must be satisfied: (i) compulsory movers should be more representative of the overall population of students than are non-compulsory movers; (ii) within each geographic area, parents should not sort endogenously across infant-only vs. infant and junior schools; (iii) compulsory movers should move exogenously across schools and peer groups between grades 2 and 6; (iv) moves should be genuine moves (i.e. not purely administrative name changes), but should not be parental moves indicative of changes in parents' family structure or employment status; (v) compulsory movers change schools between grades 2 and 3, but should not sort themselves endogenously again between grades 3 and 4, grades 4 and 5 and grades 5 and 6.

TABLE 3

Within-school achievement of non-compulsory movers

	Dependent variable:			
	Grade 2 percentile		*Grade 2 test score*	
	English	*Maths*	*English*	*Maths*
	(1)	*(1)*	*(3)*	*(4)*
Non-compulsory Mover	−0.035** (0.001)	−0.035** (0.001)	−1.459** (0.025)	−1.989** (0.026)
Observations	1,697,835	1,697,835	1,697,835	1,705,300
R^2	0.004	0.004	0.153	0.148
Year and field dummies	Yes	Yes	Yes	Yes
School effect	No	No	Yes	Yes
F-statistic	2,381	2,381	2,362	7,943

Notes: Robust standard errors in parentheses. $^{**}P < 0.01$, $^{*}P < 0.05$, $^{+}P < 0.1$.

Compulsory movers' demographics and test scores

Column (3) of Table 1 shows that compulsory movers' characteristics are significantly closer to those of the overall student population than non-compulsory movers – even though compulsory movers make up only 32% of the student population. The difference between the average test score in the student population and the average test score of compulsory movers is only 2% of a SD. The fraction of boys is virtually identical, which indicates that boys-only or girls-only schools are not more likely to serve compulsory movers. The fraction of free meals is about 1 percentage point lower among compulsory movers. The subset of compulsory movers contains slightly fewer special-needs students, more students who speak English at home and slightly fewer Blacks and Asians. In all cases however, the difference in the demographics of compulsory movers and the overall student population is smaller than the difference in the demographics of non-compulsory movers and that population. This is good news for two reasons. First, it suggests that, nationally, there is no systematic national sorting of students across infant-only and infant and junior schools. Second, if schools are differentially effective, and if we are interested in the average effectiveness of schools, a representative sample of movers is useful in identifying the average school quality.

Parental choice of infant-only schools vs. infant and junior schools

Compulsory movers are present in most Local Education Authorities. On average, about 30% of grade 2 students in an LEA are in infant-only schools. So one potential concern when focusing on compulsory movers is that, in grade 1, parents might anticipate that their child will have to change schools between grades 2 and 3 and thus sort endogenously across schools as soon as grade 1. Such sorting could present problems for our claim if infant-only schools are more (or less) likely to affect test scores.

TABLE 4

Sorting between schools with compulsory mobility and schools without compulsory mobility

	(1) Infant-only school	*(2) Infant-only school*
Maths Grade 2 score	0.002**	
	(0.001)	
Male	0.000	0.002
	(0.009)	(0.009)
Black	−0.013	−0.015
	(0.035)	(0.035)
Asian	0.026	0.021
	(0.028)	(0.028)
Mixed	−0.028	−0.029
	(0.038)	(0.038)
Special needs	0.010	−0.008
	(0.012)	(0.011)
Free school meal	−0.004	−0.010
	(0.013)	(0.013)
Observations	15,067	15,067
R^2	0.663	0.662
Area effect	yes	yes
Number of areas	6819	6819
Schools per area	3.56	3.56
F-statistic for area effects	594.0	594.48
P-value	0.000	0.000
F-statistic for Regression	16.50	10.73
P-value	0.000	0.000

Notes: Area is defined by the outward postcode and the postal district number. For instance, a school with postcode 'TR14 0DW' has area code 'TR14 0'. Standard errors are clustered by postal district area. Robust standard errors in parentheses. Clustered by area. $^{**}P < 0.01$, $^{*}P < 0.05$, $^{+}P < 0.1$.

To assess the magnitude of student sorting within each geographic area, we construct geographic areas using postal sectors.[15] There are 10,631 such areas in England and each area contains, on average, three to four schools. The following regression evaluates the amount of sorting within each area. We regress the probability of being enrolled in an infant-only school on student characteristics and an area effect, estimated by taking the within-area difference of the regression (Table 4). If there is no sorting across schools, we should expect the vector of coefficients on student characteristics to be non-significant. Results show that there is little evidence of sorting by achievement, gender or special-needs status whether or not we control for achievement. Although the coefficient for achievement is statistically significant, the difference is only 0.01% of a SD. Overall these results indicate

[15]The postal sector is the postcode's first part plus the initial digit of the second part.

TABLE 5

Distance of mobility

	Median	*SD*	*Observations*
Compulsory move			
Distance	1.33	36.29	368,114
Distance for Q1 of grade 2 scores	1.36	35.30	85,840
Distance for Q2 of grade 2 scores	1.33	34.19	86,494
Distance for Q3 of grade 2 scores	1.31	33.95	93,214
Distance for Q4 of grade 2 scores	1.29	33.70	102,005
Non-compulsory move			
Distance	20.73	86.63	149,978
Distance for Q1 of grade 2 scores	18.39	77.69	54,210
Distance for Q2 of grade 2 scores	20.35	83. 65	36,568
Distance for Q3 of grade 2 scores	21.94	88.09	33,295
Distance for Q4 of grade 2 scores	24.53	96.57	25,344

that there is no significant sorting on the basis of observable characteristics of students across infant-only vs. infant and junior schools.

Compulsory movers' mobility

We computed the mobility rate at the end of each grade based on the Pupil Level Annual Census, a register of student demographics in between the assessment years.[16] Between grades 3 and 5, the mobility rate of compulsory movers is either comparable to or smaller than the rate of non-compulsory movers. In particular non-compulsory movers tend to sort themselves into schools between grades 5 and 6 at a substantially higher rate. Compulsory movers must change schools after completion of grade 2. Both compulsory and non-compulsory movers must do so after grade 6.

Table 5 presents statistics on the distance (in miles) between the grade 2 and the grade 6 school. That distance is calculated by taking the postcodes of the two school and then measuring the distance between their respective centroids. There are more than 1.7 million postcodes in the UK, so they are very precise measures of location.[17] The postcode corresponds to the school's geographic location and not to its mailing address.

The median distance of a non-compulsory move is 20.73 miles, far enough to suggest that parental or family events – rather than student test scores and/or education – are the main drivers of mobility. Of the students who move, 6.9% change their LEA. It is interesting that the distance moved is correlated with prior test scores: students with higher grade 2 scores tend to move further from their grade 2 school.

For compulsory movers, the picture is quite different. The median distance is 1.3 miles, which suggests that moves are genuine moves (i.e. not purely administrative school name changes), and that they are not as strongly correlated with family events. With moves of relatively short distance, a student's school location could change without a corresponding residential location change. For compulsory movers, the dispersion of distances is also

[16] Figure presented in Data S1.

[17] Broadly speaking, postcodes correspond to 'blocks' in the US Census.

TABLE 6

Dynamic sorting across school types and peer groups

	(1)	*(2)*	*(3)*	*(4)*
	Foundation	*LEA control*	*Faith*	*Employer/admission*
Non-compulsory mover	−0.024**	−0.012**	0.022***	−0.023**
	(0.003)	(0.003)	(0.003)	(0.003)
Observations	1,684,976	1,684,976	1,684,976	1,684,976
R^2	0.675	0.663	0.596	0.631
F	21,428	21,479	14,073	8,072
Compulsory mover	0.003	−0.001	0.000	0.026**
	(0.004)	(0.004)	(0.006)	(0.005)
Observations	1,684,976	1,684,976	1,684,976	1,684,976
R^2	0.674	0.663	0.596	0.632
F	21,083	21,394	14,846	7,607

	(1)	(2)	(3)	(4)
	Δ Male peers	Δ Black peers	Δ Asian peers	Δ Free meal peers
Non-compulsory mover	0.003**	−0.014**	−0.014**	−0.032**
	(0.001)	(0.000)	(0.001)	(0.001)
Observations	1,675,819	1,675,819	1,675,819	1,675,819
R^2	0.001	0.012	0.005	0.016
F	27.88	308.1	147.0	452.9
Compulsory mover	0.000	0.001*	0.000	−0.000
	(0.001)	(0.000)	(0.000)	(0.001)
Observations	1,675,819	1,675,819	1,675,819	1,675,819
R^2	0.000	0.000	0.000	0.001
F	17.46	14.42	5.41	58.80

Notes: Robust standard errors in parentheses. Clustered by grade 2 school. $^{**}P < 0.01$, $^{*}P < 0.05$, $^{+}P < 0.1$.

much less. The SD is 36 miles (compared with 86 miles for non-compulsory movers), and the correlation with grade 2 scores is lower: the difference in average distance between the first and the fourth quartile of test scores is about 0.9 miles (compared with 17 miles for non-compulsory movers).

Dynamic sorting across peer groups and school types

The previous discussion indicates that non-compulsory movers may sort endogenously into schools. Although compulsory mobility provides us with a more exogenous cause of mobility than does the non-compulsory case, it remains to be shown that compulsory moves also provide a more exogenous direction of mobility. This section provides suggestive evidence that, from grade 2 to grade 6, changes in peer group composition and in school types are much more strongly correlated with non-compulsory than with compulsory mobility.

We start by sorting schools into school types (see section II's description) and regress the probability of going to each school type on the non-compulsory mover dummy. Formally, we have[18]

$$\text{School Type}_{i,6} = \beta \cdot \text{Non-compulsory Mover}_i + \alpha \cdot \text{School Type}_{i,2} + \text{Controls}_i + \text{Residual}_i$$

Here School $\text{Type}_{i,6}$ and School $\text{Type}_{i,2}$ are the school types in grade 2 and in grade 6 respectively; and Non-compulsory Mover_i is a dummy set equal to 1 when student i is a non-compulsory mover (and to 0 otherwise). The regression is at the student level. The term Controls_i is a set of student demographic controls: gender, race, free meal status, special-needs and English spoken at home.

The data summarized in the upper panel of Table 6 reveals that non-compulsory movers are significantly less likely to leave LEA schools than to leave any other school type. It is therefore noteworthy that the table's lower panel, which regresses school type on the Compulsory mover dummy, shows that this dummy is significant only for the 'Employer/Admission' category of schools. That is, conditional on a student's characteristics and previous school type, compulsory movers are not more or less likely to attend a foundation, LEA-controlled or faith school.

Next we look at endogenous sorting across peer groups. To start with, we regress the peer group composition of each student's grade 6 peer group composition on that student's grade 2 composition, with a compulsory mover dummy, and a set of controls. The results (presented in Table (6), upper panel) indicate that, conditional on peer composition in grade 2, non-compulsory movers move to schools with fewer minority students and slightly more boys. The table's lower panel presents results of the regressions with a compulsory mover dummy; in this case, compulsory movers do not significantly sort according to particular peer groups conditional on their own or their grade 2 peer group's characteristics.

Another robustness check is to look at year-to-year changes in grade composition – which, as Hoxby (2000) points out, should be random around each school's average. Data S1 presents statistical evidence of such randomness.

IV. Using compulsory movers to estimate school quality and peer effects

We estimate the education production function by focusing on compulsory movers only. This section enriches specification (1) by adding the effect of the student's peers and previous school on current achievement. We also estimate the magnitude of the endogenous mobility bias by comparing school quality estimates based on the compulsory movers only and, for the same school, the school quality estimate based on all movers.

Specifications

The education production function considered in section II was rudimentary so that we could clearly show the scope for compulsory movers to alleviate mobility bias. In this

[18] The model also augments the worker–firm equation of Abowd, Kramarz and Margolis (1999) with past firm effects, though there are obviously differences between wage and education functions.

section, we present a similar education production function for test scores in which past school quality affects current achievement; from this we will be able to decompose test scores into pupil, school and peer effects. We will compare estimates based on the entire population to estimates based on the subset of compulsory movers.

Our specifications combine two important features: (i) test scores are a function of the effects of pupil background, school quality and peers, hence school quality is estimated conditional on peers' observable characteristics and students non-time- varying unobservables; (ii) the whole history of inputs shapes each pupil's experience (c.f. Todd and Wolpin, 2003; Rivkin *et al.*, 2005): school quality and peer effects have an impact both on current and later test scores. Formally, we have:

$$\begin{aligned}\text{Test Score}_{i,f,t} = {} & \text{Controls}_{i,f,t}\beta + \text{Peers}_{-i,g(i,t)}\gamma + \text{Student}_i + \text{School}_{j(i,t)} \\ & + \text{Peers}_{-i,g(i,t-1)}\gamma_{-1} + \lambda\text{School}_{j(i,t-1)} + \lambda\text{Student}_i + \varepsilon_{i,f,t}\end{aligned} \tag{3}$$

where the past inputs are included for grade 6 observations. In each of these specifications there are two Key Stage periods ($t = 1, 2$), i denotes the N pupils with $i = 1, 2, \ldots, N$, and j denotes J the schools with $j = 1, 2, \ldots, J$. The term Test $\text{Score}_{i,f,t}$ is the standardized test score of pupil i at time t in examination field f.[19] $j(i, t)$ denotes the school that pupil i attended at time t, and $g(i, t)$ denotes the grade which pupil i attends in year t. Finally $\text{Controls}_{i,f,t}$ is a vector of dummies that control for the field, the cohort and the grade, while $\text{Peers}_{-i,g(i,t)}$ is a vector of peers' observable characteristics. The peer group is defined here as the set of students in the same school in the same grade in the same year.

The regression also includes the effect of past inputs on current achievement. Thus, $\lambda = 0$ if there is no impact of the previous school on current achievement. If $\lambda = 1$, then the model is similar to a pure value-added model in which the school effect measures the impact of the school on the student's progress $\text{Test Score}_{i,f,t} - \text{Test Score}_{i,f,t-1}$.[20] Yet in contrast with the value-added literature this specification accounts for the progress of students being affected by their individual characteristics, their previous and current peers and schools. Our estimation procedure based on an iterated technique is described in Data S1.

The endogenous mobility bias: comparing the estimates derived using compulsory and non-compulsory movers

In assessing endogenous mobility bias, one of the contributions of this paper is to establish how school quality estimated on compulsory movers differs from estimates based on all movers. Non-compulsory movers are more likely to move because of unemployment or divorce shocks, both of which negatively affect test scores (Gibbons, 2007) and hence result in a downward bias on estimates of school effects that include both compulsory and non-compulsory movers. To see this, we estimate the education production functions on the entire sample, which includes stayers in addition to non-compulsory and compulsory

[19] Since we are pooling test scores for Key Stage 1 and Key Stage 2, differences in the distributions of test scores (skewness, kurtosis, upper and lower bounds) between Key Stage 1 and 2 may affect the interpretation of the results. We checked that using the percentile of the test score in each Key Stage and each field instead of the standardized test score does not affect the results.

[20] Data S1 presents estimations of the model with a different discount factor for student and school effects.

TABLE 7

Comparison of the estimates of the school effect on compulsory movers and on non-compulsory movers

	(1)	*(2)*
	Estimated mobility bias	*Estimated mobility bias*
Fraction of non-compulsory movers among movers	−0.472* (0.192)	
<25%		Ref.
25–50%		0.112 (0.209)
50–75%		−0.101 (0.144)
>75%		−0.401** (0.144)
R^2	0.002	0.004
F	6.058	3.321
Observations	2,521	2,521

Notes: The estimated mobility bias is, for each school, the difference between the school effect estimated on all students enrolled in the schools and the school effect for this school estimated on compulsory movers only. Standard errors in parentheses. Clustered by local education authority. $^{**}P < 0.01$, $^{*}P < 0.05$, $^{+}P < 0.1$.

movers. For this comparison, a substantial number of estimates of grade 6 school effects are available in both samples. Indeed, 7,279 (out of 16,149) schools welcome compulsory and non-compulsory movers in grade 6. For these schools, the estimate of the school effect of school j based on the compulsory movers only is denoted $\widehat{\text{School}}_j^{\text{cm}}$, and the estimate of the school effect of the same school based on all movers is denoted $\widehat{\text{School}}_j^{\text{all}}$.

Table 7 presents the results from regressing the $\text{Difference}_j = \widehat{\text{School}}_j^{\text{all}} - \widehat{\text{School}}_j^{\text{cm}}$ on the fraction of non-compulsory movers among the incoming movers in grade 6 from another school in grade 2. Results suggest that identification relying on non-compulsory movers substantially biases school effects downward. A school effect identified solely via non-compulsory movers is 5% of a SD lower than the same school effect identified on compulsory movers only. Column (2) shows that the effect is nonlinear, so that the bias kicks in for schools in which more than 75% of movers are of the non-compulsory type. There are 706 (9.6%) such schools among the 7,279 schools that welcome both compulsory and non-compulsory movers. If, in addition, we include schools that welcome only non-compulsory movers, the bias concerns 8,069 (grade 6) schools.[21]

[21] Therefore 7,363 schools of the sample welcomed non-compulsory movers only in grade 6.

TABLE 8

Biases in the estimation of school quality

	Dependent variable: Estimated mobility bias			
	Subsample: Quartile of the School Effect Distribution			
	(1)	*(2)*	*(3)*	*(4)*
	1st Quartile	*2nd Quartile*	*3rd Quartile*	*4th Quartile*
School is employer of teachers & controls student admission	−0.071	−0.045	0.091	0.362**
	(0.120)	(0.069)	(0.097)	(0.099)
Controlled by foundation	0.452*	0.219+	0.033	−0.422**
	(0.190)	(0.120)	(0.137)	(0.197)
Church of England school	0.149	−0.051	−0.127	−0.338**
	(0.091)	(0.068)	(0.080)	(0.122)
R^2	0.017	0.005	0.004	0.027
F-statistic	8.661	1.563	2.071	23.86
Observations	263,915	295,026	269,854	262,586

Notes: Robust standard errors in parentheses. Standard errors clustered by local education authority. $^{**}P < 0.01$, $^{*}p < 0.05$, $^{+}P < 0.1$.

Mobility bias: school organization and curriculum

Are schools that are organized under particular structures better than others?[22] In section II we described the variety of ownership and organization structures of schools in England. In Table 8 we address mobility and quality in terms of the three (non-mutually exclusive) types of schools that are not controlled by the local education authority: schools that are controlled by a foundation; schools that are controlled by the Church of England (an example of 'faith' schools); and schools that employ their staff and have some leeway in admissions.

Most studies of school quality rightly point out that students' time-varying characteristics may confound the estimation of school quality, which administrative data sets do not include. Table 8 suggests that this bias does obtain in high-quality faith schools. The table regresses the endogenous mobility bias (as defined as in section IV), on dummies for each school type, where the reference group is the community schools controlled by LEAs. Each column lists regression results for one of the four quartiles of the school effects estimated on compulsory movers only. For instance, column (4) shows the regression for the schools that have the highest (fourth quartile) school quality measures. It can be seen that quality of faith schools in the distribution's fourth quartile of the distribution is underestimated (by about 3.4% of a SD) when we estimate on the overall sample of students. Indeed, additional evidence indicates that non-compulsory movers tend to move to faith schools with large school effects.[23] The third row of that Table 10 shows that overall, non-compulsory movers tend to bias the school effect of faith schools upward for schools in the first quartile and downward in the other quartiles. This bias is monotonic, from +1.5% of a SD to −3.4% of a SD. Thus, the variance in school effects of faith schools is underestimated: there is

[22] From a policy perspective, UK governments have set up a number of new schools with different ownership and management structures – for example, 'city academies' (Machin and McNally, 2011). In the research literature, US Catholic schools have garnered substantial attention (Card, Dooley and Payne, 2010).

[23] Results available from the authors.

TABLE 9

Estimation results on compulsory movers

	Sample:	
	Compulsory movers *Dependent variable:*	*All students*
	Test score (1)	*Test score* (2)
Effects		
SD (student)	8.720	9.328
	(0.066)	(0.022)
SD (school)	2.018	1.886
	(0.583)	(0.130)
Corr (student, school)	−0.089	−0.068
	(0.002)	(0.002)
Peers		
Fraction free meal	−1.169**	−0.134
	(0.129)	(0.108)
Fraction black	0.489	1.541**
	(0.347)	(0.197)
Fraction asian	0.906+	−0.859**
	(0.493)	(0.282)
Fraction mixed ethnicity	0.151	−0.576
	(0.419)	(0.389)
Fraction boys	0.628**	−0.006
	(0.204)	(0.081)
R^2	0.80	0.80
F-statistic	11.31	11.02
P-value	<0.000	<0.000
Observations	2,201,308	6,481,082

Notes: Estimated on the sample of compulsory movers, i.e. students who started school in an infant-only school. The estimated discount factor λ is 0.024, significant at 1% with a SE of 0.001. Data S1 presents alternative specifications with a different λ for students and schools. Bootstrap standard errors in parentheses. Standard errors clustered by school. $^{**}P < 0.01$, $^{*}P < 0.05$, $^{+}P < 0.1$.

more variance in the quality of faith schools than is indicated by regressions that use both non-compulsory and compulsory movers.

A similar finding emerges from looking at biases for the quality of foundation-controlled schools. The first quartile of the distribution of school quality is upward biased, while the 4th quartile of school quality is downward biased. Thus the variance of the quality of foundation schools is underestimated.

However, we find that non-compulsory movers tend to leave high-quality schools that employ their teachers and control admissions. As a result, the school effect of this type of school is downward biased for schools in the upper (fourth) quartile of the school effect distribution. The bias is about 3.6% of a SD, which is substantial: it amounts to approximately 3.6/2.0 = 18% of the variance in the school effects (second row of Table 9).

Bias in the estimation of variance in school quality

What is the impact of schools on achievement? More specifically: how different are the test scores of students attending schools at the top of the quality distribution from those at the bottom of that distribution? Finally, is it the schools that account for such differences. These questions have been extensively discussed in the literature (Rivkin *et al.,* 2005) relying implicitly on non-compulsory student mobility for identification.

Some clear stylized facts emerge in Table 9, which estimates specification (3). Before detailing these take-aways from the estimation tables, we first note that the variance in school effects is not substantially affected by measurement error. We can use the central limit theorem approximation: var(Measurement Error) $\approx \frac{\text{var(School)}}{n_{School}} \approx \frac{\text{var}(\widehat{\text{School}})}{n_{School}}$, where n_{School} is the average number of students per school. Since there are on average 334 students per school, the estimated variance in school effects is inflated by only 0.6% of a SD.[24]

First, the SD of pupil fixed effects is about 4.4 times the SD of school effects. This finding suggests, perhaps not surprisingly, that pupils are more heterogeneous than schools. Nevertheless, pupil fixed effects are less precisely estimated than school effects. Indeed, at most four observations per child are available and there is more noise in the estimation of student effects than in school effects. Even after correcting for the upward bias in measurement of student effects, the variance in these effects is 6.5, or about 3.2 times the variance in school effects.[25]

The second stylized fact is that the variance in school effects estimated on compulsory movers is greater (by 2% of a SD) than that estimated on all students. We observe that non-compulsory movers, who experience downward shocks to their achievement, tend to move to higher quality schools. A regression (results available from the authors) of grade 6 school quality on grade 2 school quality and a non-compulsory mover dummy suggests that non-compulsory movers move to grade 6 schools with higher school effects. Thus, we should expect that the difference between good and bad schools will be underestimated, as the estimates of school effects for 'good' schools will be confounded by the negative time-varying unobservables of the non-compulsory movers.

Finally, the variance of the part of test scores predicted by contemporaneous peer effects, var($\text{Peers}_{-i,g(i,t)}$) is small, <0.3% of a SD. Hence student effects have a much larger level of heterogeneity than peer effects.

Bias in the estimation of peer effects

Column (1) of Table (9) presents the results of estimating the education production function on the sample of compulsory movers. In contrast, column (2) presents the results of estimating that function for all students who attend a school with at least one compulsory

[24] Also, schools with compulsory movers have nearly as many (291) students, and there is no substantial difference in the bias in the estimation of variance in school effects across the sample of compulsory vs. non-compulsory movers.

[25] The variance and correlation structure of the student, school and peer effects is also estimated in one step using a method of moments estimator. This method of moments provides consistent estimators of the variance in student and school effects as well as of the correlation between student and school effects. Results are not substantially different from those reported in Table 9.

TABLE 10

Explaining peer effects estimates – conditional and unconditional results – estimated on compulsory movers only

	Dependent variable: test score Specification:	
	OLS	*Fixed effects*
	(1)	*(2)*
Effects		
SD (student)	–	−9.460
		(0.113)
SD (school)	–	2.025
		(0.563)
Corr (student, school)	–	−0.039
		(0.004)
λ	–	0.024**
Long-run effect	–	(0.001)
Peers		
Fraction free meal	−2.744**	−0.887**
	(0.016)	(0.304)
Fraction black	2.809**	–
	(0.217)	
Fraction asian	−1.011**	–
	(0.145)	
Fraction mixed ethnicity	3.599**	–
	(0.444)*	
Fraction boys	−1.903**	0.090
	(0.130)	(0.125)
R^2	0.32	0.80
F-statistic	32,408	20.11
P-value	<0.000	<0.000
Observations	2,201,308	2,201,308

Notes: Column (1) is the OLS regression controlling for student observables, without school or student fixed effect. Columns (2) control for student, school effects and past inputs. All regressions of this table are estimated on the sample of compulsory movers. Bootstrap standard errors in parentheses. Standard errors clustered by school. $^{**}P < 0.01$, $^{*}P < 0.05$, $^{+}P < 0.1$.

mover. Thus, in column (2), non-compulsory movers contribute to the identification of peer effects.

Table 4 (in section III) shows that non-compulsory movers transfer to peer groups with fewer Black and Asian peers, and fewer free meal peers. Therefore since Table 2 (in section II) suggests that non-compulsory movers experience large drops in achievement and adverse family events, we expect the peer effects of Black, Asian and free meal peers to be upward biased in column (2) of Table (9).

In fact, this is what the results suggest for Black and free meal peers. An increase of 10ppt in the fraction of free meal peers lowers test scores by 1.17% of a SD (significant at 1%) but only by 0.13% of a SD when using both compulsory and non-compulsory movers in column (2). An increase in the fraction of Black peers has no significant impact on achievement when restricting the sample to compulsory movers, but it does have a significantly positive impact when estimating on both non-compulsory and compulsory movers (1.5% of a SD for a 10% increase in Black peers).

Section III (Table 4) also suggests that non-compulsory movers tend to sort into schools with more male peers, by about 0.3 percentage points higher in grade 6 than in grade 2. This is significant because the fraction of boys does not vary significantly across schools. Given that non-compulsory movers experience downward shocks on achievement, we should expect the impact of male peers to be underestimated in the regression with all students; indeed, this is what we observe. The impact of male peers is positive and significant. A 10 percentage point increase in the fraction of male peers in the grade increases achievement by 0.6% of a SD.

Column (1) of Table 9 shows that on average, boys have a significantly positive effect on student achievement conditional on other peers' ethnicity and free meal status. A number of papers (e.g. Hoxby, 2000) have showed results suggesting boys have a negative impact on their peers. Table 10 clarifies the meaning of this result. First, column (1) shows that, in an ordinary least-squares regression controlling for student observables but not for student effects or school effects, a larger fraction of male peers is indeed correlated with lower achievement. Hence our regression results suggest that controlling for student and school effects makes a significant difference, since student and school unobservables capture some potential confounders.

Second, column (2) of Table 10 reports results of the regression with student effects, school effects, past inputs and peers' gender and free meal status. The results indicate that the positive effect of male peers holds only if we control for peers' ethnicity; otherwise, the effect is small and non-significant.

V. Conclusion

Student mobility offers the main source of identification in education production functions that control for student and school fixed effects. This mobility is unlikely to be a credibly exogenous source of identification since student mobility is correlated with observed and unobserved factors that affect student achievement (e.g. the literature describes family events like parental divorce and unemployment to mobility). This paper, therefore, focuses on the subset of students, 'the compulsory movers', who must change school because their school only caters for grades 1 and 2 students. Because our data set includes both compulsory and non-compulsory movers, we have a unique opportunity to estimate the same school effect both with and without non-compulsory movers. In doing so, it turns out that the magnitude of the endogenous mobility bias can be substantial: when all the movers to a school are of the non-compulsory type, the result is a downward bias on school effects of about 4.7% of a SD of test scores, which is about 24% of the SD of school effects. Estimates of education production functions will also underestimate or overestimate the variance of school effects. In England, non-compulsory movers tend to move to higher-quality schools (Gibbons, 2007) and so the difference between good and bad schools is underestimated.

We conclude with two additional observations. First, the addition of time-varying student observables to administrative data sets would potentially offer significant gains to learn better about what improves student performance. If education authorities across the world were able to begin doing so, our view is that this would be a significant development in the provision of administrative data in different schooling environments. Second, our econometric results are also applicable to other literatures that employ dynamic matched data, as displaced workers (Sullivan and von Wachter, 2009) could provide sources of mobility in such models.

Final Manuscript Received: June 2014

References

Abowd, J. M., Kramarz, F. and Margolis, D. N. (1999). 'High wage workers and high wage firms', *Econometrica*, Vol. 67, pp. 251–334.

Burkam, D. T., Lee, V. E. and Dwyer, J. (2009). *School Mobility in the Early Elementary Grades: Frequency and Impact from Nationally Representative Data*, Technical report, Workshop on the Impact of Mobility and Change on the Lives of Young Children, Schools and Neighborhoods.

Card, D., Dooley, M. and Payne, A. (2010). *School Competition and Efficiency with Publicly Funded Catholic Schools*, Department of Economics Working Papers No. 2010-01, McMaster University.

Card, D. and Krueger, A. B. (1992). 'Does school quality matter? Returns to education and the characteristics of public schools in the united states', *Journal of Political Economy*, Vol. 100, pp. 1–40.

Coleman, J. (1966). *Equality of Educational Opportunity*, U.S. GPO, Washington, D.C.

Department for Education and Skills. (2011). *Code of Practice on LEA Schools Relationships*, Department for Education and Skills, London.

Dustmann, C., Puhani, P. A. and Schönberg, U. (2012). *The Long-Term Effects of School Quality on Labor Market Outcomes and Educational Attainment*, CReAM Discussion Paper Series No. 1208, Centre for Research and Analysis of Migration (CReAM), Department of Economics, University College London.

Gibbons, S. (2007). *Mobility and School Disruption*, Centre for the Economics of Education Discussion Papers.

Gould, E., Lavy, V. and Paserman, D. M. (2004). 'Immigrating to opportunity: Estimating the effect of school quality using a natural experiment on ethiopians in Israel', *The Quarterly Journal of Economics*, Vol. 119, pp. 489–526.

Graham, B. S. (2008). 'Identifying social interactions through conditional variance restrictions', *Econometrica*, Vol. 76, pp. 643–660.

Hoxby, C. (2000). *Peer Effects in the Classroom: Learning from Gender and Race Variation*, NBER Working Papers No. 7867, National Bureau of Economic Research, Inc.

Hoxby, C. M. and Weingarth, G. (2007). *Taking Race Out of the Equation*, mimeo.

James-Burdumy, S. (2005). 'The effect of maternal labor force participation on child development', *Journal of Labor Economics*, Vol. 23, pp. 177–211.

Machin, S. and McNally, S. (2011). *The Evaluation of English Education Policies*, CEE Discussion Papers, Centre for the Economics of Education, LSE.

Machin, S. and Vignoles, A. (2005). *What's the Good of Education? The Economics of Education in the United Kingdom*, Princeton University Press, Princeton, NJ united states.

Manski, C. F. (2000). 'Economic analysis of social interactions', *Journal of Economic Perspectives*, Vol. 14, pp. 115–136.

Rivkin, S. G., Hanushek, E. A. and Kain, J. F. (2005). 'Teachers, schools, and academic achievement', *Econometrica*, Vol. 73, pp. 417–458.

Rockoff, J. E. (2004). 'The impact of individual teachers on student achievement: Evidence from panel data', *The American Economic Review*, Vol. 94, pp. 247–252.

Rothstein, J. (2010). 'Teacher quality in educational production: tracking, decay, and student achievement', *Quarterly Journal of Economics*, Vol. 125, pp. 175–214.

Sacerdote, B. (2001). 'Peer effects with random assignment: results for dartmouth roommates', *The Quarterly Journal of Economics*, Vol. 116, pp. 681–704.

Sacerdote, B. (2010). ‘Peer effects in education: how might they work, how big are they and how much do we know thus far?’ *Handbook of the Economics of Education*, Vol. 3, pp. 239–278.
Stevens, A. H. and Schaller, J. (2011). ‘Short-run effects of parental job loss on children’s academic achievement’, *Economics of Education Review*, Vol. 30, pp. 289–299.
Sullivan, D. and von Wachter, T. (2009). ‘Job displacement and mortality: an analysis using administrative data’, *The Quarterly Journal of Economics*, Vol. 124, pp. 1265–1306.
Summers, A. A. and Wolfe, B. L. (1977). ‘Do schools make a difference?’ *American Economic Review*, Vol. 67, pp. 639–652.
Todd, P. E. and Wolpin, K. I. (2003). ‘On the specification and estimation of the production function for cognitive achievement’, *Economic Journal*, Vol. 113, pp. F3–F33.

Supporting Information

Additional supporting information may be found in the online version of this article:

Data S1. Online appendix with robustness checks and estimation details.