The complexity of school and neighbourhood effects and movements of pupils on school differences in models of educational achievement

by

George Leckie[1]
University of Bristol, UK

Summary. Traditional studies of school differences in educational achievement use multilevel modelling techniques to take into account the nesting of pupils within schools. However, educational data are known to have more complex non-hierarchical structures. The potential importance of such structures is apparent when considering the impact of pupil mobility during secondary schooling on educational achievement. Movements of pupils between schools suggest that we should model pupils as belonging to the series of schools attended and not just their final school. Since these school moves are strongly linked to residential moves, it is important to additionally explore whether achievement is also affected by the history of neighbourhoods lived in. Using the national pupil database, this paper combines multiple membership and cross-classified multilevel models to simultaneously explore the relationships between secondary school, primary school, neighbourhood and educational achievement. The results show a negative relationship between pupil mobility and achievement, the strength of which depends greatly on the nature and timing of these moves. Accounting for pupil mobility also reveals that schools and neighbourhoods are more important than shown by previous analysis. A strong primary school effect appears to last long after a child has left that phase of schooling. The additional impact of neighbourhoods, on the other hand, is small. Crucially, the rank order of school effects across all types of pupils is sensitive to whether we account for the complexity of the multilevel data structure.

Keywords: Cross-classified models, Multiple membership-models, Multilevel modelling, Pupil mobility, School effectiveness, Value-added models

[1] Address for correspondence: George Leckie, Centre for Multilevel Modelling, University of Bristol, 2 Priory Road, Bristol, BS8 1TX, UK. E-mail: g.leckie@bristol.ac.uk

1. Introduction

Models of school differences in educational achievement typically assess the progress that pupils make between two test occasions and attempt to assess the extent to which variation between pupils is attributable to differences between schools. These models are commonly referred to as 'value-added' or 'school effectiveness' models and the current preferred practice is estimation using multilevel models (for early examples see: Aitkin and Longford, 1986; Goldstein et al., 1993; Raudenbush and Bryk, 1986).

Pupil mobility and neighbourhood effects are often discussed as important potential influences on educational achievement (ALG, 2005; DES, 2003; GLA, 2005; Ofsted, 2002). However, few value-added studies incorporate these factors into their analysis. Where studies look at the impact of whether a pupil has moved schools or not they find an overall negative association (e.g. Yang et al., 1999), but this has not been explored for different types and timings of moves. Furthermore, with the notable exception of Goldstein et al. (2007), the random part of the value-added models in these studies treat pupils as belonging to only their final schools and ignore the contribution of earlier schools attended. Similarly, although school moves are clearly linked to residential moves, no studies have incorporated this additional information into their analysis. The studies that have looked for neighbourhood effects on educational achievement have not been able to additionally model pupil movements (Fielding et al., 2006; Garner and Raudenbush, 1991). Until recently, research into pupil mobility has been held back by both a lack of data on pupil movements and also by the absence of appropriate multilevel methodology. However, the recently established national pupil database (NPD) in England and the development of cross-classified and multiple membership multilevel models now make it possible to analyse a wide range of complex non-hierarchical data structures in models of educational achievement (Browne et al., 2001; Fielding and Goldstein, 2006; Rasbash and Browne, 2001, 2008).

1.1 *Cross-classified models*

Traditional models of school effectiveness are two-level variance components models of pupils (at level 1) nested within schools (level 2). Incorporating neighbourhood as a further level is not straightforward since schools and neighbourhoods are not strictly nested within one another. Not all pupils who live in the same neighbourhood attend the same school and not all pupils from the same school live in the same neighbourhood. Rather than being nested within one another, schools and neighbourhoods are described as forming a cross-classification at level 2 within which pupils are nested. Cross-classified random effects models allow us to correctly partition the response variation between pupils, schools and neighbourhoods whilst explicitly allowing for the non-hierarchical nature of the data.

Garner and Raudenbush (1991) provide an early analysis of cross-classified data for 2500 pupils in Scotland nested within a cross-classification of 17 schools by 524 neighbourhoods. However, rather than estimating a random effects cross-classified model, they estimate a two-level random effects model of pupils (level 1) nested within neighbourhoods (level 2). In their unconditional model, neighbourhoods account for 20% of the total variation in scores. However, after they adjust for prior achievement, family background and neighbourhood social deprivation this drops to 6%. When they further incorporate schools through fixed effects of 16 dummy variables, just 4% of the remaining variation lies between neighbourhoods. In a reanalysis, Raudenbush (1993) estimates a full random effects cross-classified model that partitions the total variability between pupils, neighbourhoods and schools. A caveat is that with just 17 schools, the school component of variation is likely to be imprecisely estimated. In their unconditional model, schools and neighbourhoods together explain 20% of the total variation in scores, with neighbourhoods explaining two-thirds of this amount. Adjusting for the same predictors as Garner and Raudenbush explains most of this variation away and the ratio of school to neighbourhood residual variance increases from one-half to two-thirds. This rise is expected because the study adjusts for a neighbourhood level variable but does not adjust for any school level variables. Both studies find a strong negative effect of neighbourhood social deprivation even after adjusting for prior achievement and family background variables. More recently, Fielding et al. (2006) use cross-classified models in an analysis of a large scale dataset of over 80000 pupils in England. They find that neighbourhoods explain significant variation in pupils' educational achievement and progress, with greater variation found for smaller scales of neighbourhood. Contrasting Raudenbush (1993), they find schools more important than neighbourhoods. However, the two studies are not directly comparable as they refer to different contexts, different samples of pupils and importantly make different adjustments for pupil background characteristics.

Cross-classified models are also required to model any sustained or carryover effects of schools attended in an earlier phase of education on pupils' current progress. Goldstein and Sammons (1997) consider the persistence of junior school effects on pupil progress in secondary schools. They find the variance of junior school effects to be greater than that for secondary schools in their unconditional models and in models that adjust for prior achievement and other pupil background characteristics. They suggest that junior schools are so variable, partly because of the importance of early schooling, but also because they tend to be smaller. The smaller size of junior schools may lead them to capture pockets of heterogeneity in the sample to a greater extent than the larger secondary schools. The importance of earlier school membership is also reported by Rasbash and Goldstein (1994) and in a re-analysis by Browne et al. (2001). These two studies estimate unconditional cross-classified models for primary and secondary schools in Fife, Scotland. They find primary schools to be three times as variable as secondary schools. However, it should be noted that with just 19 secondary schools, the secondary school component of variation is likely to be imprecisely estimated. Goldstein et al. (2007), using larger datasets than all of the above studies, look at infant school effects on progress during junior schooling. In cross-classified models that adjust for prior achievement and other pupil background characteristics, they find inconsistent results across the two geographic areas they consider with the infant school variance being relatively larger in one area, but smaller in the other.

1.2 *Multiple membership models*

Between the two test occasions of a value-added analysis, pupils may change school. For these pupils, more than one school will contribute to their progress. Multiple membership models allow for this mobility. When specifying these models, an issue that arises is the relative importance, or weight, that should be attributed to each school attended. Browne et al. (2001) and Goldstein et al. (2007) both weight schools by the length of time spent in each one with the latter finding this weighting scheme to be near optimal in terms of model fit. This approach is also used by Fielding (2002) and Fielding and Yang (2006) who both estimate multiple membership models of teachers and teaching groups; weights are defined as the proportion of time teachers belong to each teaching group. These two studies also experiment with different weighting schemes and find that their main parameter estimates are relatively insensitive to the choice of weights. Goldstein et al. (2007) allow pupils to be multiple members of their junior schools. They illustrate that ignoring a multiple membership data structure leads to a known downward bias in the estimate of the corresponding (i.e. junior school) variance parameter (Goldstein, 2003). However, accounting for the junior school mobility makes little difference to the rank order of junior school effects. A potential caveat is that their analysis is limited to random-intercept models that assume the effectiveness of each school, although variable across schools, is constant across pupils within schools. It is not certain whether a similar result would apply in random coefficients models that additionally allow the effectiveness of each school to vary as a function of their pupils' characteristics.

This paper builds upon the work of Goldstein et al. (2007) to present a more detailed investigation of pupil mobility between schools, but also between neighbourhoods. The relative importance of secondary schools, neighbourhoods and primary schools on both achievement and progress are assessed for a much larger dataset than examined in previous studies. The negative association between mobility and progress is decomposed to investigate how it varies across the types and timings of moves. We also assess the importance of accounting for cross-classified and multiple membership structures for the rank order of school effects in random coefficient models. In Section 2, we introduce the general methodology for cross-classified and multiple membership models. Section 3 describes the data and variables used in the analysis. Section 4 presents the results from the analysis, and Section 5 concludes.

2. Methodology

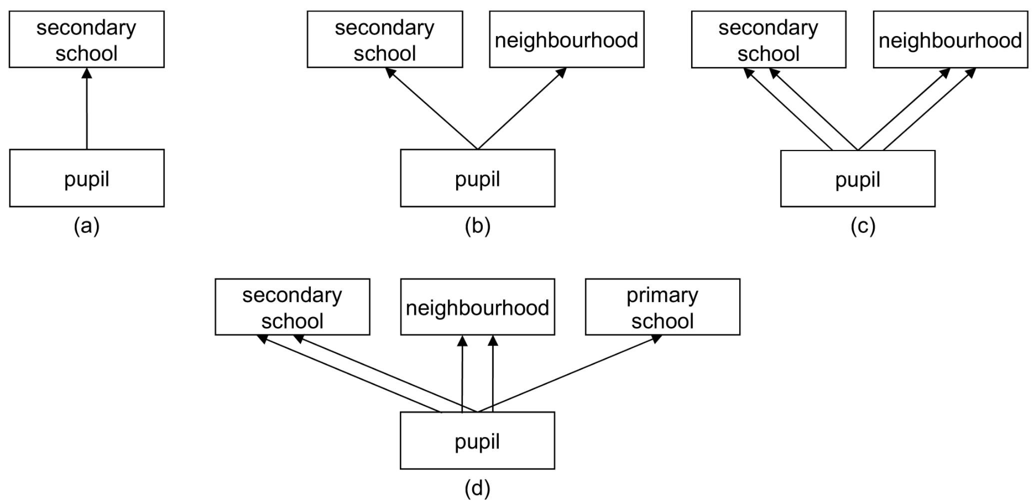
Consider a simple two-level model with an intercept and a single predictor variable. Using the 'classification' notation of Browne et al. (2001), this model can be written as

$$\begin{gathered} y_i = \beta_0 + \beta_1 x_i + u^{(2)}_{sec(i)} + e_i \\ sec(i) \in \left(1,\ldots,J^{(2)}\right), \qquad i = 1,\ldots,N \\ u^{(2)}_{sec(i)} \sim N\left(0,\sigma^2_{u(2)}\right), \qquad e_i \sim N\left(0,\sigma^2_e\right) \end{gathered} \tag{1}$$

where y_i is the test score for the ith pupil in the dataset and x_i is their prior achievement. There are two classifications: pupils and schools. The '(2)' superscripts and subscripts identify any variables or random effects that are associated with the school classification. The classification function $sec(i)$ denotes the ith pupil's secondary school. $u^{(2)}_{sec(i)}$ and e_i are, respectively, the school level and pupil level random effects which are assumed normally distributed, independent of one another, and independent of any predictor variables included in the model. Posterior estimates of the school effects are often used to rank schools in school 'league tables'.

Since classification notation does not show the multilevel structure in the data, 'classification diagrams' are typically presented in addition to the model equation (Browne et al., 2001). Fig. 1a depicts a classification diagram for the simple two-level hierarchy of model (1). The pupil and school classifications are represented by boxes whilst the single arrow from the pupil to the school classification indicates the nesting of pupils within schools. Fig. 1b depicts pupils nested within a cross-classification of schools and neighbourhoods by drawing the neighbourhood classification box at the same horizontal level as the school classification box. Fig. 1c depicts pupils as potentially belonging to multiple schools and multiple neighbourhoods by replacing each single arrow with a double arrow. Finally, Fig. 1d includes a third cross-classification with primary school.

Fig. 1. Classification diagrams for different hierarchical and non-hierarchical data structures (a) simple two-level nested model (b) cross-classified model of secondary schools with neighbourhoods (c) multiple membership model of secondary schools crossed with a multiple membership of neighbourhoods (d) multiple membership model of secondary schools crossed with a multiple membership of neighbourhoods crossed with primary schools



The final classification diagram (Fig. 1d) depicts the complex data structure of the main model presented in the analysis. This model, for the case of a single predictor, is written as follows

$$y_i = \beta_0 + \beta_1 x_i + \sum_{j \in sec(i)} w_{i,j}^{(2)} u_j^{(2)} + \sum_{j \in nbhood(i)} w_{i,j}^{(3)} u_j^{(3)} + u_{pri(i)}^{(4)} + e_i$$

where

$$sec(i) \subset \left(1,\ldots,J^{(2)}\right), \qquad nbhood(i) \subset \left(1,\ldots,J^{(3)}\right), \qquad pri(i) \in \left(1,\ldots,J^{(4)}\right), \qquad i = 1,\ldots,N$$

$$\sum_{j \in sec(i)} w_{i,j}^{(2)} = 1, \qquad \sum_{j \in nbhood(i)} w_{i,j}^{(3)} = 1$$

$$u_{sec(i)}^{(2)} \sim N\left(0,\sigma_{u(2)}^2\right), \qquad u_{nbhood(i)}^{(3)} \sim N\left(0,\sigma_{u(3)}^2\right), \qquad u_{pri(i)}^{(4)} \sim N\left(0,\sigma_{u(4)}^2\right), \qquad e_i \sim N\left(0,\sigma_e^2\right) \quad (2)$$

The classification functions $sec(i)$, $nbhood(i)$ and $pri(i)$ give sets of the ith pupil's secondary school, neighbourhood and primary school. We note that there may be more than one element in the multiple membership classifications for secondary school and neighbourhood, but in models considered here pupils 'belong' to only one primary school. The superscripts '(2)', '(3)' and '(4)' refer to the secondary school, neighbourhood and primary school classifications. The terms $w_{i,j}^{(2)}$ and $w_{i,j}^{(3)}$ are weights, each summing to one, which reflect the proportion of time a pupil has spent in each of their secondary schools and neighbourhoods, respectively. All random effects are assumed normally distributed and independent across classifications.

Care must be taken when interpreting the relative sizes of the variance components in model (2). For example, although $\sigma_{u(2)}^2$ is the variance of the secondary school effects $u_{sec(i)}^{(2)}$, the actual contribution of secondary schools to the variance for a given pupil is

$$\mathrm{var}\left(\sum_{j\in sec(i)} w_{i,j}^{(2)} u_j^{(2)}\right) = \sigma_{u(2)}^2 \sum_{j\in sec(i)} \left(w_{i,j}^{(2)}\right)^2$$

This contribution varies as a function of the number of schools a pupil attends and the time spent in each of those schools. For example, for children who attend a single school, the contribution is simply $\sigma_{u(2)}^2$ while for children who spend equal time in two schools, the contribution is just $0.5\sigma_{u(2)}^2 = \sigma_{u(2)}^2\left(0.5^2 + 0.5^2\right)$. Indeed, for pupils who attend multiple schools, the contribution is always less than that for stable pupils since the variance of a weighted sum of identically distributed random variables, with weights summing to one, is always smaller than the variance of the random variables themselves. This result has substantive appeal since we might expect that the more schools attended the more likely the positive effects of one school will be cancelled out by the negative effects of another (Fielding and Goldstein, 2006). Models which ignore the multiple membership structure will lead to biased estimates of the school effects, the extent of which increases with the degree of pupil mobility (Goldstein, 2003). Such models underestimate the true extent of between-school variation since they implicitly assume the contribution of schools to the variation of mobile and stable pupils is the same. For example, if half the pupils attend two schools for equal lengths of time, the between school variance given by a two-level model will be the average of the contribution of schools for stable ($\sigma_{u(2)}^2$) and mobile ($0.5^2\sigma_{u(2)}^2$) pupils, which is less than the true between school variation of $\sigma_{u(2)}^2$ given by a multiple membership model.

Model (2) includes a single pupil level predictor. Further predictors measured at any level can be easily added to the model. This should be done cautiously for the secondary school and neighbourhood classifications since, in the same way that we weight the secondary school and neighbourhood random effects, we should weight all school and neighbourhood fixed effect variables (Fielding, 2002; Fielding and Yang, 2005; Goldstein et al., 2007). These weighted fixed effects will better reflect the school and neighbourhood peer groups and environments that pupils have been exposed to over their secondary schooling. Model (2) describes a random intercepts model, but the model can be extended to incorporate random slopes at one or more of the higher classifications.

Estimation of cross-classified and multiple membership models by existing maximum likelihood approaches run into important computational limitations, especially when large numbers of units are involved (Browne et al., 2001). As a result, the following models are fitted using Markov chain Monte Carlo (MCMC) based algorithms as implemented in the MLwiN package (Rasbash et al., 2004). Starting values for the fixed parameters are estimated from simpler models using a maximum likelihood approach, iterative generalised least squares (IGLS, Goldstein, 1986), in MLwiN. The Bayesian deviance information criterion (DIC, Spiegelhalter et al., 2002), a model complexity measure, is used to compare the fit of models estimated by MCMC. Models with smaller DIC values are preferred to those with larger values, with differences of 10 or more considered substantial. Further details of the MCMC estimation methodology are given in Browne (2003).

3. Data

The exam data are taken from the national pupil database (NPD), a census of all pupils in the English state education system provided to us by the Department for Children Schools and Families (DCSF). We follow the cohort of pupils who took their General Certificate of Secondary Education (GCSE) exams in 2006 and key stage 2 (KS2) exams in 2001. These exams are taken in the last year of secondary schooling (age 16, academic year 11) and primary schooling (age 11, academic year 6), respectively. Successful GCSE results are often a requirement for taking General Certificate of Education Advanced level (A-levels) qualifications which are a common type of university entrance requirement. To GCSE scores, we merge data from the 2002-2006 pupil level annual school census (PLASC) datasets which give the series of schools attended and postcodes resided in between the two sets of exams. (Further information on the NPD and PLASC datasets and how to access them can be found at http://www.bris.ac.uk/Depts/CMPO/PLUG/whatisplug.htm)

The initial sample consists of the 530861 pupils who were present at all seven measurement occasions: GCSE, KS2 and in each of the five annual PLASC datasets. The analysis is limited to the 472431 pupils who took their GCSE exams in mainstream secondary schools that taught for all five years of the secondary phase of education. Thus, our sample excludes, for example, schools with non-standard age ranges such as middle schools and those that cater only for pupils with special education needs. Pupils are dropped from the sample if they have missing values for any of the variables used in the analysis. This reduces the sample by a further 4%. To ease the computational burden, we then restrict the sample to the 42681 pupils who took their GCSE exams in schools located in the South West region of England. Since our concern is with exploring the impact of mobility on models of educational progress, not inference from this sample to a larger population, this selection is felt appropriate.

3.1 *Variables used in the analysis*

The response is the total GCSE point score, capped for each pupil's eight best examination grades. We treat the response as continuous and, so that the multilevel residuals better approximate the normality assumptions of the models, monotonically transform the ranks of its values across the pupils to the corresponding expected values of order statistics from a standard normal distribution (Goldstein, 2003). Pupils who change schools score on average 0.47 less on this standardised scale than stable pupils whilst home movers score 0.24 less. These are nontrivial differences, especially given that some pupils move more than once. Prior achievement measures are derived from pupils' KS2 English, maths and science scores. To place these variables on a common scale and to ease their interpretation in the analysis, their distributions are also similarly transformed to standard normal scores.

In our models, we adjust for pupil background characteristics; these include: age, gender, English as an additional language (EAL), ethnicity, eligibility for free school meals (FSM) and an indicator of special educational needs (SEN). The FSM and SEN status of pupils varies over time with approximately 25% of pupils moving off FSM and SEN each year. In the analysis these variables are defined as the proportion of secondary schooling that pupils spent in these states rather than simply their status in the year which they took their GCSE examinations. However, we note that the timing of pupils' movements onto and off FSM and SEN may have a role in addition to the proportion of the value-added period spent in these states. Where pupils have ever been on FSM or SEN, they on average spent 60% of their secondary schooling in these states.

For each of the five years of secondary education, we know the school attended and the postcode where each pupil lives. The national statistics postcode directory (NSPD, http://www.statistics.gov.uk/geography/nspd.asp) is used to match in neighbourhood data. The chosen scale of neighbourhood is the lower super output area (LSOA), which are defined to be fairly consistent in size (they have a mean population of 1500) and to reflect as far as possible social homogeneity. Alternative spatial scales were considered, but these led to poorer model fit. We derive a range of mobility indicators and distinguish between 'compulsory' school moves - where pupils have to change schools because they reach the last year of their current school (e.g. pupils in middle schools or schools that close) - and 'non-compulsory' school moves - where pupils could have continued attending their previous school. We also identify the number of moves pupils make, the timing of these moves and whether pupils simultaneously move home or not.

School level contextual variables are included to capture the influence of pupils' peer groups. These variables are compositional variables derived for all schools in England from the original sample of pupils and not just those that took their GCSE examinations in schools in the South West. These variables are strictly school-cohort level variables as they are based only on the cohort of pupils who took their GCSE examinations in 2006. This offers a cleaner measure of the school context or school peer group for the pupils we are interested in than had we derived these variables from the sample of pupils that additionally includes younger and older cohorts of pupils who attend these schools. Variables include the average intake achievement and proportion of FSM pupils in each secondary school. We form weighted (by time spent in each school) versions of all school level variables since these are expected to better capture the influence of peer groups, especially for pupils who change school. At the LSOA neighbourhood level we include a measure of social deprivation, weighted across the series of LSOA neighbourhoods resided in: the 2004 index of multiple deprivation (IMD, ODPM, 2004). This index is a single overall measure of deprivation, aggregated from seven distinct sub domains of deprivation and is published by the government. In the analysis this variable is transformed to a standard normal score.

3.2 *Description of the non-hierarchical data structure*

The sample consists of 42681 children who took their GCSE exams in 264 secondary schools whilst living in 3175 neighbourhoods and had previously attended 3107 primary schools. The median secondary school has 161 pupils whilst the median neighbourhood has 14 pupils. Pupils are nested within a three-way cross-classification of secondary schools, neighbourhoods and primary schools. Since we observe pupils moving between secondary schools and also between neighbourhoods, there are also two multiple membership structures in the data. We cannot, however, treat pupils as multiple members of their primary schools as we only observe the final primary school they attend.

Out of 3175 neighbourhoods, 2571 (81%) have children who attend different GCSE schools with the median neighbourhood sending children to 3 different schools. Overall, 11873 out of 42681 children (28%) went to a secondary school other than the main one for their neighbourhood. Similar statistics can be calculated for primary schools and we see that the median primary school sends its pupils to 3 different secondary schools. The degree of 'imbalance' and 'sparsity' in the cross-classification - the unequal distribution of pupils across all possible school and neighbourhood combinations - is investigated but is not found to be problematic for the estimation of the cross-classified models (Fielding and Goldstein, 2006).

During secondary schooling, 8% of the sample changed schools. Adding in previously attended secondary schools raises the total number of schools in the data to 1346. 94 of these extra schools are schools in the South West that teach no pupils at GCSE and are middle schools or schools that have closed midway through the period of analysis. The remaining 988 extra schools are located outside the South West and tend to be the previous schools of pupils who have moved into the South West during their secondary schooling. The sample is restricted to pupils who took their GCSE exanimations in schools located in the South West and so the majority of these extra schools contain only one pupil. The presence of these extra schools is not problematic for the analysis as we are only interested in making inferences about GCSE schools located in the South West. We have all the pupils for these schools and, crucially, we weight these pupils by how long they attend each of their schools. Where pupils attended schools outside the South West, these earlier schools contribute very little to their GCSE scores. During secondary schooling, 27% of the sample move home with 23% also moving neighbourhood. Adding in the history of neighbourhoods lived in raises the total number of distinct neighbourhoods from 3175 to 4587. As with the extra schools, many of these extra neighbourhoods are located outside the South West and again tend to contain a single pupil. These extra neighbourhoods are incorporated into the multiple membership models in a parallel fashion, and with the same implications, as that for schools.

Table 1 describes the patterns of pupil movements between schools and between neighbourhoods. Whether we consider schools or neighbourhoods, pupils can belong to up to 5 different units during secondary schooling, with the proportion of time spent in each unit indicated by the columns Unit 1 - Unit 5. For each pupil, these proportions define the weights that are used in the multiple membership models reported in the results section. Unit 1 corresponds to the most recent school (neighbourhood) attended and Units 2-5 represent progressively less recent schools (neighbourhoods). The final four columns in the table show how pupils are distributed across the different duration patterns in terms of the schools and neighbourhoods they have attended. For example, the second row of the table informs us that 1188 or 2.78% of pupils attended a combination of two schools, where the first year (or 20%) of education (i.e. academic year 7) is spent in the first school and the remaining four years (or 80%) of education (i.e. academic years 8, 9, 10 and 11) are spent in the second school. Looking at the final two cells of the row, we see that 2631 or 6.16% of pupils attended a combination of two neighbourhoods spending 1 year in the first and 4 years in the second.

Table 1. Proportion of time spent in different secondary schools and different neighbourhoods over the 5-year secondary phase of education

Number of units	Proportion of time spent in each secondary school (neighbourhood) unit					Secondary schools		Neighbourhoods	
	Unit 1	Unit 2	Unit 3	Unit 4	Unit 5	Frequency	%	Frequency	%
1	1.0					39138	91.70	32990	77.29
2	0.8	0.2				1188	2.78	2631	6.16
2	0.6	0.4				984	2.31	1861	4.36
2	0.4	0.6				671	1.57	1743	4.08
2	0.2	0.8				317	0.74	1513	3.54
3	0.6	0.2	0.2			110	0.26	357	0.84
3	0.4	0.4	0.2			81	0.19	292	0.68
3	0.4	0.2	0.4			89	0.21	291	0.68
3	0.2	0.6	0.2			29	0.07	230	0.54
3	0.2	0.4	0.4			28	0.07	189	0.44
3	0.2	0.2	0.6			20	0.05	266	0.62
4	0.4	0.2	0.2	0.2		15	0.04	92	0.22
4	0.2	0.4	0.2	0.2		2	0.00	66	0.15
4	0.2	0.2	0.4	0.2		1	0.00	73	0.17
4	0.2	0.2	0.2	0.4		7	0.02	57	0.13
5	0.2	0.2	0.2	0.2	0.2	1	0.00	30	0.07
						42681	100.00	42681	100

4. Results

4.1 *Models of GCSE achievement*

Table 2 reports intercept-only variance components models for the normalised GCSE score. These models use cross-classified and multiple membership structures to explore the relative importance of secondary schools, neighbourhoods, primary schools and pupils in explaining GCSE achievement. Model A is the standard two-level model, allowing for the nesting of pupils within the secondary schools in which they took their GCSE examinations (see Fig. 1a). The estimated pupil and school level variances are 0.818 and 0.223 giving a variance partition coefficient (VPC, Goldstein et al., 2002) of 0.214 $= 0.223/(0.223+0.818)$; 21% of the total variation in GCSE scores lies between secondary schools.

Model B treats pupils as nested within a cross-classification of the secondary schools attended and the LSOA neighbourhoods resided in at the time of the GCSE examinations (see Fig. 1b). The DIC reduces by 1334 suggesting a substantial improvement in the fit of the model. Introducing the neighbourhood classification leads to a reduction in both the school and pupil variance terms, indicating that part of the unexplained variation in GCSE scores in model A had been wrongly attributed to these two levels. The neighbourhood variance of 0.054, although smaller than the secondary school variance, still leads to sizeable differences in the average GCSE score across neighbourhoods. The ratio of school-to-neighbourhood variation is approximately four and is consistent with Fielding et al. (2006), but not Raudenbush (1993) who finds a ratio of only one-half. However, little should be read into the difference between our results and the latter since their estimate of their between school variance is based on just 17 schools and is therefore likely to be estimated imprecisely. A neighbourhood-by-school random interaction effect (Fielding and Goldstein, 2006; Goldstein, 2003; Raudenbush and Bryk, 2002) is also considered but the within cell sample sizes are not sufficient to separately identify the interaction variance from the pupil variance.

Table 2. Parameter estimates for intercept-only variance components models of normalised GCSE scores.
Model A: Simple two-level nested model of pupils within the secondary schools they attend at the time of the GCSE examinations
Model B: Cross-classified model of the secondary schools attended with the LSOA neighbourhoods resided in at the time of the GCSE examinations
Model C: Multiple membership model of secondary schools crossed with a multiple membership of neighbourhoods
Model D: Multiple membership model of secondary schools crossed with a multiple membership of neighbourhoods crossed with primary schools

	Model A		Model B		Model C		Model D	
Fixed Part								
Constant	0.008	(0.028)	0.012	(0.028)	-0.155	(0.028)	-0.147	(0.030)
Random Part								
Secondary	0.223	(0.020)	0.204	(0.019)	0.257	(0.027)	0.248	(0.025)
Neighbourhood			0.054	(0.003)	0.064	(0.003)	0.045	(0.003)
Primary							0.033	(0.003)
Pupil	0.818	(0.006)	0.768	(0.006)	0.762	(0.006)	0.747	(0.005)
DIC	112779		111445		111260		110818	

Note: Standard errors in parentheses. MCMC estimation used a burn in of 500 and a chain length of 5000. Convergence was judged using the MCMC diagnostics in MLwiN, for example, the ESS and the Brooks Draper statistics (see Browne, 2003 for details).

Model C extends model B by introducing two multiple membership structures to account for pupil mobility between schools and between neighbourhoods (see Fig. 1c). The model sets multiple membership weights equal to the proportion of time spent in each school and in each neighbourhood (see Table 1). Although these weights are simple and intuitive, they do not convey information about the timing or the ordering of schools attended. We have carried out sensitivity checks of our models to alternative weighting schemes. In particular, we consider a range of schemes that give increasingly more weight to schools and neighbourhoods attended towards the end of secondary schooling. With each scheme, the results for model C become closer to those for model B and at the extreme, where pupils are allocated to only their most recent school, model C collapses to model B and the results are the same. For all our multiple membership models, the weighting scheme based on the proportion of time spent in each school or neighbourhood leads to approximately the best model fit. Fielding (2002), Fielding and Yang (2005) and Goldstein et al. (2007) also find that the results of their multiple membership models are robust to different weighting schemes.

Incorporating the multiple membership structures into the model leads to a modest improvement in the DIC of 185 points. The intercept term is now estimated as -0.155 compared to 0.012 in model B. This decrease is caused by the extra higher level units that are included in the model when we account for the multiple membership structures. Many of these extra schools and neighbourhoods lie outside the South West and, in our sample, often contain just one or two mobile pupils who are typically low achieving and so the mean GCSE scores for these additional units are mostly negative. The estimated intercept decreases because it is an empirical Bayes estimate that places relatively more importance on between school and between neighbourhood differences than is the case for a simple arithmetic mean. In the random part of model, schools and neighbourhoods appear to be more important than before: the between school variance increases by 26% over model B whilst the between neighbourhood variance increases by 18%. These increases are expected since ignoring multiple membership structures is known to lead to downward biases in the associated variances (Goldstein, 2003). However, a second cause of the increase is the extension of the sample, to schools and neighbourhoods outside the South West. In investigating this, we re-estimated models B and C for the subset of pupils who only live and attend schools in the South West. This removes approximately half the pupils who ever change school and one-fifth of those who ever move neighbourhood. The reduced sample results for model B are very similar to those for the full sample, but this is not the case for model C. Moving from model B to C in the reduced sample, the between school variance now increases by only 9% and the between neighbourhood variance by 17%. The extra schools clearly have a larger inflationary influence on the secondary school variance than the extra neighbourhoods do on the neighbourhood variance and this is because the extra schools account for a much higher proportion of pupils' previous schools attended than the corresponding proportion for neighbourhoods.

Model D extends model C to include a third cross-classification for primary school (see Fig. 1d). The DIC improves by a further 442 points. The primary school variance is estimated as 0.033, which is slightly smaller than that for neighbourhoods and substantially smaller than that for secondary schools. This result contrasts with the unconditional models of Goldstein and Sammons (1997) and Rasbash and Goldstein (1994) who find earlier school membership to be more important. However, we have already expressed our concerns as to the reliability of the estimated variance components in these studies given the small samples used. In summary, even after adjusting for secondary schools, primary schools and neighbourhoods explain a significant, albeit small relative to secondary schools, proportion of variation in GCSE achievement.

4.2 *Models of educational progress during secondary schooling*

Next we present the results of model E, an extension of model D that adjusts for pupils' prior achievement and background characteristics and includes random coefficients. In choosing the model specification, several models were compared with different predictors and random coefficients. Model E retains those that are statistically significant and those of substantive interest.

Fixed part of model

Table 4.3 presents the fixed part parameter estimates for model E. The model includes a composite measure of prior achievement that summarises multiple prior achievement scores along a single dimension. We choose to use a composite measure to simplify the interpretation and presentation of the analysis, particularly when we describe the random part of the model. The composite prior achievement measure is derived as the estimated linear fixed part prediction from an auxiliary model (not shown) of GCSE score on the multiple prior achievement scores (Goldstein et al., 2000; Yang et al., 1999). These scores are: age 11 English, maths and science scores, each entered as a third-order polynomial and with interactions between the linear terms. We transform this composite score to a normal score so it is scaled in standard deviation units. In model E, the effect of composite prior achievement is very strong with a one standard deviation increase associated with approximately 0.7 of a standard deviation increase in the GCSE score. The presence of the composite prior achievement measure effectively changes the interpretation of all subsequent variables in the model from explaining variation in achievement at GCSE to explaining variation in progress made over secondary schooling. Girls and younger pupils make greater progress than boys and older pupils. Those eligible for FSM and those with SEN make almost 0.3 of a standard deviation less progress whilst those speaking English as an additional language make 0.2 more progress. Asian, Chinese and other ethnic groups make considerably more progress than White and Black pupils.

Importantly, model E also includes indicators of pupil movements between schools and between homes. School moves are split into two types: compulsory and non-compulsory. A single indicator is entered for ever making a compulsory school move while four indicators of non-compulsory school moves (one for each possible move: during academic years 8, 9, 10 and 11) and four indicators of moving home are added to the model. Four interaction terms for moving school and home at the same time are also added to the specification. The mobility indicators are jointly significant giving a large improvement in the DIC of 1230 points. Before interpreting these results, it is worth stressing that the parameter estimates of these indicators should not be interpreted causally since they are likely to additionally reflect systematic differences in the unobservable characteristics of mobile and stable pupils which themselves may be important determinants of progress. For example, mobile pupils may have unobserved characteristics that lead to poorer progress irrespective of moving and, in this case, the reported associations will overstate any genuine negative causal effect of mobility.

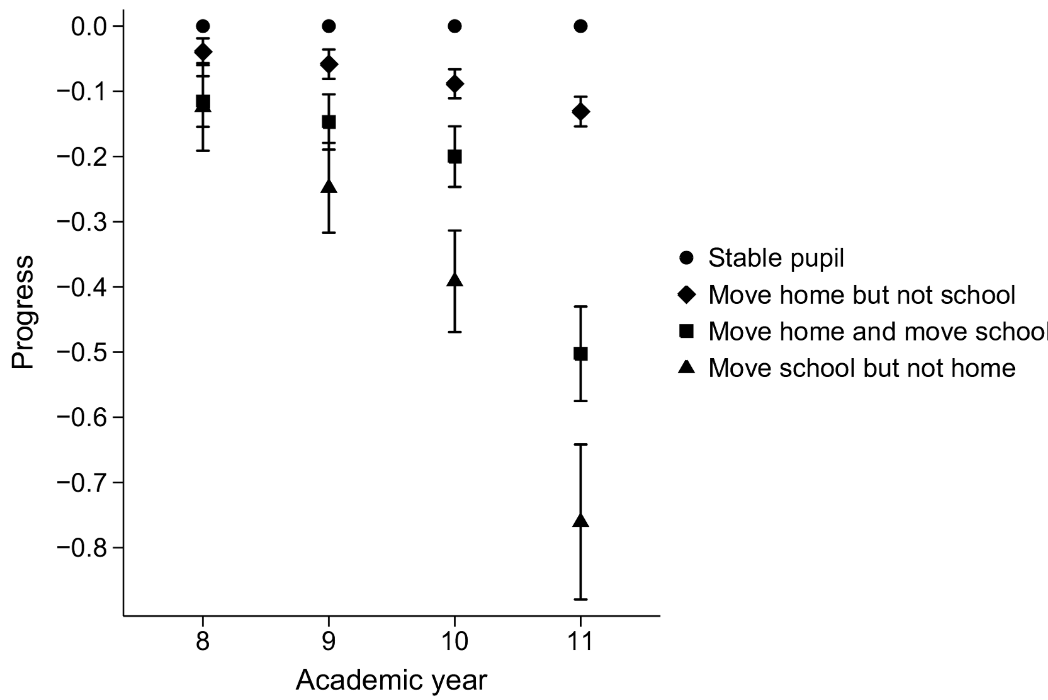
Table 3. Fixed part parameter estimates for model E

Variable	Model E	
Constant	-0.030	(0.012)
Composite prior achievement	0.722	(0.007)
Composite prior achievement squared	0.025	(0.003)
Composite prior achievement cubed	-0.017	(0.001)
Female	0.132	(0.007)
Age within cohort (i)	-0.011	(0.001)
Free school meal (FSM)	-0.279	(0.015)
Special educational needs (SEN)	-0.243	(0.018)
English as an additional language (EAL)	0.210	(0.037)
Ethnicity (ref. White)		
Asian	0.214	(0.043)
Black	0.034	(0.044)
Chinese	0.315	(0.073)
Mixed ethnic group	-0.006	(0.024)
Other ethnic group	0.288	(0.072)
Made a compulsory school change 1+ times	0.010	(0.037)
Changed schools (non-compulsory year 8)	-0.134	(0.047)
Changed schools (non-compulsory year 9)	-0.250	(0.047)
Changed schools (non-compulsory year 10)	-0.398	(0.050)
Changed schools (non-compulsory year 11)	-0.759	(0.066)
Moved home (year 8)	-0.033	(0.010)
Moved home (year 9)	-0.050	(0.011)
Moved home (year 10)	-0.079	(0.011)
Moved home (year 11)	-0.114	(0.011)
Changed schools (non-compulsory year 8) * Moved home (year 8)	0.037	(0.040)
Changed schools (non-compulsory year 9) * Moved home (year 9)	0.163	(0.042)
Changed schools (non-compulsory year 10) * Moved home (year 10)	0.290	(0.046)
Changed schools (non-compulsory year 11) * Moved home (year 11)	0.425	(0.070)
Mean composite prior achievement in pupil's secondary school	0.072	(0.012)
Proportion of FSM pupils in pupil's secondary school	0.010	(0.200)
School type (reference is a comprehensive school)		
Grammar school	0.225	(0.042)
Secondary modern school	-0.010	(0.048)
Neighbourhood social deprivation (IMD)	-0.082	(0.004)
Rural neighbourhood	0.084	(0.011)
Mean composite prior achievement in pupil's primary school	-0.059	(0.005)

Note: Standard errors in parentheses. (i) The age within cohort variable ranges in values from -6 to +6 where -6 corresponds to the youngest pupil in the academic year (born on 31st August) and +6 corresponds to the oldest pupil in the academic year (born on 1st September). A one unit change in the age variable corresponds to an age difference of one month. MCMC estimation used a burn in of 500 and a chain length of 5000. Convergence was judged using the MCMC diagnostics in MLwiN, for example, the ESS and the Brooks Draper statistics (see Browne, 2003 for details). Bayesian DIC = 73268.

Pupils who make compulsory school moves make similar progress to pupils who remain in the same school throughout their secondary schooling. However, pupils who change schools when they do not have to, make significantly less progress than stable pupils. Fig. 2 plots how the strength of the negative association varies across the timing of moves for the different types of mobile pupils. The figure shows the negative association between mobility and progress becomes more negative the closer the moves are to the GCSE exams (academic year 11). The negative association is always strongest for pupils who change schools without moving home, followed by those that move home and school at the same time. The association is weakest for those moving home but not school. A range of additional mobility variables (not shown) were also considered and led to interesting results. First, the negative association between moving school (or home) and progress is found to strengthen with the number of moves made, although relatively few pupils move more than once. Second, the exact date on which pupils change schools shows a much stronger negative association for pupils who move during the academic year compared to those who move during the summer holidays. Third, pupils who move to 'worse' neighbourhoods fare less well than those who move to 'better' neighbourhoods as defined by IMD. Finally, those pupils who migrate into the South West make relatively more progress than those pupils who move home within the South West.

Fig. 2. Negative association of mobility at different stages of secondary schooling for different types of mobile pupils.



Note: Point estimates are plotted with 95% confidence intervals.

Model E also enters school and neighbourhood contextual variables as weighted fixed effects. The addition of these weighted variables leads to a DIC that improves by 43 points compared to when these variables are simply based on the final school and neighbourhood attended. The effect of mean composite prior achievement in secondary schools is small and positive, suggesting pupils make slightly more progress when exposed to higher achieving peer groups. Pupils who attend grammar schools (5% of pupils) make substantially more progress than those who attend secondary modern (3% of pupils) or comprehensive schools (92% of pupils). A one standard deviation increase in neighbourhood social deprivation (i.e. as defined by the IMD normal score) is associated with a significant 0.082 standard deviation drop in progress. In an exploratory analysis, no non-linearities were found in this relationship. At the primary school level, we enter the average age 11 achievement and find pupils' subsequent progress decreases as the performance of their primary school increases. A possible interpretation of this result is that pupils' prior achievement scores are worth more when obtained in low achieving schools rather than in high achieving ones. For example, if two pupils from the same secondary school have the same prior achievement, but one pupil had attended a low achieving primary school, the other a high achieving one, then the model predicts the former to score significantly higher than the latter at GCSE. The inclusion of further contextual variables was explored but few were statistically significant or of substantive interest so are not presented here.

Random part of model

Table 4 presents the random part parameter estimates for model E. The coefficients of composite prior achievement, gender, FSM and SEN are allowed to vary across secondary schools to examine whether schools are 'differentially effective' for different types of pupils. We do not add random neighbourhood level coefficients since, given the low residual variation at this level, where differential effects are found they tend to be very small. This specification of the random part improves the DIC of the model by 492 points compared to the same model with only random intercepts (not shown).

Table 4. Random part parameter estimates for model E

Secondary school classification	Intercept	Composite prior achievement	Composite prior achievement squared	Female	FSM	SEN
Intercept	0.0201	0.06	0.06	-0.49	0.01	0.14
Composite prior achievement	0.0006	0.0046	-0.32	-0.09	-0.64	0.32
Composite prior achievement squared	-0.0017	-0.0005	0.0005	0.19	0.01	-0.66
Female	-0.0047	-0.0004	0.0003	0.0046	-0.07	-0.02
Free school meal (FSM)	0.0001	-0.0053	0.00002	-0.0006	0.0156	-0.42
Special educational needs (SEN)	0.0035	0.0038	-0.0026	-0.0002	-0.0091	0.0311
Neighbourhood variance	0.0032					
Primary school variance	0.0260					
Pupil variance	0.3114					

Note: Variances and covariances for each classification with correlation coefficients in the upper triangle. The reference pupil is a boy with average prior achievement, not eligible for FSM and without SEN.

Table 4 shows estimates of the variances and covariances associated with each random coefficient in model E. The differences in progress between FSM and non-FSM pupils vary substantially between secondary schools: these differences have a variance of 0.0156 (and therefore a standard deviation of 0.125) around an average of -0.279 (see Table 3). So in some schools the difference is as large as -0.529 ($= -0.279 + 2 \times 0.125$) points and in others as small as -0.029 ($= -0.279 - 2 \times 0.125$) points. Hence, relative to the average school, some schools can be seen as narrowing the gap between FSM and non-FSM pupils and some widening it. The gender differential in progress has a smaller standard deviation of 0.071 about a mean of 0.132, implying that girls do better than boys in practically all schools. The SEN difference has a very large standard deviation of 0.167 about a mean of -0.243 implying that there are a few schools where SEN pupils actually make more progress than non-SEN pupils. The correlations reported in Table 4 are also of substantive interest. For example, the negative correlation between composite prior achievement and eligibility for FSM (-0.64) indicates that pupils eligible for FSM under perform more in schools where there is a strong link between prior and current achievement. Interestingly, the negative correlation of -0.42 between the FSM and SEN differences suggests that schools with few differences between FSM and non-FSM pupils tend to have relatively large differences for SEN and non-SEN pupils and vice versa.

School effects for different types of pupils can be evaluated by calculating linear combinations of the 6 school residuals (1 random intercept and 5 random slopes). We can then investigate the extent to which schools are differentially effective for pupils with different characteristics. For example, the correlation between school effects calculated for low achieving FSM boys (with prior achievement one standard deviation below average) and high achieving non-FSM girls (prior achievement one standard deviation above average) is just 0.22. So knowing which schools are effective for low achieving FSM boys is only slightly informative about which schools are effective for high achieving non-FSM girls. Comparing school effects for more extreme groups of pupils leads to even weaker correlations. Clearly, the effectiveness of schools varies greatly for different types of pupil and should not be summarised in a single overall measure. Attempting to do so will lead to misleading inferences about schools.

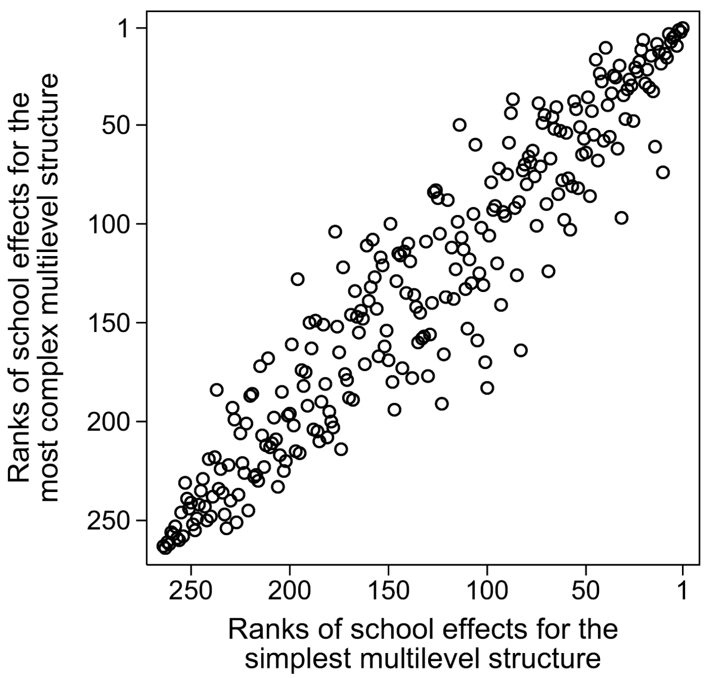
Random coefficients allow the secondary school variance (and therefore the VPC) to be a function of the predictors. At the secondary school level, the reference pupil is a boy with average prior achievement, not eligible for FSM and without SEN. For this pupil, the between school variance is 0.0201 which is smaller than the primary school variance of 0.0260. It is worth noting that primary schools are now more variable than secondary schools and are also more variable in the random intercepts version of model E. This suggests that the predictor variables explain considerably more of the initial differences between secondary schools than between primary schools. The importance of schools attended in earlier phases of schooling has been reported before in the literature (Browne et al., 2001; Goldstein et al., 2007; Goldstein and Sammons, 1997; Rasbash and Goldstein, 1994). Further analysis (not shown) suggests a pattern of greater variation between secondary schools for pupils with more extreme prior achievement, especially high achievement. This suggests that the effect of secondary schools is greatest for pupils with high intake achievements.

4.3 *Stability of estimated school effects across different non-hierarchical data structures*

Finally we examine the stability of the estimated secondary school effects from model E across the alternative data structures depicted in Fig. 1. This informs whether ignoring known complexities of the data structure leads to misleading inferences about the effectiveness of schools. Since we have random coefficient models, we can investigate whether the impact of different multilevel structures on the rank order of school effects differs for different types of pupils. For example, it may be the case that allowing for multiple membership impacts more on school effects evaluated for low achieving pupils since mobile pupils tend to have lower prior achievement than stable pupils.

In model E, the six sets of secondary school residuals are combined to compute school effects for different types of pupils and in an exploratory analysis we do this for a wide range of pupil types. For each type of pupil, the estimated school effects are highly correlated (0.94 – 0.98) across the four alternative data structures. Interestingly, the strength and patterns of these correlations appear not to vary systematically across pupil type suggesting that accounting for different complex data structures does not matter more for certain types of pupils. Allowing for primary school weakens the correlation to a greater extent than allowing for neighbourhoods or pupil mobility. The lowest correlation (0.94) always occurs when comparing the simplest (Fig. 1a) and most complex data structures (Fig. 1d). Fig. 3, a scatter plot of school effects that correspond to this correlation, reveals that there are actually substantial differences in the rank order of schools between these two data structures; approximately half the schools differ by 15 or more places. However, the inherently imprecise nature of estimating school effects (due to the small numbers of pupils within schools) will prevent many of these changes in ranks from being statistically significant (Goldstein and Spiegelhalter, 1996).

Fig. 3. Scatter plot of the rank of school effects for 'average' pupils for the models with the simplest and most complex multilevel structure.



5. Conclusions

Traditional studies of school differences in educational achievement use multilevel modelling techniques to take into account the nesting of pupils within schools. However, educational data are known to have more complex non-hierarchical structures. Neighbourhoods and the schools attended in earlier phases of education may also explain variation in pupils test scores, as may movements between schools and between neighbourhoods over time. Using GCSE data from the English national pupil database, this paper models these complexities by combining multiple membership and cross-classified multilevel models.

We find neighbourhoods and primary schools explain a significant, although small relative to secondary schools, proportion of the variation in pupils' GCSE achievement. When we explicitly model pupil mobility through multiple membership models, we correct for a downwards bias in the estimates of the secondary school and neighbourhood variances that would otherwise lead us to underestimate their importance. After adjusting for prior achievement and other pupil, school and neighbourhood characteristics, we find that pupil mobility continues to have a strong negative association with progress. This overall result has been reported before, but has not been explored for subgroups of movers. We find pupils who change school close to the GCSE exams, especially those who do not simultaneously move home, make particularly low progress. Those that move multiple times, during term time or to more deprived neighbourhoods also make significantly less progress. Interestingly, primary schools now appear as important as secondary schools in terms of the remaining unexplained progress, suggesting schools continue to have an effect on pupils long after they have left them. An essential part of our model is the inclusion of random school level coefficients which show strong differential effects for prior achievement, FSM and SEN. These results strongly suggest that attempting to summarise school effectiveness in a single overall measure will lead to misleading inferences about schools. When we account for the multiple membership and cross-classification structures, we obtain a different ordering of schools effects to that produced by the traditional two-level value-added model; half of schools move 15 or more places. However, it is important to realise that the inherently imprecise nature of estimating school effects will prevent many of these changes from being statistically interesting given the wide confidence intervals for the school effect estimates. The methodology applied in this work is relevant to other contexts in which the data have cross-classified and multiple membership structures, whilst the results demonstrate many of the issues that arise when attempting to account for such complexities.

Acknowledgements

I would like to thank Harvey Goldstein and Simon Burgess for their advice on the paper. Jon Rasbash, Fiona Steele and Stephanie von Hinke Kessler Scholder provided useful comments on an earlier version. The helpful comments from two referees and the journal editors are also gratefully acknowledged. This work was supported by the Economic and Social Research Council (PTA-042-2005-00012).

References

Aitkin, M. and Longford, N. (1986) Statistical modelling issues in school effectiveness studies. *Journal of the Royal Statistical Society: Series A,* 149, 1-43.

ALG (2005) Breaking point: examining the disruption caused by pupil mobility. Association of London Government. (http://www.londoncouncils.gov.uk/London%20Councils/Local%20Government%20Finance/Local%20Government%20Finance%20Publications/ALGmobilityreportfina.pdf)

Browne, W. J. (2003) *MCMC estimation in MLwiN,* London, Institute of Education.

Browne, W. J., Goldstein, H. and Rasbash, J. (2001) Multiple membership multiple classification (MMMC) models. *Statistical Modelling,* 1, 103-124.

DES (1991) The Parents Charter. London, Department of Education and Science

Fielding, A. (2002) Teaching Groups as Foci for Evaluating Performance in Cost-Effectiveness of GCE Advanced Level Provision: Some Practical Methodological Innovations. *School Effectiveness and School Improvement,* 13, 225-246.

Fielding, A. and Goldstein, H. (2006) Cross-classified and Multiple Membership Structures in Multilevel Models: An Introduction and Review. Research Report RR791. London, Department for Education and Skills. (http://www.dcsf.gov.uk/research/data/uploadfiles/RR791.pdf)

Fielding, A., Thomas, H., F., S., Browne, W., Leyland, A., Spencer, N. and Davison, I. (2006) Using Cross-Classified Multilevel Models to Improve Estimates of the Determination of Pupil Attainment: A Scoping Study. Research Report for Department for Education and Skills. School of Education, University of Birmingham. (http://www.socscistaff.bham.ac.uk/fielding/Dfes_scoping_report_School_of_Ed_version_August06.pdf)

Fielding, A. and Yang, M. (2005) Generalized linear mixed models for ordered responses in complex multilevel structures: effects beneath the school or college in education. *Journal of the Royal Statistical Society: Series A,* 168, 159-183.

Garner, C. L. and Raudenbush, S. W. (1991) Neighborhood effects on educational attainment: A multilevel analysis. *Sociology of Education,* 64, 251-262.

GLA (2005) Moving Home and Changing School–1: Widening the Analysis of Pupil Mobility. Briefing no. 2005/32. London, Data Management and Analysis Group, Greater London Authority. (www.bristol.ac.uk/Depts/CMPO/PLUG/publications/ewens2.pdf)

Goldstein, H. (1986) Multilevel mixed linear model analysis using iterative generalized least squares. *Biometrika,* 73, 43-56.

Goldstein, H. (2003) *Multilevel statistical models 3rd Edition,* London, Arnold.

Goldstein, H., Browne, W. and Rasbash, J. (2002) Partitioning Variation in Multilevel Models. *Understanding Statistics,* 1, 223-231.

Goldstein, H., Burgess, S. and McConnell, B. (2007) Modelling the effect of pupil mobility on school differences in educational achievement. *Journal of the Royal Statistical Society: Series A,* 170, 941-954.

Goldstein, H., Huiqi, P., Rath, T. and Hill, N. (2000) *The use of value added information in judging school performance,* Perspectives on Education Policy No. 40. London, Institute of Education.

Goldstein, H., Rasbash, J., Yang, M., Woodhouse, G., Pan, H., Nuttall, D. and Thomas, S. (1993) A Multilevel Analysis of School Examination Results. *Oxford Review of Education,* 19, 425-433.

Goldstein, H. and Sammons, P. (1997) The Influence of Secondary and Junior Schools on Sixteen Year Examination Performance: A Cross-classified Multilevel Analysis. *School Effectiveness and School Improvement,* 8, 219-230.

Goldstein, H. and Spiegelhalter, D. J. (1996) League tables and their limitations: statistical issues in comparisons of institutional performance. *Journal of the Royal Statistical Society: Series A,* 159, 385-443.

ODPM (2004) Indices of deprivation 2004. Office of the Deputy Prime Minister. (http://www.communities.gov.uk/archived/general-content/communities/indicesofdeprivation/216309/)

Ofsted (2002) Managing Pupil Mobility. London, Office for Standards in Education. (http://preview.ofsted.gov.uk/assets/65.pdf)

Rasbash, J. and Browne, W. (2001) Modelling non-hierarchical structures. IN LEYLAND, A. H. & GOLDSTEIN, H. (Eds.) *Multilevel Modelling of Health Statistics.* Chichester, Wiley.

Rasbash, J. and Browne, W. J. (2008) Non-Hierarchical Multilevel Models. IN DE LEEUW, J. & MEIJER, E. (Eds.) *Handbook of Multilevel Analysis.* New York, Springer.

Rasbash, J. and Goldstein, H. (1994) Efficient analysis of mixed hierarchical and cross-classified random structures using a multilevel model. *Journal of Educational and Behavioral Statistics,* 19, 337-350.

Rasbash, J., Steele, F., Browne, W. and Prosser, B. (2004) *A User's Guide to MLwiN version 2.0,* London, Institute of Education.

Raudenbush, S. W. (1993) A Crossed Random Effects Model for Unbalanced Data With Applications in Cross-Sectional and Longitudinal Research. *Journal of Educational and Behavioral Statistics,* 18, 321-349.

Raudenbush, S. and Bryk, A. S. (1986) A Hierarchical Model for Studying School Effects. *Sociology of Education,* 59, 1-17.

Raudenbush, S. W. and Bryk, A. S. (2002) *Hierarchical Linear Models: Applications and Data Analysis Methods 2nd Edition*, Sage Publications Inc.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and van der Linde, A. (2002) Bayesian measures of model complexity and fit. *Journal of The Royal Statistical Society: Series B,* 64, 583-639.

Yang, M., Goldstein, H., Rath, T. and Hill, N. (1999) The Use of Assessment Data for School Improvement Purposes. *Oxford Review of Education,* 25, 469-483.