THE CENTRE FOR MARKET AND PUBLIC ORGANISATION

The Centre for Market and Public Organisation (CMPO) is a leading research centre, combining expertise in economics, geography and law. Our objective is to study the intersection between the public and private sectors of the economy, and in particular to understand the right way to organise and deliver public services. The Centre aims to develop research, contribute to the public debate and inform policy-making.

CMPO, now an ESRC Research Centre was established in 1998 with two large grants from The Leverhulme Trust. In 2004 we were awarded ESRC Research Centre status, and CMPO now combines core funding from both the ESRC and the Trust.

Centre for Market and Public Organisation
Bristol Institute of Public Affairs
University of Bristol
2 Priory Road
Bristol BS8 1TX
http://www.bris.ac.uk/Depts/CMPO/

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

Modelling the Effects of Pupil Mobility and Neighbourhood on School Differences in Educational Achievement

George Leckie

January 2008

Working Paper No. 08/189

Published under the title:
The complexity of school and neighbourhood effects and movements of pupils on school differences in models of educational achievement
in
Journal of the Royal Statistical Society:
Series A 172 (3) pp. 537-554(2009)

ISSN 1473-625X

CMPO Working Paper Series No. 08/189

Modelling the Effects of Pupil Mobility and Neighbourhood on School Differences in Educational Achievement

George Leckie[1]

[1] *CMPO, University of Bristol*

January 2008

Abstract
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 (NPD), 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

JEL Classification: I2

Electronic version: http://www.bris.ac.uk/Depts/CMPO/workingpapers/wp189.pdf

Acknowledgements
I would like to thank Harvey Goldstein and Simon Burgess for their advice and comments on the paper. Jon Rasbash, Fiona Steele and Stephanie von Hinke Kessler Scholder provided helpful comments on an earlier version. This paper has benefited from presentation at a Joint Meeting of the RSS Social Statistics and General Applications Sections: Recent Advances in Multilevel Modelling Methodology and Applications. Financial support from the ESRC is gratefully acknowledged. All errors are my own.

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

CMPO is jointly funded by the Leverhulme Trust and the ESRC

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 among pupils is attributable to differences among schools. These models are commonly referred to as 'value-added' or 'school effectiveness' models and are estimated using multilevel models (for early examples see: Aitkin and Longford, 1986; Goldstein et al., 1993). In England, the Department for Children, Schools and Families (DCSF) (formerly DfES) publishes annual measures of school 'effects' from such models in school 'performance tables' with a view to informing parental school choice (See: http://www.dcsf.gov.uk/performancetables/).

Pupil mobility and neighbourhood effects are often discussed as important potential influences on educational achievement (Office for Standards in Education, 2002; Department for Education and Skills 2003; Association of London Government, 2005; Greater London Authority, 2005). However, few value-added studies incorporate these factors into their analysis. Where studies look at the impact of mobility 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), these studies ignore pupil mobility when specifying the contribution of the random effects. Thus, they treat pupils as belonging to only their final schools and ignore the contribution of earlier schools attended. Furthermore, 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 (Garner and Raudenbush, 1992; Fielding, 2006). 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 (Fielding and Goldstein, 2006).

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. Pupils are then nested within cells created by this cross-classification. 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 and 524 neighbourhoods. However, rather than estimating a random effects cross-classified model, they estimate a two-level random effects model of pupil nested within neighbourhoods and then incorporate schools as a series of 16 fixed effects. They find neighbourhood social deprivation has a strong negative association with educational attainment even after adjusting for prior achievement and family background. In a reanalysis, Raudenbush (1993) estimates a full random effects cross-classified model. Inferences for schools, in addition to neighbourhoods, now relate to the population from which these units are drawn, rather than to the sample of units themselves. This allows the variability of pupils' exam scores to be partitioned between both neighbourhoods and schools. It is worth noting that school level variables could now also be included in the model, but this is not pursued in the paper. Interestingly, the study finds neighbourhoods explain up to twice as much variation as schools. However, it should be noted that with just 17 schools, the school component of variation will be imprecisely estimated. More recently, Fielding et al. (2006) use cross-classified models in an analysis of a large scale data set 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 definitions of neighbourhood. Unlike Raudenbush (1993), they find that the importance of neighbourhoods is relatively small when compared to the importance of schools. Their study also reveals a number of dimensions of neighbourhood level deprivation and disadvantage to be significantly negatively associated with pupil progress.

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 (Rasbash and Goldstein, 1994; Goldstein and Sammons, 1997; Browne et al., 2001; Goldstein et al., 2007). Goldstein and Sammons (1997) consider the persistence of primary school effects on pupil progress in secondary schools. They assess the relative importance of the two types of school and find the variance of primary school effects to be greater than that for secondary schools. Similar results have been found for Scottish data from schools in Fife (Rasbash and Goldstein, 1994; Browne et al. 2001) and in Staffordshire in England for the persistence of infant school effects during progress in junior schooling (Goldstein et al, 2007). These papers suggest that primary (infant) schools are more variable because they tend to be more homogenous units than secondary (junior) schools.

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, acknowledging that pupils belong to more than one school. When specifying multiple-membership models, an issue that arises is the relative importance, or weight, that should be attributed to each school attended. An obvious choice is to weight schools by the length of time spent in each one. This is what is typically done in the literature (Browne et al. 2001; Fielding, 2002) and in a sensitivity analysis Goldstein et al. (2007) find this to be near optimal in terms of model fit. In addition to modelling pupils as nested within a cross-classification of junior and infant school, Goldstein et al. (2007) allow pupils to be multiple members of their junior schools. They find that ignoring junior school mobility leads to downward bias in the estimate of the junior school variance. Despite this, accounting for the multiple-membership makes little difference to the rank order of school effects. However, their analysis is limited to random-intercept models, which treat schools as equally effective for all types of pupils. It is not certain whether a similar result would apply in random coefficients models which allow schools to be differentially effective for different types of pupils.

This paper builds upon the work of Goldstein et al. (2007) to present a more detailed investigation of pupil mobility between schools and also between neighbourhoods. The relative importance of secondary schools, neighbourhoods and primary schools on both raw achievement and progress are assessed for a much larger data set than examined in previous studies. The negative association between mobility and progress is decomposed to see how it varies across the types and timings of these 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 as well as random intercept 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

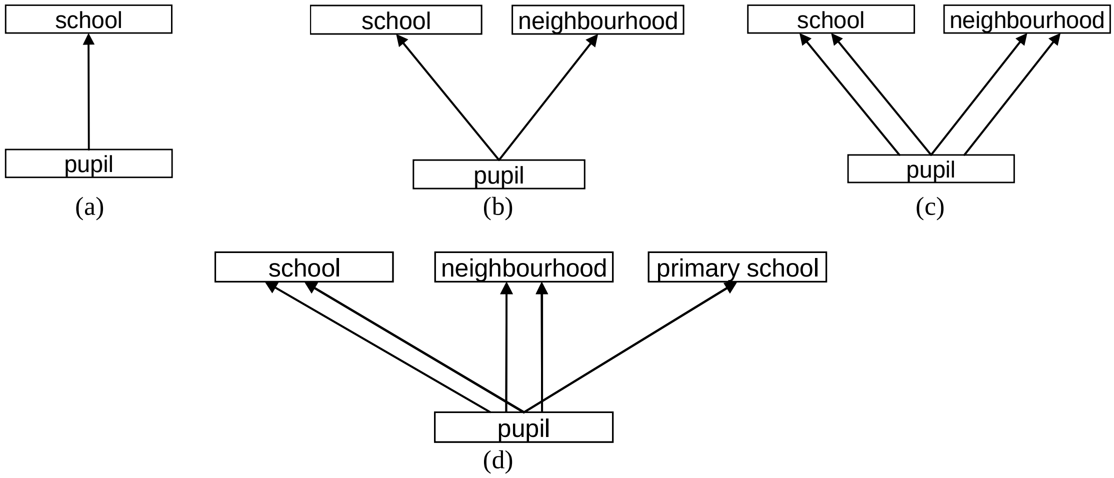
Consider a simple two-level variance components model with an intercept and a single predictor variable. Using the ‘classification’ notation of Browne et al. (2001), this model can be written as

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

There are two classifications: pupils as classification 1 and schools as classification 2. The ‘(2)’ superscript and subscripts identify any variables or random effects that are associated with the school classification. Subscript *i* refers to the *i*th pupil in the data set. The classification function $sec(i)$ denotes the secondary school that pupil *i* attends. y_i is the test score for the *i*th pupil in the data set, x_i is their prior achievement. $u^{(2)}_{sec(i)}$ 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 described above. 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 (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 school



The final classification diagram (Fig. 1d) gives 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

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

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

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

The classification functions $sec(i)$, $nbhood(i)$ and $pri(i)$ refer to the secondary school, neighbourhood and primary school for pupil *i*. 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 (2). For example, although $\sigma^2_{u(2)}$ is the variance of the school effects $u^{(2)}_{sec(i)}$, the actual contribution of schools to the variance for a given pupil is given by:

$$\text{var}\left(\sum_{j \in sec(i)} w^{(2)}_{i,j} u^{(2)}_j\right) = \sigma^2_{u(2)} \sum_{j \in sec(i)} \left(w^{(2)}_{i,j}\right)^2$$

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 $\sigma^2_{u(2)}$ while for children who spend equal time in two schools, the contribution is just $\sigma^2_{u(2)}\left(0.5^2 + 0.5^2\right) = 0.5\sigma^2_{u(2)}$. Indeed, for pupils who attend multiple schools, the contribution is always less than that for stable pupils. The statistical reason for this is that 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 also 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, 2002). There are important implications for models which ignore the multiple-membership of schools and instead assign pupils to the final school they attend. Such models will 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 of pupils are mobile and 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 ($\sigma^2_{u(2)}$) and mobile ($0.5\sigma^2_{u(2)}$) pupils. This estimate is less than the true between school variation of $\sigma^2_{u(2)}$ given by a multiple membership model. It follows that ignoring multiple membership will lead to biased estimates of the school 'effects', the extent of which increases with the degree of pupil mobility (Goldstein, 2003).

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; Goldstein et al., 2007). These weighted fixed effects should 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; however, the model can be extended to incorporate random slopes at one or more of the higher classifications. For example, setting the coefficient of prior achievement to be random at the secondary school classification will allow the effectiveness of schools to vary over the distribution of intake ability. More generally, model (2) could be extended to include, for example, further classifications, discrete responses and multivariate responses (Goldstein et al., 2003, Bryk and Raudenbush, 2003). However, we do not pursue such extensions in this paper.

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 technique are given in Browne (2004).

3. Data and Variables

The exam data are taken from the national pupil database (NPD), a census of all pupils in the English state education system. The NPD holds information on pupils' test score histories and a limited number of pupil level characteristics. From this database, we extract 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 A-levels and are a common type of university entrance requirement. To GCSE scores, we merge data from the 2002-2006 pupil level annual school census (PLASC) data sets which give the series of schools attended and postcodes resided in between the two sets of exams. (Further information on the NPD and PLASC data sets 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 yearly PLASC data sets. The analysis is limited to the 472431 pupils who took their GCSE exams in standard secondary schools that taught for all five years of the secondary phase of education. 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, and is the same measure as that used in published school performance tables. This measure is considered fairer than the uncapped score since it lowers the scores of pupils who score highly simply by taking many examinations. We treat the response as continuous and transform it to a standard normal score so that the multilevel residuals better approximate the normality assumptions of the models (Goldstein, 2003). The mean GCSE point score is equivalent to eight grade C’s whilst a 1 standard deviation difference is equivalent to a 2 grade difference in each of the eight examinations. Pupils who changed schools score on average 0.47 of a standard deviation less 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 transformed to standard normal scores.

The data contain information on a number of pupil level characteristics, which we adjust for in our models. Variables 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. Where pupils have ever been on FSM or SEN, they on average spent 60% of their secondary schooling in these states. 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.

For each of the five years of secondary education, we know the school attended and the postcode where each pupil lives. Using these postcodes, we link in administrative neighbourhood data using the national statistics postcode directory (NSPD). The chosen scale of neighbourhood is the lower super output area (LSOA). LSOAs 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 fits in the analysis. Changes in the school and postcode variables across consecutive time periods are the basis of our measures of pupil mobility between schools and between neighbourhoods. Crucially, we can decompose binary indicators of school and home mobility to provide a fuller picture of the association between mobility and achievement. First, we can identify whether 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). Following Machin et al. (2006), we term these moves “compulsory school moves” while remaining moves are termed “non-compulsory school moves”. Second, we can identify the number of moves pupils make, the timing of these moves and whether pupils simultaneously move home or not.

School level compositional variables - constructed from the pupil level data – are included to capture the influence of pupils’ peer groups. These variables include the average intake achievement and proportion of FSM pupils in each secondary school. A potential difficulty is that pupils who change secondary school are exposed to multiple peer groups and therefore have different values for these variables over time. We therefore form weighted versions of all school level variables since these are expected to better capture the influence of peer groups. At the neighbourhood level we include the index of multiple deprivation (IMD), a scale which incorporates information along seven dimensions of the local community: income, employment, health and disability, education, skills and training, barriers to housing and services, living environment, and crime.

3.2 Description of the non-hierarchical data structure

The sample consists of 42681 children who, at the point of sitting their GCSE exams, attend 264 secondary schools and live in 3175 neighbourhoods. The median secondary school has 161 pupils whilst the median neighbourhood has 14 pupils. At the point of sitting their KS2 exams, these pupils attend 3107 primary schools. 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. Descriptive statistics for these non-hierarchical structures are described below.

3.2.1 Cross-classification between secondary school, neighbourhood and primary school

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; 3 in 10 children would have to change schools to obtain a strict system of neighbourhood schooling. 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 ‘inbalance’ 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).

3.2.2 Multiple-membership of schools and multiple-membership of neighbourhoods

The sample pupils took their GCSE exams in 264 secondary schools located in the South West. However, 8% of pupils changed schools during the period of analysis. Adding in previously attended secondary schools raises the total number of schools in the data set to 1346. Of the extra 1082 schools, 94 are schools in the South-west that taught no pupils at GCSE. Further investigation found these schools to be a combination of middle schools and schools that have closed midway through the period of analysis. The remaining 988 schools are located outside the South West and tend to be the former schools of pupils whose families have moved into the

South West during their secondary schooling. Since our sample only contains those pupils who took their GCSE exanimations in schools located in the South west, the majority of these ‘out of sample’ schools contain a single pupil. This is not problematic for the analysis as we are only interested in making inferences about schools located in the South West. For these schools, we have all the pupils who took GCSE exams and, crucially, we weight these pupils by how long they attend these schools. It is worth noting that, had the sample not been a single contiguous area (i.e. the South West), the proportion of all schools that would be out of sample schools would be considerably higher and may have led to model estimation problems.

Turning attention to pupils’ residential neighbourhoods, we see that over their secondary schooling 27% of pupils 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 additional neighbourhoods are located outside the South West and, where this is the case, they again tend to contain a single pupil.

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	Proportion of secondary schooling spent in:	Secondary schools	Neighbourhoods

of units	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	100.00

4. Results

The first analysis (section 4.1) reports intercept-only 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 raw GCSE achievement. The second analysis (section 4.2) extends the most complex of these models to adjust for pupils' prior achievement and background characteristics and, in doing so, switches the focus of the analysis from educational achievement to progress. Pupil movements between schools and neighbourhoods are examined in detail in the fixed part of these models whilst 'differential effectiveness' is explored in the random part. The third analysis (section 4.3) compares the estimated school 'effects' across a range of multilevel structures. Importantly, this is done in the context of random coefficient models to see whether accounting for the multilevel structure of the data matters more for school effects evaluated for different types of pupils.

4.1 Relative importance of secondary school, neighbourhoods and primary schools in explaining raw GCSE achievement

Table 2 presents results from a series of variance components models with only an intercept term. These models decompose the total variation in the normalised GCSE score into four parts, corresponding to the different classifications in the data: secondary schools, neighbourhoods, primary schools and pupils.

Table 2. Parameter estimates for intercept-only variance components models of normalised GCSE scores. The four models are (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 school

	A		*B*		*C*		*D*	
Variable	*Estimated coefficient*	*Standard error*	*Estimated coefficient*	*Standard error*	*Estimated coefficient*	*Standard error*	*Estimated coefficient*	*Standard error*
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	

Notes: MCMC estimation used a burn in of 500 and a chain length of 5000.

Model A is the standard two-level model, allowing for nesting of pupils within secondary schools (see Fig. 1a). The model produces a school level variance of 0.223 and a pupil level variance of 0.818. This gives an intra-unit, or in this case intra-school, correlation of 0.214 {0.223 / (0.223 + 0.818)}, which is the estimated correlation between the GCSE scores of two randomly selected pupils attending the same school. Interpreted as a variance partition coefficient (VPC, Goldstein, 2002), 21% of the total variation in GCSE scores lies between secondary schools.

Model B treats pupils as nested within a cross-classification of secondary schools and LSOA neighbourhoods (see Fig. 1b). The DIC reduces by 1334 points to 111445 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 had been wrongly attributed to these two levels. The ratio of school-to-neighbourhood variation is approximately four and is consistent with the findings of Fielding et al. (2006), but not Raudenbush (1993) who found, for one educational authority in Scotland, neighbourhoods to explain almost twice as much variation as schools. Despite the greater importance of secondary schools, the neighbourhood variance of 0.054 still leads to sizeable differences in the average GCSE score across neighbourhoods. For example, the difference in average GCSE scores between pupils in a low scoring neighbourhood (neighbourhood effect 1 standard deviation below average) and a high scoring neighbourhood

(neighbourhood effect 1 standard deviation above average) is $2 \times \sqrt{0.054} = 0.465$. This value is equivalent to a one grade improvement in seven out of eight GCSE subjects. A neighbourhood-by-school random interaction effect (Goldstein, 2003; Raudenbush, 2003; Fielding and Goldstein, 2006) is also considered but the within cell sample sizes are not sufficient to separately identify the interaction variance from the pupil variance. A small significant interaction variance is identified when we consider the larger middle super output area (MSOA) scale of neighbourhood (they have a mean population of 7200), but overall this model fits the data less well than the model reported here.

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). Alternative weighting schemes were examined but these led to slightly worse fits of the data. Incorporating the multiple-membership structures into the model leads to a modest decrease in the DIC of 185 points. Larger improvements in the DIC would be expected when modelling populations with greater mobility (e.g. data sets measuring progress during primary schooling). Interestingly, with no modification to the fixed part of the model, the intercept term is now estimated as -0.155 compared to 0.012 in model B. This decrease arises from accounting for the multiple-membership structures in the data, which increase the number of schools and neighbourhoods included in the model. Many of these additional higher level units lie outside the South West and, in our sample, often contain just one or two mobile pupils. Since mobile pupils tend to achieve low GCSE scores, it follows that the mean GCSE scores for these additional units are also low. The reason this decreases the intercept is because the estimate of the intercept 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. Looking at the random part of the model, we see the between school variance increases by 26% over model B whilst the between neighbourhood variance increases by 18%. Schools and neighbourhoods are clearly more important sources of variation than implied in the simpler model. Model C therefore demonstrates how ignoring multiple-membership structures in multilevel models can lead to severely downward biased estimates of higher level variances.

Model D extends model C to include a third cross-classification for primary school (see Fig. 1d). The DIC falls by a further 442 points, suggesting this model structure provides the best fit of the data compared to the simpler structures allowed in models A, B and C. 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. Since the total variance is approximately one, this estimate can be interpreted as the proportion of variation that lies between primary schools and also the approximate correlation between pupils who attend the same primary school, but different secondary schools and neighbourhoods. 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 raw GCSE achievement.

4.2 Multilevel models of pupil progress

In this section we present the results of model E, an extension of model D with explanatory variables and random coefficients. In choosing the model specification, several models were compared with different explanatory variables and random coefficients. The selected model E retains those that are statistically significant and those of substantive interest. The fixed part of model E (Table 3) adjusts for pupil prior achievement, pupil background characteristics, indicators of pupil mobility and the characteristics of pupils' schools and neighbourhoods. The random part of model E (Table 4) has the same multilevel structure as model D, but, by including random coefficients, also allows the effects of important pupil level variables to vary across secondary schools.

Table 3. Fixed parameter estimates for model E

Variable	*Estimated coefficient*	*Standard error*
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	-0.012	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 (reference is 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 attainment in pupil's GCSE school	0.072	0.012
Percent of FSM pupils in pupil's GCSE school	0.0001	0.002
Grammar school	0.225	0.042
Secondary modern	-0.010	0.048
Neighbourhood deprivation	-0.082	0.004
Rural neighbourhood	0.084	0.011
Mean composite prior attainment in pupil's primary school	-0.059	0.005

Notes: MCMC estimation used a burn in of 5000 and a chain length of 50000. Bayesian DIC = 73268.

4.2.1 Fixed part of model

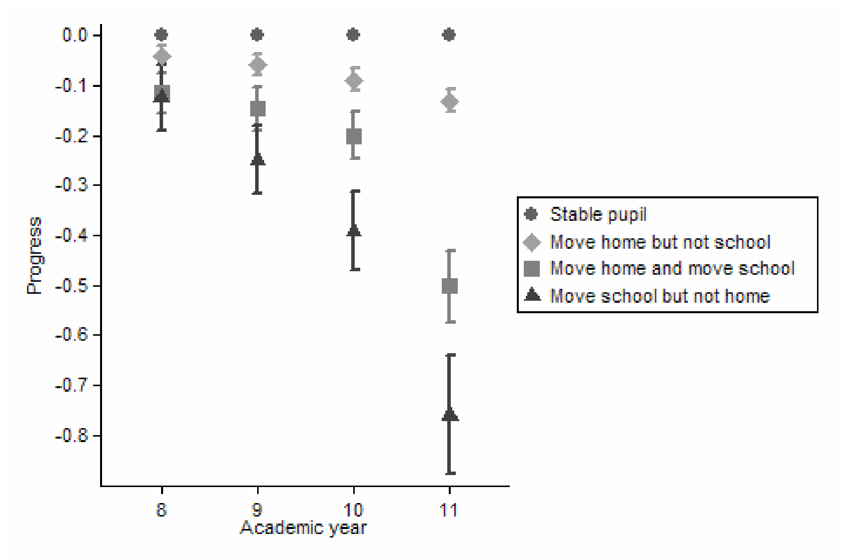
Model E includes a composite measure of prior achievement that summarises multiple prior achievement scores along a single dimension (Yang et al., 1999; Goldstein et al., 2000). We choose to use a composite measure to simplify the interpretation and presentation of the analysis, particularly when we describe the random part of the model. The composite prior achievement measure is derived from the linear fixed part prediction of an auxiliary model (not shown) of GCSE score on age 11 English, maths and science scores. 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.

Model E adds a standard set of pupil background characteristics to the fixed part of the model: gender, age, FSM eligibility, SEN status, EAL and ethnicity. 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. This relative poor performance of white and black pupils has been reported elsewhere in the literature (e.g. Nuttall et al., 1989).

Importantly, model E also includes indicators of pupil movements between schools and between homes. 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. In this case, the reported associations will overstate any genuine negative causal effect of mobility. In model E, 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.

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. To ease the interpretation of the parameters estimates, 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 increases monotonically the closer the moves are to the GCSE exams. 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.

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



Notes. Point estimates are plotted with 95% confidence intervals.

A range of alternative combinations of mobility variables (not shown) are also considered in the fixed part of the model, the findings of which are briefly listed here. First, the negative association between moving school (and/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.

Model E also enters higher level contextual and compositional variables into the fixed part of the model. For each pupil, the secondary school and neighbourhood level predictors are entered as weighted fixed effects to reflect the proportion of time spent in each school and neighbourhood. 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. Secondary school compositional variables include the mean age 11 intake achievement of a school and the percentage of its pupils eligible for FSM. The effect of the former is small and positive whilst the latter is small and negative. This potentially suggests that pupils make slightly more progress when exposed to higher achieving peer groups but less progress when exposed to more impoverished peer groups. Grammar school pupils make substantially more progress than comprehensive pupils and those who attend secondary modern schools. At the neighbourhood level, a 1 standard deviation increase in deprivation is associated with a significant 0.09 standard deviation drop in progress. At the primary school level, we enter the mean age 11 achievement of pupils and find that pupils' subsequent progress decreases as the raw performance of their primary school increases. One possible interpretation of this result is that pupils' prior achievement scores are worth more when obtained in low achieving schools compared to high achieving ones. The inclusion of further higher level variables, both compositional and contextual, were explored but few were statistically significant or of substantive interest so are not presented here.

4.2.2 Random part of model

Since it is well established in the literature that value-added models should study the effectiveness of sub-groups within schools as well as the overall effectiveness of schools (Nuttall et al., 1989), we incorporate random coefficients to examine whether schools are 'differentially effective' for different types of pupils. The coefficients of composite prior achievement, gender, FSM and SEN are allowed to vary across schools. Random coefficients on both the linear and squared terms of composite prior achievement are required to adequately model the differential effectiveness of schools across the intake ability distribution. A random school level coefficient on IMD was considered, but there was no evidence that the neighbourhood deprivation effect varies over schools. We note that random neighbourhood level coefficients could also be added to the model. However, given the low overall residual variation at this level, where differential effects are found they tend to be very small.

Table 4 presents the random parameter estimates for model E and this specification improves the DIC of the model by 492 points compared to the same model with only random intercepts (not shown). It is worth noting that in the random intercepts model, the total unexplained variation is just 0.361, compared to a variance of 1 for the normalised GCSE score. This illustrates the substantial explanatory power of the predictors (especially composite prior achievement). In addition, the VPCs for that model are: 0.04, 0.01, 0.07 and 0.88 for secondary school, neighbourhood, primary school and pupil respectively. Interestingly, there is now more unexplained variation between primary than secondary schools whilst almost no variation lies between neighbourhoods. We shall return to these points later.

Table 4. Random parameter estimates for model E

Secondary school classification	Intercept	Composite prior achievement	Composite prior achievement squared	Female	FSM	SEN
Intercept	0.0201					
Composite prior achievement	0.0006 (0.06)	0.0046				
Composite prior achievement squared	-0.0017 (-0.54)	-0.0005 (-0.32)	0.0005			
Female	-0.0047 (-0.49)	-0.0004 (-0.09)	0.0003 (0.19)	0.0046		
Free school meals (FSM)	0.0001 (0.01)	-0.0053 (-0.64)	0.00002 (0.01)	-0.0006 (-0.07)	0.0156	
Special educational needs (SEN)	0.0035 (0.14)	0.0038 (0.32)	-0.0026 (-0.66)	-0.0002 (-0.02)	-0.0091 (-0.42)	0.0311
Neighbourhood variance	0.0032					
Primary school variance	0.0260					
Pupil variance	0.3114					

Notes: Variances and covariances for each classification with correlations in parentheses. 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 classification in model E. The differences in progress for 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 points (-0.279 + 2*0.125) and others as small as -0.029 points (-0.279 - 2*0.125). 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. Although not explored in this paper, it would be interesting to study to what extent the higher progress of SEN pupils in these schools is due to above average performance of SEN pupils or to a below average performance of 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). By correlating different sets of these school effects, 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 1 standard deviation below average) and high achieving non-FSM girls (prior achievement 1 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 tends to lead 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. The importance of schools attended in earlier phases of schooling has been reported before in the literature (Rasbash and Goldstein, 1994; Goldstein and Sammons, 1997; Browne et al., 2001; Goldstein et al., 2007). 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.

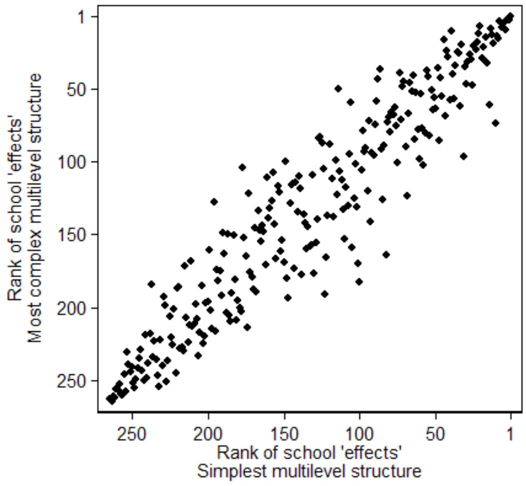
4.3 Stability of estimated school 'effects' across different multilevel structures

Finally we examine the stability of the estimated secondary school 'effects' from model E across the alternative multilevel structures depicted in Fig. 1. This will inform whether ignoring the known complexities of the data structure leads to misleading inferences about the effectiveness of schools. Goldstein et al. (2007) do this for random intercept models of pupil progress that evaluate the overall effectiveness of schools. They compare junior school level residuals from their junior by infant cross-classified, multiple-membership model with junior school level residuals from a simple two-level model and find a high correlation of 0.98. Here we have random coefficient models that measure the effectiveness of schools for different types of pupils. This allows us to 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 random school level intercept and five random school level coefficients give rise to six sets of school residuals. These 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 multilevel structures. Allowing for primary school weakens the correlation to a greater extent than allowing for neighbourhoods or pupil mobility and the lowest correlation (0.94) always occurs when comparing the simplest (Fig. 1a) and most complex structures (Fig. 1d). Interestingly, the strength and patterns of these correlations appear not to vary systematically across pupil type. This suggests that accounting for different multilevel data structures does not matter more for certain types of pupils.

Fig. 4 shows that the correlation of 0.94 between school effects from the simplest and most complex structures actually hides substantial changes in the rank order of schools. Fig. 4 is drawn for male pupils with average prior achievement (i.e. the reference pupil in the school level random effects variance-covariance matrix) and shows the rank order of half the schools changing 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. 4. Scatter plot of the rank of school effects for ‘average’ pupils from 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 structure. 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. Our conclusions are summarised below.

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 large 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 and higher level 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 and/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.

References

Aitkin, M. & Longford, N. (1986) Statistical modelling issues in school effectiveness, Journal of the Royal Statistical Society A, 149, pp. 1-42.

Association of London Government (2005) Breaking Point, Examining the disruption caused by pupil mobility

Browne, W. J. (2004) MCMC Estimation in MLwiN, version 2.0. London: Institute of Education.

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

Bryk, A. S. and Raudenbush, S. W. (2002) Hierarchical Linear Models. Newbury Park: Sage.

Department for Education and Skills (2003) Managing Pupil Mobility: Guidance, Nottingham: DfES.

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 (2): 225-246

Fielding, A and Goldstein, H. (2006) Cross-classified and Multiple-membership Structures in Multilevel Models: An Introduction and Review, DFES report. http://www.dfes.gov.uk/research/data/uploadfiles/RR791.pdf

Fielding, A., Thomas, H., .Steele. F., Browne W., Leyland A., Spencer, N., Davison, I. (2006) Using Cross-classified multilevel models to improve estimates of the determination of pupil attainment: a scoping study.

Garner, C. and Raudenbush, S. W. (1991) Neighbourhood effects on educational attainment, Sociology, of Educational.

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

Goldstein, H. (2003) Multilevel Statistical Models, 3rd edn. London: Arnold.

Goldstein, H., Huiqi, P., Rath, T. and Hill, N. (2000) The Use of Value-added Information in Judging School Performance. London: Institute of Education.

Goldstein H, Browne W, & Rasbash J. (2002) Partitioning Variation in Multilevel Models. Understanding Statistics, 1(4), 223-231.

Goldstein, H., Burgess, S. and McConnell, B. (2007) Modelling the effect of pupil mobility on school differences in educational achievement. J. R. Statist. Soc. A, 170: 941-954

Goldstein, H., Rasbash, J., Yang, M., Woodhouse, G., Pan, H., Nuttall, D. and Thomas, S. (1993) A multilevel analysis of school exam results. Oxford Review of Education., 19, 425–433.

Goldstein, H. and Sammons, P. (1997) The influence of secondary and junior schools on sixteen year exam performance: a cross-classified multilevel analysis. School Effect. School Improvement, 8, 219–230.

Goldstein, H. and Spiegelhalter, D. (1996) League tables and their limitations: statistical issues in comparisons of educational performance (with discussion). Journal of the Royal Statistical Society, A. 159: 385-409.

Greater London Authority (2005) Moving Home and Changing School – 1: Widening the Analysis of Pupil Mobility, Data Management and Analysis Briefing no. 2005/32, London: GLA.

Machin, S., Telhaj, S. and Wilson, J. (2006) The Mobility of English School Children. Fiscal Studies, 27, 253-80

Nuttall, D., Goldstein, H., Prosser, R. and Rasbash, J. (1989) Differential School Effectiveness, Chapter 6 in International Journal of Educational Research, special issue Developments in School Effectiveness Research, 13, 769-776.

Office for Standards in Education (2002), Managing Pupil Mobility, London: Ofsted.

Rasbash, J. and Goldstein, H. (1994) Efficient Analysis of Mixed Hierarchical and Cross-Classified Random Structures Using a Multilevel Model. Journal of Educational and Behavioural Statistics, 19 (4): 337-350

Rasbash, J., Steele, F. and Browne, W. (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 Statistics, Vol. 18, No. 4., pp. 321-349.

Spiegelhalter, D. J., Best, N.G., Carlin, B. P. and van der Linde, A. (2002) Bayesian measures of model complexity and fit (with discussion). J. R. Statistical. Soc. 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 (4): 469-483.