Analysing Differential School Effectiveness Through Multilevel and Agent-Based Modelling

Elio Marchione, Mauricio Salgado and Nigel Gilbert

CRESS: Centre for Research in Social Simulation, Department of Sociology, University of Surrey, Guildford, Surrey, GU2 7XH, United Kingdom,
email: {e.marchione, m.salgado, n.gilbert}@surrey.ac.uk

Abstract. Multilevel Models (MLM) have pioneered the analysis of hierarchical data of two or more levels. Agent-Based Models (ABM) are also used to analyse social phenomena in which there are two or more levels involved. This paper addresses a comparison between MLM and ABM. To provide a basis of comparison, we focus on differential school effectiveness analysis, where MLM has been well studied, using data from the *London Educational Authority's Junior Project*. A MLM is fitted and an ABM of pupils' educational attainment using a social network structure is built. The paper reports the results of both models and compares their performances in terms of predictive and explanatory power. Although the fitted MLM outperforms the proposed ABM, the latter still offers a reasonable fit and provides a causal mechanism to explain differences in school performance that is absent in the MLM.

Key words: Agent-based modelling, differential school effectiveness, multilevel modelling, social simulation.

1 Introduction

During the last thirty years education researchers have developed models for judging the comparative performance of schools, in studies of what has become known as *differential school effectiveness* [1, 2]. These variable-based models, which have achieved great sophistication, allow the researchers to identify the extent to which schools improve pupils' educational attainment. Among those models, Multilevel Models (MLM) are very popular, since they allow the analysis of data that have a hierarchical structure, with two or more 'levels' (e.g., pupils and schools) [3]. However, despite their sophistication, variable-based models do not provide causal explanations for the observed social phenomenon [4]. Whether a MLM (or any statistical model) is able to identify causal effects depends largely on the availability of longitudinal data on likely causal influences, such as teacher-pupil and pupil-pupil interactions, and detailed knowledge about the underlying processes [5]. Of course, such information, which would allow researchers to formulate more complex models (closer to causal models), is rarely available. In the absence of such data, MLM are still well-suited to identify differences among groups, but they cannot explain why those differences might emerge in the first place, since they do not uncover the generative mechanisms that bring them about. When researchers want to understand why some social phenomenon emerges but they do not have access to precise longitudinal data, agent-based modelling (ABM) might be the best alternative. ABM is a computational method to experiment with models composed of autonomous agents that interact within an environment [6]. For instance, researchers might use ABM to explain differential school effectiveness by focusing on the dynamics of the social networks that shape and are shaped by pupils' interactions within and outside school. Whilst ABM is explanatory, MLM is a sophisticated way of description and hypotheses testing. The comparison and integration of multivariate analysis, such as MLM, and the modelling of generative mechanisms, such as ABM, is an important methodological issue.

This paper explores that possibility by formalising an ABM to explain differences in school effectiveness. It describes an ABM to understand the effects of pupils' interactions in educational attainment using a network structure and a methodological strategy to assist with the comparison between MLM and ABM. We begin this paper with a brief account of MLMs in education research (Section 2). Then, we describe the data we are using (Section 3) and we fit a MLM to evaluate possible group effects and the extent to which differential school effectiveness is present in the data (Section 4). Later, we present an ABM to explain differential school effectiveness, describing the model entities, interactions and main dynamics (Section 5). The last part of the paper presents a comparison between the modelling techniques taking into account their predictive power (Section 6). It concludes with some remarks about how the analysis could be extended (Section 7).

2 Multilevel models in education research

In the context of educational research, MLM were developed to adjust simple comparisons of school mean values by using measures of pupils' prior achievement and other variables to take account of selection and other procedures that are associated with pupils' achievement, but not related to any effect that the schools themselves may have on achievement [7, 8]. A simple two-level, random intercept model based on data from a random sample of schools can be written as follows, where subscript i refers to the pupil, and j to the school:

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + u_j + e_{ij}, \qquad u_j \sim N\left(0, \sigma_u^2\right),\ e_{ij} \sim N\left(0, \sigma_e^2\right); \tag{1}$$

where y_{ij} and x_{ij} respectively are the response variable and prior attainment, and u_j is an underlying school effect (which is associated with school organization, teaching, etc.). This model assumes that e_{ij} and u_j are uncorrelated and also uncorrelated with any explanatory variable—i.e. it assumes that any possible dependences that may result from, for example, school selection mechanisms are accounted for. Posterior estimates $\hat{u}_j$ with associated confidence intervals are typically used to rank schools in 'league tables' or used as 'screening devices' in school improvement programmes.

When a MLM is used, it is assumed that the group level makes a difference that explains the total variance of the dependent variable [9]. It is useful to measure how important the group level differences are (i.e., to identify the importance of the 'school effect'), or the proportion of the total variance accounted by the group level. A convenient summary of this effect is the 'interclass-correlation' coefficient (ICC), given by the formula

$$\rho = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2} \tag{2}$$

Model 1 can be elaborated by introducing further covariates such as socio-economic background or peer group characteristics, to make additional adjustments, satisfy the distributional assumptions or investigate interactions. In addition, it is typically found that models such as Model 1 require random coefficients, where, for example, the coefficient of prior achievement varies randomly across schools. In this case, using a more general notation, we have

$$\begin{gathered} y_{ij} = \beta_{0ij} + \beta_{1j} x_{ij}, \\ \beta_{0ij} = \beta_0 + u_{0j} + e_{ij}, \\ \beta_{1j} = \beta_1 + u_{1j}, \\ e_{ij} \sim N\left(0, \sigma_e^2\right), \quad \begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \sim N\left(0, \Omega\right), \Omega = \begin{pmatrix} \sigma_{u0}^2 & \\ \sigma_{u01} & \sigma_{u1}^2 \end{pmatrix}. \end{gathered} \tag{3}$$

The Multilevel Model 3 may also be extended to include further levels of hierarchy, such as education board or authority, and random factors which are not contained within a simple hierarchy, such as area of pupil residence or school attended during a previous phase of education. Such designs are known as 'cross-classification'. [10].

An ABM that hopes to explain school effectiveness should describe a similar pattern, that is, it should reproduce the school effects or differences in the school effectiveness that are in the data as shown by a pattern of high interclass-correlation. The advantage of complementing MLM with a 'bottom-up' approach lies not only in its power to replicate some previous discoveries, but also in the the possibility of testing hypothesised causal mechanisms that might bring about the differences in school effectiveness.

3 Data

We use a subsample from the *The London Education Authority's Junior School Project Data* for pupils' mathematics progress over 3 years from entry to junior school to the end of the third year in junior school [1]. This was a longitudinal study of around 2000 children. Our subsample consists of 887 pupils from 48 schools, with five relevant variables, namely:

- *School ID*, an identification number assigned to each school, from 1 to 48,
- *Social Class*, a variable representing father's occupation, where 'Non Manual Occupation' = 1 and 'Other Occupation' = 0,
- *Gender*, a variable representing pupils' gender, where 'Boy' = 1 and 'Girl' = 0, and

– *Math 3* and *Math 5*, pupil's score in math tests in year 3 and in year 5 respectively, with a range from 0 to 40.

These data enable us to formulate a two-level model (pupils grouped in schools). In order to establish whether a MLM is appropriate, we estimated an *unconditional means model* [11], which does not contain any predictors but includes a random intercept variance term for groups, and which is defined as $Y_{ij} = y_{00} + u_{0j} + r_{ij}$, where the dependent variable is a function of a common intercept y_{00} and two error terms: the between-group error term, u_{0j}, and the within-group error term, r_{ij}. This model is useful since we can get two estimates of variance from it: τ_{00} for how much each group's intercept varies from the overall intercept (y_{00}), and σ^2 for how much each individual score differs from the group mean. An analysis of this model showed that the ICC (see Equation 2) equals 0.119, so an important portion of the variance ($\approx 12\%$) is explained by the pupils' group (i.e., school) membership. Further, the overall group mean reliability test [12] of the outcome variable equals 0.67, although several schools have quite low estimates. In fact, just 22 over 48 schools have group mean reliability over 0.7, which is the conventional value to determine whether groups can be reliably differentiated. Finally, we get from our unconditional means model that the intercept variance τ_{00} is significantly different from zero, $\chi^2(3) = 52.3$, $p < .0001$. Therefore, the analysis shows that fitting a MLM is a sensible decision.

However, given the great heterogeneity in group mean reliability among the schools, subsequent analysis and modelling was confined to those 22 schools that had high estimates in this test, representing 558 pupils. By doing so, we will base our exploratory analysis on data that contains schools that are reliably different one from another.

4 Fitting a Multilevel Model

The multilevel models used for the analysis of the second maths test scores (year 5) were elaborated to take into account relevant background factors and prior attainment (i.e., maths scores in year 3). The MLM were built in the Statistical Software R [13], using the package `nlme`. The parameter estimation was carried out by using the algorithm Log-Cholesky [14]. Models were compared in order to evaluate their overall fit. In Table 1, *Model 0* is a base model, with no predictors but just random intercepts. *Model 1* considers one predictor, previous attainment, and the intercepts of the groups were allowed to vary randomly. *Model 2* adds two background factors for each pupil, gender and social class, to the previous model. Finally, *Model 3* considers previous attainment, background factors and, additionally, the coefficients for previous attainment, which were allowed to vary randomly across the 22 schools. The results shown in Table 1 establish that *Model 3*, which allows random coefficients for previous attainment, has a significantly better fit to the data than *Model 0*, *Model 1* and *Model 2*.

The results obtained from fitting Model (3) are shown in Table 2. The average intercept across all the schools, β_0, equals 12.65 (std. error 1.79) and the average slope for *Math 3* across the 22 schools β_1 equals 0.6 (std. error 0.05). Both parameters are significant. The individual school slopes, u_{1j}, vary around the average slope with a standard deviation estimated as 0.14. The intercepts of the individual schools, u_{0j}, also differ,

Table 1. Comparison of Fitted Models

	df	AIC	BIC	log Lik
Models				
Model 0	3	3858.127	3871.257	−1926.064
Model 1	4	3660.438	3677.945	−1826.219
Model 2	6	3659.913	3686.174	−1823.957
Model 3	8	3657.157	3692.170	−1820.578
Tests	χ^2	p-value		
0 vs 1	199.689	< 0.001		
1 vs 2	4.525	0.104		
2 vs 3	6.757	0.034		

with a standard deviation estimated as 6.04. In addition, there is a negative covariance between intercepts and slopes, σ_{u01}, estimated as -0.98, suggesting that schools with higher intercepts tend to have lower slopes. Finally, the pupils' individual scores vary around their schools' lines by quantities e_{ij}, the level 1 residuals, whose standard deviation is estimated as 5.17.

The two control variables included in the model, gender and social class, perform differently. Only social class (i.e., '*Nonman*' in Table 2) makes a contribution to the model, with an estimated regression coefficient of 1.17 (std. error 0.53, $p < 0.05$). This means that pupils whose father's occupation is non-manual have an expected advantage of 1.17 points in *Math 5* in comparison to those students whose father's occupation is manual. On the other hand, gender (i.e., '*Boy*' in Table 2) does not contribute to the predictive power of the model, since its regression coefficient is not significantly different from zero.

With the information obtained from the MLM, predictions might be made for every pupil in one of the 22 schools. For example, consider a boy student from school 32, whose previous attainment in mathematics at year 3 was 22, and whose father's occupation is classified as manual. From the MLM we know that the group-intercept for this school, $\hat{u}_{0,32}$ is 6.7869 and its group-slope for previous attainment $\hat{u}_{1,32}$ is -0.1418. These values may be incorporated into Equation 3 to obtain the predicted value in *Math 5* for this student as ≈ 29.5.

5 An Agent-Based Model

The ABM we propose addresses the problem of explaining the differences in school effectiveness by taking into account the inputs of knowledge or feedback that every pupil receives from her or his social environment in relation to one specific subject they are supposed to learn. Thus, the model considers the relevant social network in which the pupil is embedded. Furthermore, in order to establish comparisons and possible integrations between this ABM and the MLM explained in Section 4, we empirically calibrated the former using the same data we referred to in Section 3. The ABM was built in NetLogo 4.1.2 [15].

Table 2. Parameters of Random Slope Model for Maths Attainment in Year 5

Parameters (Outcome Variable: Math 5)			
		Random Effects Parameters Estimate	
St. Dev. (σ)	Intercept (u_{0j})	6.04	
	Math 3 (u_{1j})	0.14	
Cov. (σ_{u01})	Math 3*Intercept	-0.98	
	Residual (e_{ij})	5.17	
		Fixed Effects	
		Estimate	Std. Error
Coefficients (β_n)	Intercept (β_0)	12.65^{***}	1.79
	Math 3 (β_1)	0.60^{***}	0.05
	Nonman (β_2)	1.17^{*}	0.53
	Boy (β_3)	-0.02	0.44

Note. $^{***} = p < 0.001$, $^{**} = p < 0.01$, $^{*} = p < 0.05$.

5.1 Theoretical framework

The importance of taking into account the network in which a pupil is embedded in order to explain her or his educational attainment is well established in the literature. Since the observational study carried out by Rist [16] in the seventies, educational researchers have been aware of the impact the student-teacher relationship might have on pupils' learning. Thus, schools where teachers have higher expectations regarding the future of their students might perform better compared to others where teachers have lower expectations [17]. These expectations determine which pupils are defined by the teacher as 'fast learners' and which ones as 'slow learners'. In this way, teachers' behaviour contributes to a 'self-fulfilling prophecy', that is, pupils who are considered 'slow learners' in advance receive less attention and educational feedback, and consequently, they perform worse compared to pupils who are considered 'fast learners'. Equally important are the pupils' characteristics within the classroom, for which the effect on children's educational achievement has also been well documented. Beckerman and Good [18] showed that classrooms in which more than a third of the children were 'high-aptitude' students and less than a third were 'low-aptitude' performed better than those classrooms in which the opposite was true. Their results indicated that both high- and low- aptitude students in the first kind of classroom had greater achievement gains than comparable students in less 'favourable classrooms'. These findings are consistent with the 'peer-effect' hypothesis, something that has been modelled using Social Network Analysis [19] (however, see [20] for disconfirmatory evidence of peer-effect on educational achievement). Finally, the cultural capital that pupils' families possess has an important effect on students' performance [21, 22]. Thus, previous research suggests that we should focus on three dimensions to explain school differential effectiveness:

(a) the educational feedback that pupils receive from their teachers; (b) pupil-pupil interactions and (c) pupils' cultural capital.

The previous three social dimensions of education are the elements we aim to model. To do so, we define a social mechanism to explain how these dimensions are related within the schools. This mechanism gives an account of the way in which pupils interact among them and create groups with other similar to them in those dimensions. We refined Resnick and Wilensky's model [23] to replicate this group formation mechanism. Students form groups with others similar to them (the homophily principle [24]) following group formation rules present at the school level. We assume that these rules are stable and similar for all the individuals within the school [25]. We are not interested in giving an account of the emergence of these rules; we take for granted they exist. In the next section we describe the proposed ABM in detail.

5.2 Model description

The ABM was designed following two basic assumptions. The first concerns the way in which pupils' learning of one specific subject evolves over time. It seems reasonable to assume that this learning can be modelled as a logarithmic function of the educational feedback received on the subject. Thus, there is an initial period of rapid increase, followed by a period where the growth in learning slows (evidence supporting this pattern of learning may be found in [26]).

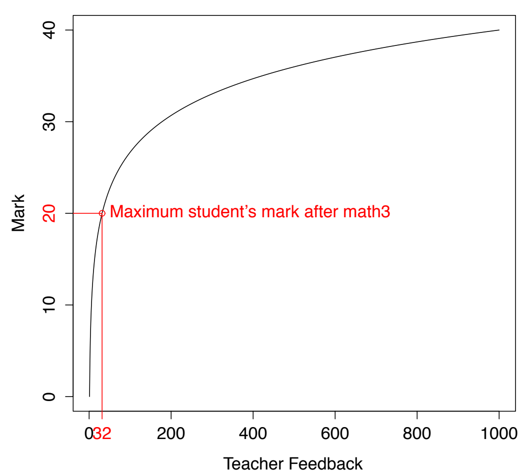

Fig. 1. Simulated Pupils' Learning Curve

In order to model pupils' learning in maths from year 3 to year 5, we define a students' learning curve. Firstly, we assume that learning maths is a continuous process in which the student receives feedback on the subject from the teacher, that is, the amount of teaching time or teachers' reward-directed behaviour towards the pupils. This learning process starts at the first maths lesson, *lesson* 0, and finishes when the knowledge of maths is measured in year 5 (or *Math 5*). Although we do not have any real measurement of the educational feedback involved in this process, it is plausible to know it (Rist measured these elements in his observational study, [16]). Given that such information is absent in our data, we arbitrarily define 1,000 as the amount of feedback that the entire learning process involves. Figure 1 shows the students' learning curve employed in the ABM. Simulated students' marks are worked out as a function of the amount of teachers' feedback that students have undertaken. We also assume that when the test *Math 3* is applied, students have learned half of the topics they were supposed to learn. Further, since both *Math 3* and *Math 5* range between 0 and 40, we transform *Math 3* by dividing it by 2.

Secondly, we assume that the feedback that students receive from their teachers depends on the socialisation processes within their schools. By socialisation we mean all those practices and rules that eventually generate stable groups of students. A group is stable when its members do not want to leave it, that is, they are 'happy' as members of the specific group. Let g_k be a stable group in a school j and s_{ik} a student in such a group. Let $math3_k$ be the average of *Math 3* marks of group g_k, then the amount of feedback that the students in group k receive is given by the following equation:

$$t_k = \left(e^{2 \cdot math3_k}\right)^{\frac{1}{\vartheta}} \tag{4}$$

where ϑ in the exponent allows us to fit a logarithmic function that maps 'Teacher Feedback' into 'Mark' (see Figure 1). Under this condition, $\vartheta \approx 5.790593$; since we know that $log(1,000^{\vartheta}) \approx 40$. Then, the simulated student's score $simMath5_{ik}$ is shown in Equation 5, where $t_{ik} = t_k + t_{math3,ik}$ and $t_{math3,ik}$ is the amount of feedback the pupils in group k have had when their attainment is measured as *Math 3*.

$$simMath5_{ik} = log\left(t_{ik}^{\vartheta}\right) \tag{5}$$

The second assumption is related to the group formation mechanism. There is an initial number of spots where students can hang out. Every school has a threefold tolerance criteria which is adopted by the students to decide whether to stay in a specific group or to move to the next one. Pupils who belong to the same spot establish a group. If they are in a group that has, for example, a higher percentage of people of the opposite sex than the school's tolerance, then they are considered 'uncomfortable", and they leave that group for the next spot. Movement continues until everyone at the school is "comfortable" with their group. The final number of groups might be smaller than the number of spots. Taking into account the available data (see Section 3), we defined three tolerance levels: *Educational tolerance*, that reflects the students' tolerance of accepting others with different attainments in *Math 3*; *Gender tolerance* indicates the students' tolerance for people of the opposite sex; and *Social class tolerance*, the pupils' tolerance for different social class. Tolerance levels range between 0 and 1 and corresponds to the proportion of similar pupils within each group. Figure 2 shows the student network at the end of a simulation for school 32. Male and female pupils are coloured blue and pink respectively; rounded and squared shaped nodes represent low and high social class respectively; and previous attainment in *Math 3* is labelled on students' icons. In this scenario, education, gender and class tolerances are 0.9, 0.3 and 0.9 respectively. There are 39 students in school 32 and these form themselves into 15 groups.

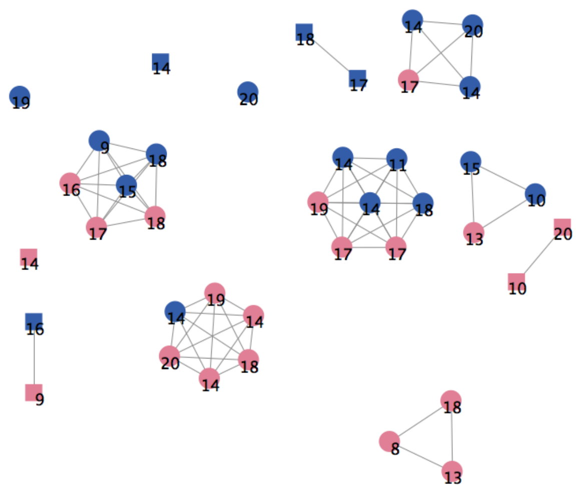

Fig. 2. Simulated Students Social Network in School # 32

5.3 Model Calibration

We initialised the ABM with the pupils' performance in *Math 3* and we explored the parameter space given by the three tolerance levels. Our objective was to find a set of tolerance levels for each school that minimises the differences between the data and the simulations results. Let d_j be such a difference for school j. Then,

$$d_j = \sum_{i=1}^{n_j} |math5_{ij} - simMath5_{ij}| \,/2 \tag{6}$$

where $math5_i$ and $simMath5_i$ are the score in *Math 5* of student i obtained from the real data and from the simulations respectively. In the example shown in Figure 2, $d_{32} = 2.231$, which means that the simulated score in *Math 5* differs, on average, from the data by ± 2.231 units. In order to explore the parameter space of the model, we ran 126,720 simulations. This represents all the possible combinations of the three tolerance levels (varying among 0.3, 0.5, 0.7 and 0.9) and the number of spots (varying among 15, 20 and 25) across the 22 schools. In order to have more robust results, we ran each setting 30 times and then took the average of d_j over all 22 schools as the aggregate outcome.

6 Comparing MLM and ABM

Table 3 shows the results for the parameter setting that minimises d_j. We present the average distance (in the same units as the real data) between the predicted scores and the real scores in *Math 5* for both the multilevel model ('MLM (d_j)') and the simulation ('ABM (d_j)') respectively. The results are grouped according to the 22 schools we included in our study. As well, in this table we show the number of groups ('Final Groups') in which all the pupils were happy with their group membership, given the values in the 'Tolerance Levels' for education, gender and social class (the last three columns of Table 3). Recall that these three last variables were set as simulation parameters, and the specific values presented in the table correspond to those combinations at the school level that minimise the distance between the simulated and the real data scores in *Math 5*.

Table 3. Calibration Results

					Tolerance Levels		
School Id	Num. Pupils	MLM (d_j)	ABM (d_j)	Final Groups	Edu.	Gender	Soc. Class
1	25	2.88	3.36	13	90%	50%	30%
4	24	2.26	3.12	12	90%	90%	50%
5	25	1.53	2.26	12	90%	70%	90%
8	26	1.41	2.82	12	90%	70%	30%
9	21	1.67	2.91	12	90%	70%	30%
11	22	2.21	3.10	12	90%	30%	70%
12	19	3.03	3.55	12	90%	50%	30%
20	28	1.60	2.62	12	90%	30%	70%
22	18	2.18	3.63	10	90%	30%	70%
23	21	1.43	3.19	12	90%	90%	50%
25	20	2.60	3.50	11	90%	30%	50%
26	19	1.85	2.79	12	90%	70%	50%
29	20	2.30	3.36	12	90%	70%	30%
30	35	1.03	2.56	14	70%	90%	70%
31	22	2.30	3.60	12	90%	70%	50%
32	39	1.72	2.71	15	90%	30%	90%
33	25	1.22	3.04	12	90%	30%	90%
35	27	1.01	2.44	13	90%	70%	30%
41	38	2.46	3.25	16	90%	30%	70%
45	30	1.58	2.62	12	90%	30%	70%
46	62	2.24	2.96	15	90%	90%	70%
47	22	1.85	3.61	12	90%	50%	90%

Firstly, by comparing the average between the two models, we see that the predictions of the multilevel model outperform the predictions of the agent-based model, so the former is more accurate. However, the prediction errors of the ABM are not high; in fact, the overall distance equals 3.04 on a scale of 40 points. Thus, the ABM, despite its simplicity, offers a reasonable fit to the data. Secondly, the simulation results suggest a high educational tolerance, since most of values equal 90% (except from school 30, in which the tolerance level equals 70%). On the other hand, the tolerance levels of social class and gender vary across the schools. Therefore, the group formation mechanism in our simulation seems to be ruled by the variables social class and gender, and previous attainment in maths does not constitute a variable that discriminates between groups. Thirdly, the hypothesised mechanism that bring about the differences in school effectiveness, based on social interactions among pupils and group formation according to tolerance levels defined at the school level, seems to be justified. The simulation results indicate that the mechanism of group formation helps to minimise the distance between the predicted and the real scores, allowing a better fit with the data. For instance, when we compare the number of groups with the number of pupils, we can see that in general we have fewer groups than students in each school (for a graphical example, see Figure 2). If the numbers of groups made no difference in the simulation, then the number of groups and the number of pupils would tend to be similar (at least in those schools with 25 or fewer pupils, which is the maximum number of groups the ABM calibration allowed). This is clearly not the case. Therefore, the pupils' social networks seem to be important to explain the differences in effectiveness among schools.

7 Concluding Remarks

In this paper we have presented and compared the results of two models to address differential school effectiveness. The first one is a MLM, where the hierarchical nature of educational processes is considered. The second one is an ABM, where the social mechanisms that might generate school effects in pupil attainments are formalised and explored. We found that the MLM provides more accurate predictions compared to the ABM, However, the differences in the prediction are small and range between 1 and 2 units of marks. The next research step is to reduce this prediction gap between the two models by further refining the ABM. There are important differences between the two modelling techniques. Whereas MLM is data driven, ABM is both data driven and theory based, so the latter allows us to formalise and falsify *in silico* plausible mechanisms that might bring about the observed differences in performance across schools. Finally, whereas MLM takes few seconds, the proposed ABM took several hours for coding, calibration and analyses.

Acknowledgements

This paper is part of an ESRC-NCRM Collaborative Project named *'Multilevel and Agent-Based Modelling: Comparison and Integration'*. We thank Dr. Edmund Chattoe-Brown, Professor Fiona A. Steele and Professor Sylvia Richardson for their helpful review and comments on early versions of this paper.

References

1. Nuttall, D.L., Goldstein, H., Prosser, R., Rasbash, J.: Differential school effectiveness. International Journal of Educational Research **13**(7) (1989) 769–776
2. Sammons, P., Nuttall, D., Cuttance, P.: Differential school effectiveness: Results from a reanalysis of the inner london education authority's junior school project data. British Educational Research Journal **19**(4) (1993) 381–405
3. Paterson, L., Goldstein, H.: New statistical methods for analysing social structures: an introduction to multilevel models. British Educational Research Journal **17**(4) (1991) 387 – 393
4. Hedström, P., Swedberg, R.: Social mechanisms. Acta Sociologica **39**(3) (1996) 281–308
5. Snijders, T.A.B.: Models for longitudinal network data. In Carrington, P.J., Scott, J., Wasserman, S., eds.: Models and Methods in Social Network Analysis. Cambridge University Press, New York (2005) 215–247
6. Gilbert, N.: Agent-Based Models. Sage Publications, Inc (2007)
7. Goldstein, H., Burgess, S., McConnell, B.: Modelling the effect of pupil mobility on school differences in educational achievement. Journal of the Royal Statistical Society: Series A (Statistics in Society) **170**(4) (2007) 941–954
8. Steele, F., Vignoles, A., Jenkins, A.: The effect of school resources on pupil attainment: a multilevel simultaneous equation modelling approach. Journal of the Royal Statistical Society: Series A (Statistics in Society) **170**(3) (2007) 801–824
9. Gelman, A., Hill, J.: Data Analysis Using Regression and Multilevel/Hierarchical Models. 1 edn. Cambridge University Press (2006)
10. Leckie, G.: The complexity of school and neighbourhood effects and movements of pupils on school differences in models of educational achievement. Journal of the Royal Statistical Society: Series A (Statistics in Society) **172**(3) (2009) 537–554
11. Singer, J.D.: Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. Journal of Educational and Behavioral Statistics **23**(4) (1998) 323–355
12. Bliese, P.D.: Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In Klein, K.J., Kozlowski, S.W.J., eds.: Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions. Jossey-Bass, San Francisco (2000) 349–381
13. R Development Core Team: R: A language and environment for statistical computing (2011) http://www.R-project.org.
14. Pinheiro, J.C., Bates, D.M.: Unconstrained parameterizations for Variance-Covariance matrices. Statistics and Computing **6** (1996) 289—296
15. Wilensky, U.: NetLogo. Center for Connected Learning and Computer-Based Modeling. Northwestern University, Evanston, IL. (2011) http://ccl.northwestern.edu/netlogo.
16. Rist, R.: Student social class and teacher expectations: The Self-Fulfilling prophecy in ghetto education. Harvard Educational Review **70**(3) (2000) 257–302
17. de Vos, H.: Using simulation to study school effectiveness. In: The Annual Meeting of the European Council on Educational Research, Bath, England (1995)
18. Beckerman, T.M., Good, T.L.: The classroom ratio of high- and Low-Aptitude students and its effect on achievement. American Educational Research Journal **18**(3) (1981) 317–327
19. Calvó−Armengol, A., Patacchini, E., Zenou, Y.: Peer effects and social networks in education. Review of Economic Studies **76**(4) (2009) 1239–1267
20. Foster, G.: It's not your peers, and it's not your friends: Some progress toward understanding the educational peer effect mechanism. Journal of Public Economics **90**(8-9) (2006) 1455–1475

21. Bourdieu, P., Passeron, J.C.: Reproduction in Education, Society and Culture. Second edition edn. Sage Publications Ltd (1990)
22. Sullivan, A.: Cultural capital and educational attainment. Sociology **35**(4) (2001) 893–912
23. Resnick, M., Wilensky, U.: Diving into complexity: Developing probabilistic decentralized thinking through role-playing activities. Journal of Learning Sciences **7**(2) (1998)
24. McPherson, M., Smith-Lovin, L., Cook, J.M.: Birds of a feather: Homophily in social networks. Annual Review of Sociology **27**(1) (2001) 415–444
25. Akerlof, G.A., Kranton, R.E.: Identity and schooling: Some lessons for the economics of education. Journal of Economic Literature **40**(4) (2002) 1167–1201
26. Baloff, N.: Extension of the learning curve – some empirical results. Operational Research Quarterly (1970-1977) **22**(4) (1971) 329–340