Multilevel Modelling and Agent-Based Modelling: Comparison and Integration
Final Report

Mauricio Salgado, Elio Marchione and Nigel Gilbert

October 11, 2011

Contents

1 Introduction 1
  1.1 Objectives . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 2

2 Description of the database 2

3 Fitting a Multilevel Model 3

4 The proposed ABM 4
  4.1 A mechanism-based explanation . . . . . . . . . . . . . . . . . . 4
  4.2 Main Results: Comparing MLM and ABM . . . . . . . . . . . . 5

5 Conclusions 7

6 Collaboration 7

7 Outputs 8
  7.1 Conference papers: . . . . . . . . . . . . . . . . . . . . . . . . . 8
  7.2 Working papers: . . . . . . . . . . . . . . . . . . . . . . . . . . . 8
  7.3 Poster presentations: . . . . . . . . . . . . . . . . . . . . . . . . . 8

1 Introduction

Multilevel models (MLM) have pioneered the analysis of data that have a hierarchical structure with two or more ‘levels’. They have been developed within a statistical paradigm, primarily as a method of describing and analysing large data-sets. Agent-based models (ABM) are also used to analyse social phenomena in which there are two or more ‘levels’ involved, often called the micro- and macro- levels. ABM were developed from a non-statistical background, drawing on artificial intelligence and physics. Agent-based modelling usually follow a
deductive or abductive methodology, testing a model against data, while multi-
level modelling is often inductive, deriving a model from data. MLM allow the
user to make inferences with known confidence, which is generally not true of
ABM, while ABM are capable of modelling non-linear, complex systems with
emergent behaviour. Thus to some extent the two modelling ‘paradigms’ are
interested in the same kind of issues but approach them from entirely different
directions and have different strengths.

1.1 Objectives

The aim of this short study was to clarify the similarities and differences be-
tween the two styles of modelling, attending to the modelling of levels, and to
investigate whether there is value in integrating some aspects. Although
there is some literature comparing the merits of equation-based models and
ABM, there is none comparing MLM with ABM. This study has provided one
of the first attempts to link these two methods and fill this gap in the scientific
literature.

To provide a basis for the comparison, the study took an example data set, for
which a MLM has been well studied, and then developed an ABM, based on this
MLM. The data set used was the “Inner London Educational Authority (ILEA)”
school data consisting of examination records from 140 secondary schools in
years 1985, 1986 and 1987 [8]. The project was carried out between September
2010 and April 2011. In the following sections the main results are presented.

In carrying out the project, there were no significant deviations from the
proposal, except that the start date was put back due to administrative delays
and that there was a ‘no-cost’ extension to allow for attendance at the ESSA
conference to present the results of the work in September 2011.

2 Description of the database

We use a subsample from 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 [8]. 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,
- Occupational 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, pupils’ scores in maths 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 [13], which does not contain any predictors but includes a random intercept variance term for groups. An analysis of this model showed that the ‘interclass-correlation’ coefficient (ICC) equals 0.119, so an important portion of the variance (~12%) is explained by the pupils’ group (i.e., school) membership. Further, the overall group mean reliability test [1] of the outcome variable equals 0.67, although several schools have quite low estimates. In fact, just 22 of the 48 schools have a group mean reliability over 0.7, which is the conventional value to determine whether groups can be reliably differentiated. Finally, the intercept variance \( u_{0j} \) is significantly different from zero, \( x^2(3) = 52.3, p < 0.001 \). 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.

### 3 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 [10], using the package nlme. The parameter estimation was carried out by using the algorithm Log-Cholesky [9]. A MLM with random coefficients that considered previous attainment, background factors (occupational class and gender) was fitted. In this model, the coefficients for previous attainment and the group intercepts were allowed to vary randomly across the 22 schools. Table 1 shows the results obtained from fitting this MLM. The average intercept across all the schools, \( \beta_0 \), equals 12.65 (std. error 1.79, \( p < 0.001 \)) and the average slope for Math 3 across the 22 schools \( \beta_1 \) equals 0.6 (std. error 0.05, \( p < 0.001 \)). 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, with a standard deviation estimated as 6.04. In addition, there is a negative covariance between intercepts and slopes, \( \sigma_{u01} \), estimated as -0.98, suggesting that schools with higher intercepts tend to have lower slopes.

The two control variables included in the model, gender and occupational class, perform differently. Only occupational class (i.e., ‘Nonman’ in Table 1) 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 1) does not contribute to the predictive
Table 1: Parameters of Random Slope Model for Maths Attainment in Year 5

<table>
<thead>
<tr>
<th>Parameters (Outcome Variable: Math 5)</th>
<th>Random Effects Parameters</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>St. Dev. ($\sigma$)</td>
<td>Intercept ($u_{0j}$)</td>
<td>6.04</td>
</tr>
<tr>
<td></td>
<td>Math 3 ($u_{1j}$)</td>
<td>0.14</td>
</tr>
<tr>
<td>Cov. ($\sigma_{u01}$)</td>
<td>Math 3*Intercept</td>
<td>-0.98</td>
</tr>
<tr>
<td></td>
<td>Residual ($e_{ij}$)</td>
<td>5.17</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fixed Effects</th>
<th>Estimate</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficients ($\beta_n$)</td>
<td>Intercept ($\beta_0$)</td>
<td>12.65***</td>
</tr>
<tr>
<td></td>
<td>Math 3 ($\beta_1$)</td>
<td>0.60***</td>
</tr>
<tr>
<td></td>
<td>Nonman ($\beta_2$)</td>
<td>1.17*</td>
</tr>
<tr>
<td></td>
<td>Boy ($\beta_3$)</td>
<td>-0.02</td>
</tr>
</tbody>
</table>

Note: *** = $p < 0.001$, ** = $p < 0.01$, * = $p < 0.05$.

power of the model, since its regression coefficient is not significantly different from zero.

4 The proposed ABM

We proposed a simple ABM that 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 ties in which the pupil is embedded. In order to establish comparisons between this ABM and the MLM explained in Section 3, we calibrate the former empirically using the same data we referred to in Section 2. The ABM was built in NetLogo 4.1.2 [14].

4.1 A mechanism-based explanation

The ABM assumes that a combination of friendship dynamics based on homophily [6, 4] and self-fulfilling prophecy [7] based on teacher expectations bias [11] can produce differential achievement among students and schools [2]. This explanatory mechanism can be established as follows. Firstly, there is a peer effect among pupils, which is brought about by the pupils’ tendency to sort themselves in groups with similar others. This lateral group formation mechanism affects their individual learning and progress, producing groupings of pupils with different academic performances. Secondly, the differences among
groups determine the way in which teachers interact with their pupils, since
groups of high-performance pupils capture more attention and receive more
feedback from teachers compared to groups of low-performance pupils. This
vertical behavioural mechanism also affects the pupils’ academic performance.
This explanatory mechanism is proposed as a likely one, and additional research
is needed to confirm or falsify it in real world settings. The proposed ABM pre-
sented in this report takes into account the two mechanisms described above
as the model microspecification [3]. Further details on the model description,
etties and main interactions can be found in [12] and [5].

4.2 Main Results: Comparing MLM and ABM

Table 2 shows the results for the parameter setting that minimises $d_j$. We
present the average distance (in the same units as the data) between the pre-
dicted scores and the real scores in Math 5 for both the multilevel model (‘MLM
($d_j$)’) and the simulation (‘ABM ($d_j$)’). The table also shows 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 occupational class (the last three columns of Table 2. 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 data scores in Math 5.

Comparing the averages 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 distance averaged over all schools equals 3.04 on a scale of
40 points. Thus, the ABM, despite its simplicity, offers a reasonable fit to the
data.

The simulation results suggest a high educational tolerance, since most of
the values equal 90% (except from school 30, in which the tolerance level equals
70%). On the other hand, the tolerance levels of occupational class and gender
vary across the schools. Therefore, the group formation mechanism in our sim-
ulation seems to be ruled by the variables occupational class and gender, while
previous attainment in maths does not discriminate much between groups.

The hypothesised mechanism that bring about the differences in school ef-
ectiveness seems to be justified. The simulation results indicate that the mech-
anism 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. 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 al-
lowed). This is clearly not the case. Thus, the sorting mechanism that has been
implemented in this simulation and the groups reflecting that mechanism seem
to be important in explaining the differences in effectiveness among schools.
Table 2: Calibration Results

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

In this report we have presented and compared the results of two models to address differential school effectiveness. The first was a MLM, where the hierarchical nature of educational processes is considered. The fitted model, controlling for some pupil background characteristics, was able to identify significant differences in the effectiveness of schools. However, with the available longitudinal data, it is not possible to model likely causal influences that might explain these differences. As we discussed, this is a common problem in multilevel modelling, in which, because of practical and conceptual problems, causal mechanisms are typically ignored. By building a simple ABM, we have tried to complement the statistical findings with a mechanism-based explanation. Because of the lack of a canonical mechanism-based explanation describing why school performances might emerge, we proposed a simple one, based not on direct empirical data but on plausible dynamics that previous research has found likely to explain differences in academic performance among pupils. These social mechanisms referred to the effects of friendship ties and teacher expectations bias on pupils’ educational achievement. They were the ABM microspecification.

After comparing the predictive and explanatory power of both models, we found that the MLM provides more accurate predictions than the ABM. However, although the ABM is rather simple, the differences in the predictions are small. The analyses from the two models made clear that, whereas MLM is data driven, ABM is both data driven and theory based, so the latter allows researchers to formalise and falsify in silico plausible mechanisms that might bring about the observed differences in performance across schools. In short, although the fitted MLM outperforms the proposed ABM, the latter still offers a reasonable fit and provides a causal mechanism to explain differences in the identified school performances that is absent in the MLM.

The next research step might be to falsify the proposed mechanism in ‘real world settings’. This would help to interpret and explain the differences established using MLM. That is, in order for researchers to determine whether friendship ties and teacher expectation bias provide an adequate explanation of differential school effectiveness, as we have computationally demonstrated here, one needs to gauge the results with peer effects and teacher expectations estimates based on actual (and not simulated) peer networks. This project has shown that both the generative sufficiency of ABM and the inferential nature of MLM can be complemented to provide sociological research with more powerful tools. Our work presents a first step in this direction.

6 Collaboration

We are very grateful to Professor Sylvia Richardson (BIAS, UCL), Professor Fiona Steele (LEMA, Bristol) and Dr Edmund Chattoe-Brown (SIMIAN, Leicester) for their advice and assistance throughout the project.
7 Outputs

7.1 Conference papers:


7.2 Working papers:


7.3 Poster presentations:


References


