

Harvey GOLDSTEIN

Chapter 1

MULTILEVEL MODELLING OF EDUCATIONAL DATA

Opening. Educational science was the first social science to develop fully multilevel modelling, although it had already been long used in statistics under the form of “random-effect models” or “mixture models” (Eisenhart *et al.*, 1947). In fact, for many years, the literature on education research had been the forum for substantive discussions on the most relevant analytical unit for measuring scholastic attainment: should statistical analysis focus on the class or on the student? As Harvey Goldstein explains in this first chapter, Aitkin *et al.* (1981) came up with a genuine multilevel model inspired by an earlier study (Bennett, 1976), which recognised only students and teaching styles as the units of analysis, ignoring teacher groupings and class groupings: Aitkin and his colleagues introduced the effect of these groupings while at the same time admitting a specific student effect. They thus managed to show the important role played by the teacher: this erased the sizeable differences found between students subjected to a “formal style of teaching” and to an “informal” style in the earlier study, which also referred to a “mixed style of teaching”. The choice between the class and the student ceased to be relevant. In fact, Aitkin and his co-authors argued, the effects of both aggregation levels should be examined together, so that their respective actions can be properly separated. Since that date, education researchers have accepted the need to take aggregation levels into account in order to understand the differences observed between students.

To take these analyses one step further, we must ask the following, more general question: are the inequalities observed between school-leavers greater than, equal to, or lesser than those observed when they entered the system, and should these differences be attributed to the system or to other, external reasons (Kereckhoff, 1991)? The search for a clear answer to this question leads us to envisage multilevel models far more complex than the previous one, which was hierarchical, did not examine all the aggregation levels needed to complete the analysis, and confined its study to a student’s progress in a single school year.

First, it is important to realise that a ranking of levels is not always satisfactory for analysing the most general situations in which students find themselves. It is often useful to consider “cross-classifications”, which involve no hierarchy. For example, when we classify students according to the school they attend and the environment in which they live, there is no reason to consider a hierarchy between these classifications. They should be treated as being at the same level, but we need to identify their effects separately. The new models offered should therefore accommodate cross-classifications.

Second, it is important to distinguish as many aggregation levels as possible in order to make sure that the dominant levels are effectively taken into consideration, and to avoid erroneous conclusions, even if we subsequently reduce that number to the most significant levels. In the example discussed above, teaching style—which can, incidentally, be broken down into more than three types—seemed important for analysing students’ performance; but the introduction of another aggregation level—the teacher—sharply diminishes the style effect or even changes it in some cases. The effective way to accommodate this is to consider each student as belonging to a large number of units at different levels, leading to “multiple membership models”. Naturally these must be able to separate cross-classifications from hierarchical classifications, while bringing both into play at the same time. In the end, this should make it possible to distinguish the effects of the school system from the effects of other systems or external influences.

Third, to properly introduce students’ entire life histories and link their level when entering the school system to their level on leaving it, we need to adopt an event-history approach, with which we can assess the influence of past events on scholastic achievement over time. Students can change classes, teachers, schools, regions, countries, etc. To model a result obtained at a given moment, we need to take into account students’ entire past histories—for example, all the schools they attended and all of their teachers. Such a model obviously requires far fuller information than that typically available to education researchers. But this information is vital to an understanding of how the school system in which students are involved, their human environment,

the constraints they face, and their own personalities influence their final scholastic attainment and the structure of outflows from the school system.

In this chapter, Harvey Goldstein explores the ways of factoring in these various effects and the models that can shed new light on the dynamics of the education process. By defining the units more flexibly, we can change our way of conceptualising the effects of institutions and of individuals' environments. Likewise, we can introduce an effect of the performance or of the actions of a school in a locality on the performance and the actions of other, nearby schools—which, in a way, means adding a form of spatial autocorrelation. This also requires the introduction of new ways of understanding the education system, as these methods must allow a simultaneous modelling of responses at different aggregation levels, without setting a logical priority of one over another. For example, we can quantify the link between student performances and teacher attitudes without requiring the latter to logically precede the former (or vice versa). This ultimately leads Goldstein to examine how the new approaches affect the epistemology of education.

D. C.

1. FUNDAMENTALS: UNITS AND LEVELS

While this chapter concentrates on educational data, many of the issues are more generally relevant. This is especially the case when we consider what we mean by a “unit” and a “level”.

Traditionally, pupils or students have been viewed as well-defined units, often at the lowest level of a data hierarchy. Likewise schools have typically been considered as well-defined units with students “nested” or grouped within them. Thus, we would ordinarily say that a particular student “belongs to” a particular school over a period of time. We immediately see, however, that this does not always accord with reality. For example, a student may move from one school to another during a study, or a school may change by splitting into two schools or being merged with another school. Likewise, if we consider the unit of a “class” within a school, this may vary in its form and composition frequently, and any particular student may experience several different teachers with which she or he is studying a particular subject during a year. We would regard such a student as “belonging to” or having a “multiple membership” of the set of teachers and we shall describe such structures in more detail later. The same kinds of issues arise in demographic studies (Goldstein *et al.*, 2001), health and other human sciences. In other words our definition of a “unit” has to incorporate temporal information (for what period does it exist in a particular form) and the relationships *between* units have to be specified in terms of membership at particular periods. This has clear implications for longitudinal studies, but also applies in many cases to purely cross-sectional data. Suppose, for example, that we are comparing student achievement test scores among schools at one point in time. It is well known that the use of such data to rank order schools is problematical because it fails to take account of students' prior “intake” achievements (see also Section 3). In addition, however, it assumes that the students can be assigned unequivocally to the schools they belong to at the time of the test. Typically there will be significant amounts of mobility among schools, so that many of the students will have been in their assigned schools a relatively short time and it will be unreasonable to expect those schools to have greatly influenced the student test scores.

It is important to distinguish between actual changes in the definition of a unit and merely changes in a unit's characteristics. Thus, for example, a school may change the number of students it has or its teachers, but may still be regarded as the same unit, and of course, when modelling the effects of school size or composition such changes can be taken into account. Even so, there may be borderline cases; a school may undergo reorganisation to such an extent that it changes nearly all of its staff, has a new name and perhaps new accommodation. If, at the same time, its student composition also changes with some moving to other schools and new students entering it may be more useful to consider this as a change of unit, with the new unit only coming into existence at that time. A choice in such a case will partly depend upon the research questions being asked and also upon the modelling feasibility of different approaches which I shall be discussing later.

There is another kind of unit whose definition is closely tied in with the definition of a "level". This is where a collection of lower level units belongs to a higher level unit which itself is defined solely (or partly) in terms of the particular units that belong to it. Thus, for example, a friendship group is defined solely by its members and will change when any are lost or acquired. Families and households have similar properties (Goldstein *et al.*, 2001). At any one time an individual person may belong to several friendship groups (or learning groups for example) and this can be regarded as a multiple membership structure. Because group formation can change over the period of study, it will often be the case that there will exist many more groups than there are individuals. In such a situation we may encounter modelling estimation problems in terms of separating out different effects: in terms of levels we may regard the groups as lower level units generating measurements which are "nested" within individuals, where each individual can be thought of as generating a set of friends. For example, in the case of learning groups within a classroom, if the results of a group project are being measured then the basic response is measured at the group level rather than that of the student and, so long as there is adequate movement of individuals among groups, this can be modelled as a multiple membership structure with groups at level 1 and students at level 2. This example will be discussed in more detail in Section 4. In household studies a similar situation can arise when a household characteristic, such as electricity use, is being measured over time with individuals changing households. In both these cases we can, in principle, obtain estimates of the relative contributions of groups (households) and individuals and also estimate (posterior) effects for each individual or group. The ability of such models in education to take account of group learning and performance provides a powerful tool for many situations that formerly have posed considerable analytical problems.

In order precisely to define a complex data structure of the kind I have been describing we need to specify the membership relationships among all the units involved. In fact there is just a small number of types of relationship involved and these can conveniently be set out diagrammatically for any particular structure. I shall describe how this may be done below when looking at examples.

Methodology and Epistemology of Multilevel Analysis

Approaches from Different Social Sciences

Courgeau, D. (Ed.)

2003, XII, 238 p., Hardcover

ISBN: 978-1-4020-1475-8