Contextual and Random Coefficient Multilevel Models. A Comparison

G. Rohwer, H.-P. Blossfeld February 2010

Abstract We discuss multilevel models focusing on individuals belonging to institutional units. It is assumed that the individual members of the institutional units cannot be identified by referring to positions. Modeling therefore requires analytical models relating to generic individuals. In this framework, we compare models using contextual variables and random coefficient multilevel models. We argue that random coefficient models do not offer substantial advantages when the goal is to explain individual outcomes.

1 Introduction

Multilevel models exist in different forms and are used for different applications. In this paper, we are interested in models for human individuals that take into account that individual outcomes are influenced by institutional units the individuals belong to. So there are two levels, individuals and institutional units. We consider mainly two modeling approaches: individual-focused models including contextual variables (derived from the institutional units), and random-coefficient multilevel (RCML) models.

The conceptualization of RCML models is often based on the assumption that the individual units associated with a higher-level unit can be identified through positions. This is reasonable, for example, when dealing with repeated measurements that can be identified by time points (serial numbers). The assumption is crucial because it is required for thinking of a joint distribution of the outcome variables of individuals belonging to the same higher-level unit. We therefore stress that in the applications we have in mind this assumption cannot be made.

Our approach is based on the notion of analytical models aiming to predict and, via interpretation, explain individual outcomes for generic individuals (meaning individuals defined by values of variables). Such models can easily include contextual variables and will then be called contextual multilevel models. The article argues that RCML models do not offer substantial advantages when the goal is to explain individual outcomes. Actually, much of the discussion of RCML models in the literature concerns the modeling of institutional units.[1] This is outside the scope of the present article. Our argumentation concerns the conceptual set-up of multilevel models for individual outcomes. Except for some brief remarks dealing with supposed implications of "dependencies among observations," questions concerning the estimation of model parameters will not be discussed.

[1] For example, most of the research questions referred to by Raudenbush and Bryk (2002) to illustrate their approach to RCML models concern statistically constructed properties of institutional units (e.g. "school effectiveness").

2 Descriptive and analytical models

We distinguish between descriptive and analytical models. Descriptive models serve to describe a given set of data or a population. Such models describe (aspects of) a statistical distribution that is defined for a sample or a population.[2] For example, given data containing information about household incomes, one can fit a lognormal distribution to describe the income distribution. This would be a descriptive model describing the income distribution in the sample. Instead, one can refer to a population of households and set up a model that uses a lognormal distribution for describing the income distribution in the population. Sampled data might then be used to estimate this descriptive model for the population.

In contrast to descriptive models, analytical models do not serve to describe data (or a population) but to formulate theoretical hypotheses. Very often the hypothesis concerns dependency relations between variables. For example, the question motivating the research might be how the educational success of children depends on conditions. This question cannot be answered by a descriptive model but requires an analytical model that formulates a hypothesis about a dependency relation. Moreover, in most applications, the hypothesis does not concern a particular child (or group of children) but a generic child, that is, any child that can be imagined to exemplify the theoretical process.

Analytical models are often formulated as regression models. A simple linear model could be written as

$$y = \alpha + x\beta + e \tag{1}$$

Understood as an analytical model, the equation formulates a hypothesis about the dependency of values of a variable y on values of a variable x. To account for the fact that, based on values of x, values of y can be predicted only probabilistically, the model includes a random variable e. Being part of an analytical model, this random variable represents the uncertainty in the prediction of y when using this model. Of course, one can assume $\mathrm{E}(e) = 0$, and this allows one to consider (1) as a hypothesis about the dependency of expected values of y on values of x: $\mathrm{E}(y|x) = \alpha + x\beta$.

Regression models can be, and often are, used for descriptive purposes. For example, having sampled values of the variables x and y for n individuals, say (x_i^*, y_i^*) for $i = 1, \ldots, n$,[3] one can set up a regression model

$$y_i^* = \alpha + x_i^*\beta + e_i \tag{2}$$

This would be a descriptive model that describes an aspect of the joint distribution of the variables found in the data. Being part of a descriptive model, e_i is not a stochastic variable but represents a residual from fitting the model. In fact, values of e_i can only be defined by using some method of estimating the parameters of the model; only then, having determined estimates $\hat{\alpha}$ and $\hat{\beta}$, one can define: $e_i^* := y_i^* - \hat{\alpha} - x_i^*\hat{\beta}$.

[2] The term 'population' is here used to denote a finite set of units which actually exist or have existed in the past. This understanding is required, in particular, in order to think of data as being a sample that is (randomly) drawn from a population.

[3] To distinguish values from variables, they are referred to by starred letters.

3 Varieties of multilevel models

Following the remarks in Section 2, we distinguish between descriptive and analytical multilevel models. We propose to understand analytical multilevel models as models formulating hypotheses about dependency relations that involve two or more different kinds of units.

For the present discussion, we distinguish four kinds of units: *Individuals;* in this paper these are always human individuals. *Institutional units;* for example: households, firms, schools. *Structured units;* these are groups (sets) of two or more individuals whose members can be identified by positions. For example, a couple consisting of a man and a woman. *Statistical units;* these are collections (sets) of individuals delimited by a common property. In contrast to structured units, members of a statistical unit cannot be distinguished by structurally defined positions (only by additional variables). As an example one can think of occupational groups, considered as sets of people having the same occupation.

The distinction between institutional and statistical units is important because, referring to an institutional unit, one can often think of a group of individuals being in some sense associated with the institutional unit. For example, the group of members of a household, or the group of employees of a firm. If considered as units of analysis, these groups of individuals are statistical units and must be distinguished from the institutional units referred to in their definition.

Based on the above definitions of kinds of units, we distinguish the following kinds of models:

a) Multilevel models for individuals. These are models having a dependent variable that refers to an individual and take into account that the process that generates values of that variable is influenced by the individual's being a member of, or in some way associated with, an institutional unit.
b) Models for institutional units. These are models having a dependent variable that refers to an institutional unit and take into account that the process that generates values of that variable is influenced by individuals (being in some way associated with the institutional unit). For example, one can think of a model that tries to explain the mean value (or some other aspect of the distribution) of wages paid by firms.
c) Models for structured units. These are models having a multidimensional dependent variable, say $(y_1, \ldots, y_n)$, where the components relate to the individual members of a structured unit. For example, when the unit is a couple, one can define a variable (y_1, y_2) where y_1 relates to the man and y_2 relates to the women. (Allowing the notion of structured units to refer to any kind of objects, one can also think of models for repeated measurements where time points can be used to identify individual measurements.)
d) Models for statistical units. Since the individual members of a statistical unit cannot be identified, dependent variables cannot refer to identifiable individuals. Instead, dependent and explanatory variables refer to statistical distributions defined for the statistical unit; models can therefore properly be called population-level models.[4] As examples one can think of diffusion models concerning the spread of some property in a population of individuals.

[4] A framework for the conceptualization of such models is discussed in Rohwer (2010).

The present paper focuses on analytical multilevel models for individuals. The models that will be discussed concern a dependent variable that is defined for a generic individual which belongs in some way to an institutional unit. References to institutional units can be made in one of two ways: a model can refer to a generic institutional unit or to a fixed population of institutional units.

4 Variance partitions and explanations

Discussions of multilevel models often start from a hierarchical data set. For example, a hierarchical data set containing values of two variables, y and x, for $i = 1, \ldots, n$ individuals could be given as follows:

$$(y_i^*, x_i^*, l_i^*) \tag{3}$$

In this notation, l_i^* provides the label of the institutional unit the individual i belongs to (say, $l_i^* \in \{1, \ldots, m\}$). We assume that y is the variable of interest, and x is a (possibly multidimensional) explanatory variable. In a formal sense, also the labels l_i^* can be considered as values of a variable, say l. There is, however, an important difference between x and l: in contrast to values of x, values of l cannot contribute to an explanation of values of y.

Think of an individual i having the value y_i^*. Why? Referring to the value x_i^* could, possibly, contribute to an answer; but the label, l_i^*, of an institutional unit (say, a school) the individual belongs to has no explanatory content. Of course, something that characterizes the institutional unit (e.g., class size) might contribute to the explanation of y_i^*. This information, however, does not derive from the label, but from knowing the value of an explanatory variable (that could be identical for several institutional units).

While labels cannot contribute to explanations, they can be used to partition the variance of a variable of interest, say $\mathrm{V}(y)$, into two components: a mean value of variances that are specific for each institutional unit (label), and the variance of the mean values. Such variance partitions are often made the starting point for introducing multilevel models. In fact, several authors suggest that the main task of multilevel analyses is to contribute to an explanation of "variability" (in a variable of interest) through methods of variance partition (e.g., Snijders and Bosker 1999, p. 1; Healy 2001; Browne et al. 2005; Stanat and Lüdtke 2008, p. 325; Heck and Thomas 2009, pp. 11, 51). This approach could be helpful when the goal is prediction. For example, knowing the conditional mean values $\mathrm{M}(y|l = j)$, the label l_i^* could be used for the prediction of i's value of y. However, variance partitions w.r.t. institutional units cannot contribute to explanations.

Being interested in explanations requires an approach that is conceptually different from partitions of variance. One has to start from an analytical model that refers to a generic individual. The meaning of "variability," when referring to the dependent variable of an analytical model, cannot be defined by referring to a sample (or population) of individuals. Considering the dependent variable of an analytical model as a random variable, its variance derives from the "error term," that is, a random variable representing the uncertainty of predicting values of the dependent variable.

5 Causal conditions and selection effects

Having defined the dependent variable to be considered in an analytical model (to be used for explanations), the next step is to think of processes that can generate values of the variable. This is the starting point for the consideration of conditions on which the processes might depend. The basic idea of multilevel models is that the process generating an individual's value of a dependent variable also depends in some way on the institutional unit to which the individual belongs. There are mainly two complementary possibilities:

a) One can think of features of the institutional unit that could be causally relevant conditions for processes generating values of the dependent variable. For example, one could assume that processes generating an individual's abilities in reading depend on features of the school in which the learning takes place, e.g., properties of the curriculum, qualification of the teachers, class size.
b) One can often assume that an individual is influenced by some or all other persons belonging to the same institutional unit. In the just mentioned example, one could assume that the individual's learning also depends on his or her interactions with other persons in the school.

Variables representing both kinds of circumstances will be called *contextual variables* if they are used to refer to conditions for a process that generates an individual's value of a dependent variable. In this understanding, contextual variables characterize individuals *situated in a context*. What makes these variables specific is that their definition (that is, the definition of possible values and their meanings) requires reference to an institutional unit.

Processes generating an individual's value of a variable must be distinguished from selection processes. For example, the educational level of a child's parents may be assumed to be one of the conditions for the process through which the child acquires its reading capabilities. Now imagine that the school selects children according to the educational level of their parents. Obviously, the selection process is conceptually different from the process through which the child learns reading. On the other hand, it might well be possible to consider the selection process as a substantial process that generates some feature of the institutional unit (the school) which, in turn, constitutes a causal condition for the child's learning process.[5]

6 Models with contextual variables

A linear version of an analytical multilevel model that employs contextual variables can be written as follows:

$$y = \alpha + x\,\beta + z\,\gamma + x\,z\,\delta + e \tag{4}$$

z is a contextual variable, x records characteristics of the individual that do not require reference to an institutional unit (both are possibly multidimensional variables); in addition, there is a random variable e having expectation $\mathrm{E}(e) = 0$.

[5] The consideration of selection effects becomes of critical importance when setting up models for institutional units. This is outside the scope of the present article.

Notice that there is no formal distinction between contextual and other explanatory variables. The formulation in (4) is completely symmetrical with respect to x and z. This is not surprising, of course, because this modeling approach uses contextual variables, like any other variables, to characterize conditions for a process that generates an individual outcome. These models will be called *contextual multilevel models*.

Understanding (4) as an analytical model means that it is intended to predict values of the dependent variable for a generic individual. Accordingly, the random term is understood as representing the uncertainty in making such predictions. Actually, the prediction is made by using the expectation of y and possibly adding some assessment of the uncertainty. The model that is actually used for predictions may be written as

$$\mathrm{E}(y|x,z) = \alpha + x\,\beta + z\,\gamma + x\,z\,\delta \tag{5}$$

showing how it predicts the expectation of y depending on values of x and z.

As formulated in (4), the model assumes that the random variable e representing the uncertainty is independent of the explanatory variables x and z. Note that this is a feature of the model, not of a (hierarchical) data set; in particular, this feature has nothing to do with properties of a sampling scheme that might be used to generate data for the estimation of model parameters.

It is quite possible to set up a model where the uncertainty of prediction depends on variables. A simple approach assumes a parametric density function for the dependent variable. For example, one could use a normal density function $\phi(y;\sigma)$, where y is the mean and σ is the standard deviation. If y is made dependent on covariates, e.g. in the linear form assumed in (5), but σ is treated as a single parameter, the model would again imply that the variance of the uncertainty is independent of covariates. On the other hand, it is quite possible to make also σ dependent on variables. As a result one would get a model where the variance of the distribution representing the uncertainty is no longer independent of explanatory variables. Instead of OLS, one could then use the maximum likelihood method to find estimates of the model parameters.

7 Models with labels of institutional units

Analytical models with contextual variables concern a generic individual, that is, an individual which is only characterized by values of variables. For example, a child of age eight that visits a school where it is learning reading. Moreover, also the institutional unit is referenced in a generic way. The model only takes into account values of contextual variables which do not identify particular institutional units.

A different approach to multilevel modeling starts from assuming a population of institutional units. For example, a collection of identifiable schools, or a collection of countries (e.g., the countries of the European Union). In order to refer to a collection of institutional units, we use the notation $\Omega = \{\omega_1,\ldots,\omega_m\}$, m being the number of units.

This presupposition allows one to define (at least conceptually) a separate model for each institutional unit:

$$y = \alpha_j + x\,\beta_j + e_j \quad (j = 1,\ldots,m) \tag{6}$$

Note that these models refer to identifiable institutional units but are, nonetheless, analytical models. Given the label j of an institutional unit, the model concerns a generic individual supposed to belong to that unit. By using dummy variables (say $d_j = 1$ if an individual belongs to ω_j, and $d_j = 0$ otherwise), one can also formulate a single model, often called a *fixed-effects multilevel model*. For example,

$$y = \Sigma_j\, d_j \alpha_j + x\, \Sigma_j d_j\, \beta_j + \Sigma_j d_j e_j \qquad (7)$$

would be equivalent to the full set of m separate models; by adding constraints one could define more restricted models.

A fixed-effects multilevel model can be understood as an analytical model that uses labels of institutional units as additional information for predicting values of a dependent variable defined for a generic individual. Alternatively, the model can be understood as a descriptive model aiming at the description of a population of institutional units without explicitly referring to a population of individuals. The description concerns properties of the institutional units defined by the regression models in (6) and therefore depends on the specification of these models. As mentioned, these are analytical models, they do not describe populations (or samples) of individuals belonging to the institutional units.

An obvious drawback of this modeling approach is that it does not allow including contextual variables which are defined with respect to the institutional units in Ω. So the question arises how to interpret differences between the m models. This mainly depends on whether the labels of the institutional units are informative or not. In some applications, it could be sensible to use informative labels; for example, when comparing countries. However, in many applications the number of institutional units is large and labels are not informative (think, e.g., of households, schools, and firms). Differences between the institutional units are then difficult to interpret. Since the labels are not informative, they cannot suggest ideas about variables which could have contributed to generating the differences. It is not even possible to conclude that the differences are due to unobserved contextual variables; at least some part may well be due to variables omitted from the model (6) that is used for the comparisons.

In short, fixed-effects multilevel models cannot be used to explain differences between institutional units in terms of variables characterizing the units; and consequently, they cannot contribute to explaining differences between individuals belonging to different institutional units in terms of (contextual) variables. These models are therefore seldom useful for analytical purposes.

8 Random coefficient multilevel models

In contrast to the modeling approach discussed in the previous section, random coefficient multilevel models do not use labels of institutional units as values of variables. Setting up this kind of multilevel model proceeds in three steps. The first step specifies a model for a generic individual. Using previous notations, this *level-1 model* could be written as

$$y = \alpha_0 + x\, \beta_0 + e_0 \qquad (8)$$

Assuming then that the processes generating values of y take place in the context of an institutional unit, the second step consists in specifying a *level-2 model* that makes the parameters of (8) dependent on properties that characterize the institutional unit. In our example, this model consists of two parts corresponding to the two parameters of (8) and could be specified as

$$\alpha_0 = \alpha + z\,\gamma + e_\alpha \quad \text{and} \quad \beta_0 = \beta + z\,\delta + e_\beta \qquad (9)$$

where z is a variable characterizing an institutional unit, and it is assumed that $\mathrm{E}(e_\alpha) = \mathrm{E}(e_\beta) = 0$. The third step consists of combining (8) and (9), resulting in the model

$$y = \alpha + x\,\beta + z\,\gamma + x\,z\,\delta + (e_\alpha + x\,e_\beta + e_0) \qquad (10)$$

This is called a *random coefficient multilevel* (RCML) *model*. Different from models discussed in the previous section, labels of institutional units do not occur. In fact, except for the formulation of the stochastic part, this model has the same structure as (4).

There is, however, also a conceptual difference. This becomes visible when asking how to understand the random variables, e_α and e_β, in the level-2 model. Why not simply assume $\alpha_0 = \alpha + z\,\gamma$ and $\beta_0 = \beta + z\,\delta$ that would make (10) completely identical to (4)?

Our understanding of these random variables is based on the following interpretation of the modeling approach:[6] It is intended to predict the value of y for a generic individual. It is known (or assumed) that the process generating that value depends on features of the institutional unit the individual belongs to, say ω_{j^*} (a member of Ω). This allows the further assumption that there is a model corresponding to (8), say

$$y = \alpha_{j^*} + x\,\beta_{j^*} + e_{j^*} \qquad (11)$$

that, if possible, should be used for the prediction. This is not possible, however, because the parameters, α_{j^*} and β_{j^*}, assumed for the particular institutional unit the individual belongs to are not known. Nevertheless, one can think of the level-2 model as providing estimates of the unknown parameters α_{j^*} and β_{j^*}.

Based on this interpretation, one can understand the random variables that are used in the formulation of the level-2 model as representing the uncertainty in the prediction of level-1 model parameters assumed to exist for the particular institutional unit a generic individual belongs to. This interpretation also highlights the conceptual difference between the modeling approaches:

a) The contextual multilevel model (4) assumes that the process generating an individual's value of y depends on conditions which can be represented by contextual variables, leaving it open whether and how the process might depend on further properties of the institutional unit the individual belongs to.
b) The RCML model (10), like a model using labels, assumes that the process generating an individual's value of y depends on the *particular* institutional unit the individual belongs to,[7] and it uses contextual variables, or any other variables characterizing institutional units, to estimate model parameters assumed for that unit.

---

[6] This interpretation is suggested, e.g., by DiPrete and Forristal 1994, p. 336; Hox 2000, p. 16; Heck and Thomas 2009, p. 78.

[7] For example, Raudenbush and Willms (1995, p. 308) define a 'school effect' as "the extent to which attending a particular school modifies a student's outcome." The contextual model would ask, instead, how a student's outcome depends on variables characterizing a school.

Note that the proposed interpretation of the random variables in the level-2 model is not in terms of sampling from a population of institutional units. Thinking in terms of sampling is sometimes proposed in the literature (e.g., Goldstein 2003, p. 15; Healy 2001), but would not be compatible with using RCML models as analytical models relating to generic individuals.[8] Neither would it make sense, then, to start from a randomly drawn institutional unit, nor would it be reasonable to start from a generic individual which is then placed into a randomly drawn institutional unit. An interpretation in terms of sampling would also be in contradiction to applications where model estimation is based on information about all of a small number of institutional units.[9]

9 Comparing the modeling approaches

Leaving aside the conceptual difference, the contextual multilevel model (4) and the RCML model (10) are very similar. In fact, except for different formulations of the stochastic parts, they are formally identical. An important common feature is that both models do not employ labels of institutional units. This implies that all substantial interpretations must be based on explicitly defined explanatory variables.

It is sometimes said that the RCML model allows formulating the hypothesis that the parameters of an individual-level model like (8) may vary between institutional units. However, the same assumption is implied in the contextual multilevel model (4). In fact, referring to an institutional unit characterized by a value z^*, both models imply the same individual-level model:

$$y = (\alpha + z^*\gamma) + x\,(\beta + z^*\delta) + e_0 \tag{12}$$

Starting from (4), e_0 is conceptually identical with e. On the other hand, following the RCML approach, one first uses (9) to predict $\alpha_0 = \alpha + z^*\gamma$ and $\beta_0 = \beta + z^*\delta$, and then inserts these values into (8).

An argument often given to suggest RCML models is that these models can be used to show how relationships between individual-level variables depend on the institutional context (e.g., Blien, Wiedenbeck, and Arminger 1994, p. 270; Goldstein 2003, p. 15; Raudenbush and Bryk 2002, p. 8). Mason, Wong, and Entwisle (1984, pp. 74-75) gave the following formulation:

> Our fundamental assumption is that the micro values of the response variable in some way depend on context and that the *effects* of the micro determinants *may* vary systematically as a function of context.

Referring to the RCML model (10), the interest concerns how the relationship between the conditional expectation $\mathrm{E}(y|x,z)$ and x depends on z. Again, both the RCML model and the contextual multilevel model (4) give the same answer:

$$\partial \mathrm{E}(y|x,z)/\partial x = \beta + z\delta \tag{13}$$

[8] Thinking in terms of sampling from a population of institutional units could be justified with being interested in the distribution of parameters of level-1 models in that population. However, RCML models presuppose parametric forms of the distributions of the level-2 random variables, and consequently cannot be used to learn something about the form of the distribution of parameters in the population of institutional units.

[9] Such applications are widespread in comparative political research; see, e.g., Dallinger 2008; Pichler and Wallace 2009.

The fact that both models give the same answer is a consequence of referring to institutional units not by labels, but only by explanatory variables. This implies that we do not distinguish between institutional units having identical values of the variable z.

10 Level-1 and level-2 variables

One often finds the suggestion that RCML models can be used to assess the relative importance of factors attributable to individuals and factors attributable to institutional units (e.g., Teachman and Crowder 2002; Heck and Thomas 2009, p. 14). The basic idea is that these models allow interpreting *level-1 variables* (defined by being included in the level-1 model) as representing factors attributable to an individual and *level-2 variables* (defined by being included in the level-2 model) as representing factors attributable to an institutional unit.

In our view, there are several reasons why this distinction between variables should not be used for substantial conclusions. The first point to note is that the distinction between level-1 and level-2 variables only reflects the stepwise procedure of setting up the model without having a substantial meaning. In fact, in the combined model (10), there is no longer any distinction between level-1 and level-2 variables.

Furthermore, there is no correspondence with the distinction between contextual variables and variables that can be defined without reference to an institutional unit. Contextual variables could be used in the level-1 model; on the other hand, it is not required that level-2 variables can be interpreted as contextual variables. Following the interpretation of RCML models proposed in Section 8, one can use any variables that might help to estimate parameters postulated for the institutional unit an individual belongs to. It is not required that the variable is in any sense a causal condition for a process that generates values of the dependent variable.

A further point concerns the random variables included in the level-2 model. By definition, these are level-2 variables and (therefore) often interpreted as representing unobserved influences attributable to an institutional unit (e.g., Kreft and de Leeuw 1998, p. 43; Snijders and Bosker 1999, p. 47). Correspondingly, the random variable included in the level-1 model is interpreted as representing influences attributable to the individual. Based on these interpretations, it is proposed that the variances of the random variables can be used to assess the relative importance of unobserved level-1 and level-2 variables (e.g., DiPrete and Forristal 1994, p. 338; Rice et al. 1998; Dallinger 2008; Heck and Thomas 2009, pp. 83, 88-9; Kim, Solomon and Zurlo 2009, p. 270).

An obvious objection derives from the fact that the level-2 model is based on having previously defined a level-1 model. All parameters of the level-2 model, including the variances of the random variables, depend on the specification of the level-1 model. Adding further level-1 variables will change these parameters and, in particular, can well lead to a decrease in the variances of the level-2 random variables. This shows that, even if accepting the meaningfulness of the distinction between level-1 and level-2 variables, no reliable conclusions can be drawn from the variance components in the stochastic part of the model.

The most important point, in our view, is that statements concerning the contribution of different kinds of variables should be based on explicitly defined variables (in contrast to interpreting variance components in terms of "unobserved variables"). This can well be done with models incorporating contextual variables. Such models also show that it is not, in general, possible to think of separable influences to be associated with different kinds of variables.

To illustrate, we use model (4) (using instead the RCML model (10), one would be led to the same conclusions). Assume two individuals having, respectively, values x_1^* and x_2^* of the variable x, and values z_1^* and z_2^* of the variable z. The model then predicts the following difference in the expected values of the dependent variable:

$$\begin{aligned} &\mathrm{E}(y|x_1^*, z_1^*) - \mathrm{E}(y|x_2^*, z_2^*) = \\ &(x_1^* - x_2^*)\,\beta + (z_1^* - z_2^*)\,\gamma + (x_1^* z_1^* - x_2^* z_2^*)\,\delta \end{aligned} \tag{14}$$

Due to the interaction effect, it is not possible to think of this difference as resulting from two independent sources (one attributable to the individuals and another one attributable to the institutional units). Even when comparing two individuals belonging to the same institutional unit ($z_1^* = z_2^* = z^*$), this would not be possible. One would get the equation

$$\mathrm{E}(y|x_1^*, z^*) - \mathrm{E}(y|x_2^*, z^*) = (x_1^* - x_2^*)\,(\beta + z^*\,\delta) \tag{15}$$

showing how the difference between the expectations of the individual scores still depends on the institutional context.

11 How to formulate models?

We have argued that one should clearly distinguish between descriptive and analytical models. Unfortunately, confusion easily results from the widespread habit of writing multilevel models in terms of variables referring to a sample of individuals; for example,

$$y_{ij} = \alpha + x_{ij}\beta + e_{ij} \quad (i = 1, \ldots, n_j; j = 1, \ldots, m) \tag{16}$$

where i and j refer, respectively, to individuals and to institutional units. The notation could be useful for descriptive models, that is, when x_{ij} and y_{ij} are meant to represent data, implying that also e_{ij} represents a fixed quantity (which must be defined by an estimation procedure for the parameters α and β). However, confusion is likely to occur when the notation is intended to set up a model that assumes e_{ij}, and consequently y_{ij}, to represent random variables. The notation then seems to allow thinking about a joint distribution of the variables e_{ij}, making it possible to formulate assumptions about correlations among its components. However, it is in no sense clear how to understand this joint distribution.

In contrast to an analytical model expressing a hypothesis about a generic individual, the formulation (16) refers, for each label j, to a plurality of individuals ($i = 1, \ldots, n_j$). In some applications, these individuals can be considered as members of a structured unit, and this would then allow thinking of corresponding variables having a joint distribution (in fact, this would lead to an *analytical* model for structured units). However, in most applications the institutional units referred to in multilevel models cannot be considered as structured units. This implies that the index i cannot be interpreted as referring to variables defined by an analytical model.

One is therefore led to view the formulation (16) as referring to a sample of individuals. This understanding might suggest to think of a sampling scheme that could be used for generating samples providing values of (x_{ij}, y_{ij}). However, this approach will not lead to an interpretation of the variables e_{ij} in terms of sampling. Even if the reference to a sampling scheme would allow interpreting (x_{ij}, y_{ij}) as random variables w.r.t. the sampling scheme,[10] this interpretation would not imply a definition of the random variables e_{ij}.

In fact, the definition of these variables requires two things: (a) the specification of a model for the prediction of values of y_{ij} based on values of x_{ij}, and (b) deciding about an estimation method for this model that can be used to find values of the e_{ij} variables. This shows that the random variables e_{ij} get their meaning, not from a sampling scheme, but from an analytical model that aims to predict values of a dependent variable from values of explanatory variables. They are used to capture the uncertainty in making the predictions.

We are led to the conclusion that, when using analytical models for generic individuals (in contrast to models for structured units), there is no sound foundation for thinking in terms of variables having subscripts referring to individuals. Furthermore, using subscripts referring to institutional units could be sensible, but implies treating these subscripts as labels of identifiable units. In any case, there is no valid notion of a joint distribution of the e_{ij} variables.

12 Dependencies among observations?

Authors proposing RCML models often argue with "dependencies among observations" of individuals belonging to the same institutional unit. For example, Kreft and de Leeuw (1998, p. 9) say:

> Observations that are close in time and/or space are likely to be more similar than observations far apart in time and/or space. Therefore, students in the same school are more alike than students in different schools, due to shared experiences, shared environment, etc. The sharing of the same context is a likely cause of dependency among observations.

Many similar statements can be found in the literature (e.g., de Leeuw 2002, p. xx; Blien, Wiedenbeck and Arminger 1994, pp. 268-9; Diez-Roux 1998, p. 220; Pickett and Pearl 2001, p. 117; Raudenbush and Bryk 2002, p. 100; Hox 2002, p. 5; Gorard 2003; Kim, Solomon and Zurlo 2009, p. 266; Heck and Thomas 2009, pp. 12, 76). Unfortunately, the expression 'dependency among observations' has no well-defined meaning.

One context for attempting an understanding is sampling theory. A clustered sampling scheme will lead to "dependencies among observations" in the sense that units belonging to the same cluster have higher second-order inclusion probabilities (compared with units belonging to different clusters). Being interested in the estimation of population parameters, the calculation of standard errors should take into account these dependencies (as reflected in the inclusion probabilities to be derived from the sampling scheme).

However, this understanding of "dependencies of observations" does not apply to the estimation of the parameters of an analytical multilevel model. These are not population parameters, but quantities postulated by setting up a model. Correspondingly, the random variables in these models are not defined w.r.t. a sampling scheme,

[10] Actually, since samples are not structured units, the subscript i does not make sense. Furthermore, using the subscript j would require the presupposition of a population of *identifiable* institutional units.

but reflect the uncertainty in using the model for predictions. In other words, analytical models relate, not to data-generating processes (literally understood), but to substantial processes that generate facts (possibly observed and then taken as data by a data-generating process).

Observations for estimating the parameters of an analytical model might come from a clustered sampling scheme, but this would be irrelevant for the definition and estimation of standard errors of the parameter estimates. For example, parameters of the contextual model (4) can be estimated with OLS regardless of the sampling scheme used to generate observations. Alternatively, one can start from assuming a parametric distribution for the dependent variable and then use maximum likelihood estimation. As already mentioned, this approach provides the opportunity to model heteroscedastic error terms. In any case, neither OLS nor maximum likelihood estimation will lead to "wrong" standard errors. It should be stressed that, when referring to analytical models, standard errors cannot be defined by referring to a sampling distribution (derived from a sampling scheme). A reasonable alternative could be to think in terms of "precision" that can be obtained from the given observations. However, leaving aside technical details, this line of reasoning makes standard errors always conditional on the data used to estimate model parameters.

In fact, proponents of RCML models most often do not argue with sampling schemes but with hierarchical structures. Reasoning in terms of (dependencies among) *observations* is then no longer appropriate. Instead, one has to think about how individuals might depend on relationships with other individuals belonging to the same institutional unit. Explicit modeling of such relationships would require the definition of a structured unit (allowing one to represent relationships between individuals by variables). When developing analytical models for a generic individual, the only option is to use contextual variables describing how the generic individual depends on other individuals and their properties.

13 Conclusion

We conclude that analytical multilevel models focusing on generic individuals can be set up as contextual models. These are regression models that use contextual variables to represent conditions deriving from the individual's belonging to an institutional unit. Assuming that the individuals belonging to an institutional unit cannot be identified by referring to positions, their interdependencies cannot be modeled in terms of a joint distribution. The obscure talk of "dependencies among observations" cannot (therefore) be given a clear meaning. Instead, one has to use contextual variables to capture an individual's dependence on other individuals and their properties.

Compared with contextual models, RCML models do not offer substantial advantages when the goal is to explain individual outcomes. The distinction between level-1 and level-2 variables suggested by RCML models should not be used for substantial conclusions. In particular, there is no reliable meaningful interpretation for the variance components associated with the error variables.

References

Blien, U., Wiedenbeck, M., Arminger, G.: Reconciling Macro and Micro Perspectives by Multilevel Models: An Application to Regional Wage Differences. In: I. Borg, P. P. Mohler (Eds.), Trends and Perspectives in Empirical Social Research, 266–282. Berlin: de Gruyter (1994)

Browne, W. J., Subramanian, S. V., Jones, K., Goldstein, H.: Variance Partitioning in Multilevel Logistic Models that Exhibit Overdispersion. J. R. Stat. Soc. **A 168**, 599–613 (2005)

Dallinger, U.: Sozialstaatliche Umverteilung und ihre Akzeptanz im internationalen Vergleich: Eine Mehrebenenanalyse. Z. Soziol. **37**, 137–157 (2008)

Diez-Roux, A. V.: Bringing Context Back into Epidemiology: Variables and Fallacies in Multilevel Analysis. Am. J. Publ. Health **88**, 216–222 (1998)

de Leeuw, J.: Series Editor's Introduction to Hierarchical Linear Models. In: S. W. Raudenbush, A. S. Bryk, Hierarchical Linear Models, 2nd ed. Thousand Oakes: Sage (2002)

DiPrete, T. A., Forristal, J. D.: Multilevel Models: Methods and Substance. Annu. Rev. Sociol. **20**, 331–357 (1994)

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

Gorard, S.: What is Multi-level Modelling For? Br. J. Educ. Stud. **51**, 46–63 (2003)

Healy, M. J. R.: Multilevel Data and Their Analysis. In: A. H. Leyland, H. Goldstein (eds.), Multilevel Modelling of Health Statistics, 1–12. New York: Wiley (2001)

Heck, R. H., Thomas, S. L.: An Introduction to Multilevel Modeling Techniques, 2nd ed. New York: Routledge (2009)

Hox, J. J.: Multilevel Analyses of Grouped and Longitudinal Data. In: T. D. Little, K. U. Schnabel, J. Baumert (eds.), Modeling Longitudinal and Multilevel Data, 15–32. Mahwah: Lawrence Erlbaum (2000)

Hox, J. J.: Multilevel Analysis. Techniques and Applications. Mahwah: Lawrence Erlbaum (2002)

Kim, T. K., Solomon, P., Zurlo, K. A.: Applying Hiearchical Linear Modeling (HLM) to Social Work Administration Research. Adm. Soc. Work **33**, 262–277 (2009)

Kreft, I., de Leeuw, J.: Introducing Multilevel Modeling. London: Sage (1998)

Mason, W. M., Wong, G. Y., Entwisle, B.: Contextual Analysis Through the Multilevel Linear Model. In: S. Leinhardt (ed.), Sociological Methodology 1983-84, 72–103. San Francisco: Jossey-Bass (1984)

Pichler, F., Wallace, C.: What are the Reasons for Differences in Job Satisfaction across Europe? Individual, Compositional, and Institutional Explanations. Eur. Sociol. Rev. **25**, 535–549 (2009)

Pickett, K. E., Pearl, M.: Multilevel Analyses of Neighborhood Socioeconomic Context and Health Outcomes: A Critical Review. J. Epidemiol. Community Health **55**, 111–122 (2001)

Raudenbush, S. W., Willms, J. D.: The Estimation of School Effects. J. Educ. Behav. Stat. **20**, 307–335 (1995)

Raudenbush, S. W., Bryk, A. S.: Hierarchical Linear Models: Applications and Data Analysis Methods, 2nd ed. Thousand Oaks: Sage (2002)

Rice, N., Carr-Hill, R., Dixon, P., Sutton, M.: The Influence of Households on Drinking Behaviour: A Multilevel Analysis. Soc. Sci. & Med. **46**, 971–979 (1998)

Rohwer, G.: Models in Statistical Social Research. London: Routledge (2010)

Snijders, T. A. B., Bosker, R. J.: Multilevel Analysis. London: Sage (1999)

Stanat, P., Lüdtke, O.: Multilevel Issues in International Large-Scale Assessment Studies on Student Performance. In: F. J. R. van de Vijver, D. A. van Hemert, Y. H. Poortinga (Eds.), Multilevel Analysis of Individuals and Cultures, 315–344. New York: Lawrence Erlbaum (2008)

Teachman, J., Crowder, K.: Multilevel Models in Family Research: Some Conceptual and Methodological Issues. J. Marriage Fam. **64**, 280–294 (2002)