

Multilevel modeling for longitudinal data: concepts and applications

Joseph F. Hair, Jr.

University of South Alabama, Mobile, Alabama, USA, and

Luiz Paulo Fávero

University of São Paulo, São Paulo, Brazil

Multilevel
modeling for
longitudinal
data

459

Received 19 April 2019
Revised 26 July 2019
Accepted 12 August 2019

Abstract

Purpose – This paper aims to discuss multilevel modeling for longitudinal data, clarifying the circumstances in which they can be used.

Design/methodology/approach – The authors estimate three-level models with repeated measures, offering conditions for their correct interpretation.

Findings – From the concepts and techniques presented, the authors can propose models, in which it is possible to identify the fixed and random effects on the dependent variable, understand the variance decomposition of multilevel random effects, test alternative covariance structures to account for heteroskedasticity and calculate and interpret the intraclass correlations of each analysis level.

Originality/value – Understanding how nested data structures and data with repeated measures work enables researchers and managers to define several types of constructs from which multilevel models can be used.

Keywords Multilevel modeling, Longitudinal data, Fixed effects, Random effects, Stata

Paper type Research paper

1. Introduction

Regression models for longitudinal data are very useful when the researcher wishes to study the behavior of a given phenomenon in the presence of nested data structures with repeated, or longitudinal, measures.

While in nested structures of clustered data certain explanatory variables do not present variation between observations (representing a level of analysis) coming from a given group (representing another level of analysis), in data structures with repeated measures there is also the temporal evolution, a fact that enables the researcher to investigate the individual reasons that may lead each of the observations to present different behaviors of the dependent variable, for the same group or for distinct groups, over time (Fávero, 2010; Martins & Terra, 2015; Misangyi, Lepine, Algina, & Goeddeke, 2006).

© Joseph F. Hair, Jr. and Luiz Paulo Fávero. Published in *RAUSP Management Journal*. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licenses/by/4.0/legalcode>

Joseph F. Hair, Jr. and Luiz Paulo Fávero contributed equally to all aspects of this paper.



For example, certain school data that does not vary among students, such as location and size, can be compared with data from other schools; and certain student data, such as sex and religion, that do not vary over time, can be compared with data from other students, which allows the different influences in the dependent variable to be analyzed. In all of these situations (nested data without or with repeated measures), datasets provide structures from which hierarchical models can be estimated.

Multilevel regression models have become considerably important in several fields of knowledge, and the publication of papers that use estimations related to these models has become more and more frequent (Goldstein, 2011). The reason for the importance of multilevel modeling is due mainly to the determination of research constructs that consider the existence of nested data structures, in which certain variables show variation between distinct units that represent groups but do not assess variation between observations that belong to the same group. In addition, the computational development and investments that data analysis software developers have made in the processing capacity to estimate multilevel models have also provided support to researchers who are increasingly interested in this type of approach (Gelman & Hill, 2007; Hough, 2006; Santos, Fávero, & Distadio, 2016; Serra & Fávero, 2018).

Theoretically, researchers can define a construct with a greater number of levels of analysis, even if the interpretation of model parameters is not something trivial. For instance, imagine the study of school performance, throughout time, of students nested into schools, these nested into municipal districts, these into municipalities, and these into states of the federation. In this case, we would be working with six analysis levels (temporal evolution, students, schools, municipal districts, municipalities and states).

The main advantage of multilevel models over traditional regression models estimated, for instance, by ordinary least squares (OLS), is the possibility of considering a natural nesting of data (Steenbergen & Jones, 2002), that is, multilevel models enable us to identify and analyze individual heterogeneities, and heterogeneities between the groups, to which these individuals belong, making it possible to specify random components in each analysis level (Heck & Thomas, 2009).

Multilevel models correct for the fact that observations in the same group are not independent and thus, compared to OLS models, lead to unbiased estimates of standard errors (SEs). But one could say that the same can be obtained with clustered standard errors in OLS. Indeed, if the number of clusters is plentiful (i.e. above 20), clustered SEs in OLS models and multilevel models are equally adequate for precision estimates of group-level effects. On the other hand, if there are less than 20 clusters, researchers should avoid using clustered SEs and adopt multilevel modeling. Furthermore, if researchers are also interested in testing whether group-level covariates moderate individual-level effects, multilevel models may prove to be the most appropriate choice (Arceneaux & Nickerson, 2009; Steenbergen & Jones, 2002).

According to Courgeau (2003), within a model structure with a single equation, there seems to be no connection between individuals and the society in which they live. In this sense, the use of level equations enables the researcher to “jump” from one science to another: students and schools, families and neighborhoods, firms and countries. Ignoring this relationship means to elaborate incorrect analyzes about the behavior of the individuals and, equally, about the behavior of the groups. Only the recognition of these reciprocal influences allows the correct analysis of the phenomena.

This is in line with what is called by Mathieu and Chen (2011) the multilevel paradigm, which refers to a way of thinking: considering management phenomena in context and looking for driving variables not only from the focal unit of analysis but also from levels

above and below. Such an approach often implies the development of multidisciplinary theories and investigations, what is the spirit articulated by [Hitt, Beamish, Jackson, and Mathieu \(2007\)](#) when discussing the built of theoretical and empirical bridges across levels through multilevel modeling. Most modern-day multilevel studies seek to associate relationships across proximal layers, such as team attributes and members' attitudes or environmental conditions and variable performance.

In an effort to make multilevel modeling more accessible, we provide the syntax for the mixed procedures in Stata for each step and show how to test and compare these designs in the model-building process. Previous discussions involving multilevel data have illustrated the use of multilevel modeling in programs such as MLn ([Kreft & de Leeuw, 1998](#)), R ([Bliese & Ployhart, 2002](#)), HLM ([Raudenbush, Bryk, Cheong, Congdon, & Du Toit, 2004](#)), and SAS ([Littell, Milliken, Stroup, & Wolfinger, 2004](#); [Singer, 1998](#)).

In this paper, our focus will be on hierarchical linear models (HLM), also known as linear mixed models (LMM). According to [West, Welch, and Galecki \(2015\)](#), the term "linear mixed models" comes from the fact that these models present linear specification and the explanatory variables include a mix of fixed and random effects. That is, they can be inserted into components with fixed effects, as well as into components with random effects. While the estimated fixed effects parameters indicate the relationship between explanatory variables and the metric dependent variable, the random effects components can be represented by the combination of explanatory variables and non-observed random effects.

Our main objectives are:

- to introduce the concepts of nested data structures;
- to define the type of model to be estimated based on the characteristics of the data;
- to estimate parameters through several methods in Stata;
- to interpret the results obtained through several types of existing estimations for multilevel models; and
- to define the most suitable estimation for diagnosing and forecasting effects in each of the cases studied.

2. Nested data structures

Models that take into account the presence of nested structures in the data offer benefits to researchers since they make possible the study of the sources of variance, in different levels, of an outcome variable.

[Raudenbush and Bryk \(2002\)](#) discuss the applications of multilevel modeling from nested data structures in various areas of knowledge, particularly education. In this field, works of [Aitkin and Longford \(1986\)](#), [Raudenbush and Bryk \(1986\)](#), [Garner and Raudenbush \(1991\)](#), [Raudenbush \(1993\)](#), [Rumberger and Thomas \(1993\)](#), [O'Connell and McCoach \(2008\)](#), and [Goldstein \(2011\)](#) deserve mention.

Multilevel modeling is also extensively used in strategy literature to compare existing variances at firm and industry levels for firm performance composition, such as [Schmalensee \(1985\)](#), [McGahan and Porter \(1997\)](#), [Brush and Bromiley \(1997\)](#), [Mauri and Michaels \(1998\)](#), [Brush, Bromiley, and Hendrickx \(1999\)](#), [Chang and Singh \(2000\)](#), [Bowman and Helfat \(2001\)](#), [McGahan and Porter \(2002\)](#), [Ruefli and Wiggins \(2003\)](#), [Short, Ketchen, Palmer, and Hult \(2007\)](#) and [Short, McKelvie, Ketchen, and Chandler \(2009\)](#). Other authors have analyzed the country-of-origin effect on performance, notably [Christmann, Day, and Yip \(1999\)](#), [Lee \(2003\)](#), [Hawawini, Subramanian, and Verdin \(2004\)](#), [Makino, Beamish, and](#)

Zhao (2004a), Makino, Isobe, and Chan (2004b), Misangyi *et al.* (2006), Goldszmidt, Brito, and Vasconcelos (2007), Fávero (2008), and Holcomb, Combs, Sirmon, and Sexton (2010).

Therefore, multilevel regression models enable us to formally investigate the behavior of a certain dependent variable Y , which represents the phenomenon we are interested in, based on the behavior of explanatory variables, whose changes may occur for clustered data, between observations and between groups to which these observations belong, and for data with repeated measures throughout time. In other words, there must be variables that have data that change between individuals that represent a certain level. But these variables remain unchanged for certain groups of individuals, and these groups represent a higher level.

First, imagine a dataset with data on n individuals, and each individual $i = 1, \dots, n$ belongs to one of the $j = 1, \dots, J$ groups, obviously $n > J$. Therefore, this dataset can have certain explanatory variables X_1, \dots, X_Q that refer to each individual i , and other explanatory variables W_1, \dots, W_S that refer to each group j ; but they are invariable for the individuals of a certain group. Table I shows the general model of a dataset with a two-level clustered/nested data structure (individual and group).

Based on Table I, we can see that X_1, \dots, X_Q (columns 4 to 6) are level-1 variables (data change between individuals), and W_1, \dots, W_S (columns 7 to 9) are level-2 variables (data change between groups; however, not for the individuals in each group). Furthermore, quantities of individuals in groups 1, 2, \dots, J (column 2) are equal, respectively, to $n_1, n_2 - n_1, \dots, n - n_{j-1}$ (column 1). Figure 1 shows the existing nesting between the level-1 units (individuals) and the level-2 units (groups), which characterizes the existence of clustered data.

Imagine another dataset in which, in addition to the nesting presented for clustered data, there is temporal evolution. That is, data with repeated measures. Thus, besides the individuals that will now belong to level 2 and therefore will be called $j = 1, \dots, J$, nested into

(Observation) (individual i) level 1 (1)	Group j level 2 (2)	Y_{ij} (3)	X_{1ij} (4)	X_{2ij} (5)	...	X_{Qij} (6)	W_{1j} (7)	W_{2j} (8)	...	W_{Sj} (9)
1	1	Y_{11}	X_{111}	X_{211}	...	X_{Q11}	W_{11}	W_{21}	...	W_{S1}
2	1	Y_{21}	X_{121}	X_{221}	...	X_{Q21}	W_{11}	W_{21}	...	W_{S1}
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
n_1	1	Y_{n_11}	X_{1n_11}	X_{2n_11}	...	X_{Qn_11}	W_{11}	W_{21}	...	W_{S1}
$n_1 + 1$	2	$Y_{n_1+1,2}$	$X_{1n_1+1,2}$	$X_{2n_1+1,2}$...	$X_{Qn_1+1,2}$	W_{12}	W_{22}	...	W_{S2}
$n_1 + 2$	2	$Y_{n_1+2,2}$	$X_{1n_1+2,2}$	$X_{2n_1+2,2}$...	$X_{Qn_1+2,2}$	W_{12}	W_{22}	...	W_{S2}
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
n_2	2	Y_{n_22}	X_{1n_22}	X_{2n_22}	...	X_{Qn_22}	W_{12}	W_{22}	...	W_{S2}
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
$n_{j-1} + 1$	J	$Y_{n_{j-1}+1,J}$	$X_{1n_{j-1}+1,J}$	$X_{2n_{j-1}+1,J}$...	$X_{Qn_{j-1}+1,J}$	W_{1J}	W_{2J}	...	W_{SJ}
$n_{j-1} + 2$	J	$Y_{n_{j-1}+2,J}$	$X_{1n_{j-1}+2,J}$	$X_{2n_{j-1}+2,J}$...	$X_{Qn_{j-1}+2,J}$	W_{1J}	W_{2J}	...	W_{SJ}
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
n	J	Y_{nJ}	X_{1nJ}	X_{2nJ}	...	X_{QnJ}	W_{1J}	W_{2J}	...	W_{SJ}

Table I.
General model of a dataset with a two-level clustered/nested data structure

Note: The dataset will have a balanced nested data structure if $n_1 = n_2 - n_1 = \dots = n - n_{j-1}$

$k = 1, \dots, K$ groups (which now belong to level 3), we will also have $t = 1, \dots, T_j$ periods in which each j individual is monitored. Consequently, this new dataset can have the same explanatory variables X_1, \dots, X_Q that refer to each j individual. But now they are invariable for each j individual during the periods of monitoring. Moreover, the dataset can also have the same explanatory variables W_1, \dots, W_S that refer to each group k . But W_1, \dots, W_S are also invariable throughout time for each group k . Table II provides the logic to describe a dataset with a three-level nested data structure with repeated measures (time, individual and group).

Based on Table II, we can now see the variable that corresponds to the period is a level-1 explanatory variable (column 1), since the data change is in each row of the dataset, and that X_1, \dots, X_Q (columns 5 to 7) become level-2 variables (data change between individuals, but not for the same individual throughout time), and that W_1, \dots, W_S (columns 8 to 10) become level-3 variables (data change between K groups (column 3), but not for the same group throughout time). Furthermore, quantities of periods in which individuals $1, 2, \dots, J$ (column 2) are monitored are equal, respectively, to $T_1, T_2 - T_1, \dots, T_j - T_{j-1}$ (column 1).

Similar to what was shown for the case with two levels, Figure 2 enables us to see the existing nesting between the level-1 units (temporal variation), the level-2 units (individuals), and the level-3 units (groups), which characterizes a data structure with repeated measures.

Through Tables I and II, as well as through the corresponding Figures 1 and 2, we can see that the data structures present absolute nesting. That is, a certain observation can be nested into only one group, and this group into only another higher-level group, and so on.

In the next section, we will estimate multilevel models with repeated measures in Stata, whose econometric development is in Appendix 1. Appendix 2 is intended for the presentation of the commands in Stata.

3. Estimation of multilevel models with repeated measures in Stata

This section gives researchers the opportunity to estimate multilevel models through Stata Statistical Software®. For our example, we will use the step-up strategic multilevel analysis proposed by Raudenbush and Bryk (2002), and Snijders and Bosker (2011). That is, we first studied the variance decomposition from the definition of a null model (non-conditional model), so that afterwards, a random intercepts model and a random intercepts and slopes model could be estimated. Finally, from the definition of the random nature of the error terms, we estimated the complete model by including level-2 variables into the analysis. We, therefore, estimate a three-level hierarchical linear model, in which the nesting of data will be characterized due to the presence of repeated measures. Thus, there is temporal evolution in the behavior of the dependent variable.

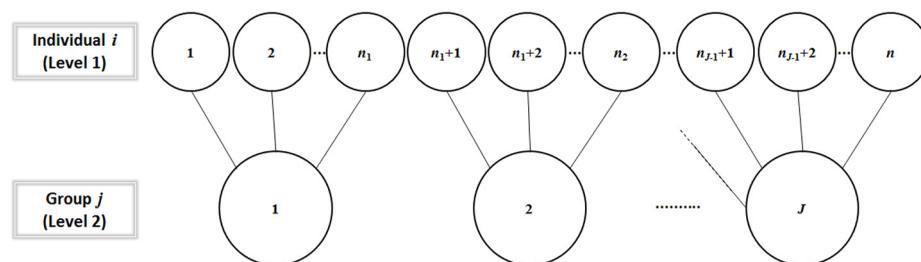


Figure 1. Two-level nested structure of clustered data

Table II.
General model of a
dataset with a three-
level nested data
structure with
repeated measures

Period t (repeated measure) level 1 (1)	(Observation) (individual j) level 2 (2)	Group k level (3)	Y_{ijk} (4)	X_{1ijk} (5)	X_{2ijk} (6)	...	X_{Qijk} (7)	W_{1k} (8)	W_{2k} (9)	...	W_{SK} (10)
1	1	1	Y_{111}	X_{111}	X_{211}	...	X_{Q11}	W_{11}	W_{21}	...	W_{S1}
2	1	1	Y_{211}	X_{111}	X_{211}	...	X_{Q11}	W_{11}	W_{21}	...	W_{S1}
...
T_1	1	1	Y_{T_11}	X_{111}	X_{211}	...	X_{Q11}
$T_1 + 1$	2	2	$Y_{T_1+1,21}$	X_{121}	X_{221}	...	X_{Q21}
$T_1 + 2$	2	2	$Y_{T_1+2,21}$	X_{121}	X_{221}	...	X_{Q21}
...
T_2	2	1	Y_{T_221}	X_{121}	X_{221}	...	X_{Q21}	W_{11}	W_{21}	...	W_{S1}
...
$T_{J-1} + 1$	J	K	$Y_{T_{J-1}+1,JK}$	X_{1JK}	X_{2JK}	...	X_{QJK}	W_{1K}	W_{2K}	...	W_{SK}
$T_{J-1} + 2$	J	K	$Y_{T_{J-1}+2,JK}$	X_{1JK}	X_{2JK}	...	X_{QJK}	W_{1K}	W_{2K}	...	W_{SK}
...
T_J	J	K	Y_{T_JK}	X_{1JK}	X_{2JK}	...	X_{QJK}	W_{1K}	W_{2K}	...	W_{SK}

Note: The dataset will have a balanced longitudinal data structure if $T_1 = T_2 - T_1 = \dots = T_J - T_{J-1}$

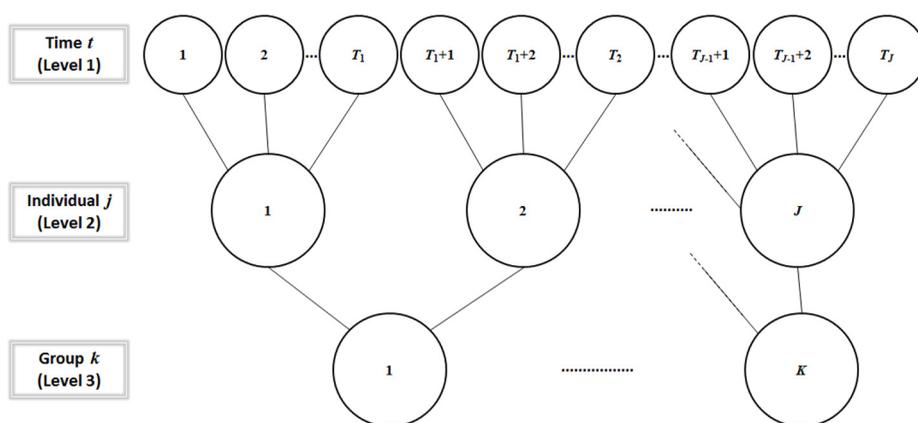


Figure 2. Three-level nested structure with repeated measures

3.1 Hypotheses and data

A dataset was constructed by a professor interested in monitoring students' school performance for a certain period of time, in order to investigate if there is variability in this performance throughout time between students within the same school, and between those from different schools. In addition, if there are certain student and school characteristics that explain this variability. This dataset follows the logic of the seminal work developed by [Raudenbush, Rowan, and Kang \(1991\)](#).

A total of 15 schools volunteered to provide data on their students' school performance (scores from 0 to 100) in the last four years, a total of 610 students. In addition, the professor also obtained each student's gender in the dataset in order to verify if there are differences in school performance resulting from this variable. The variable regarding professors' years of teaching experience, for each school, also was included in the study. The dataset PerformanceTimeStudentSchool.dta can be found in [Fávero and Belfiore \(2019\)](#).

It is important to mention that, although traditional maximum likelihood estimation methods for multilevel modeling have been shown to provide biased estimates when the number of clusters is below 30, methods such as restricted maximum likelihood (REML) estimation have shown potential to perform well with ten clusters or fewer ([McNeish & Stapleton, 2016](#)). In this paper, as discussed below, we estimate all models through REML.

First, we test three fundamental hypotheses regarding the nature of student's school performance over time, following the logic proposed by [Raudenbush and Bryk \(2002\)](#) and [Short, Ketchen, Bennett, and Du Toit \(2006\)](#):

- H1. There is significant variance in student's school performance over time, both within students and schools.
- H2. Student's school performance follows a linear trend over time.
- H3. There is significant variance around a linear performance trend.

Student and schools' characteristics may influence student's school performance over time. Thus, we test the following:

- H4. There is a significant relationship between students' characteristics and student's school performance over time.

H5. There is a significant relationship between schools' characteristics and student's school performance over time.

We have a balanced longitudinal data structure since all 610 students are monitored in the four periods. Figure 3 enables us to analyze the temporal evolution of the school performance of the first 50 students in the sample. From the trends in the lines, we can see that the temporal evolutions of the school performance have different intercepts and slopes between students. These different intercepts and slopes justify the use of multilevel modeling and provide reasons to include intercept and slope random effects in Level 2 of the models that will be estimated.

Figure 4 shows the temporal evolutions of the average school performance. The increasing trend over time provides further justification for estimating a three-level

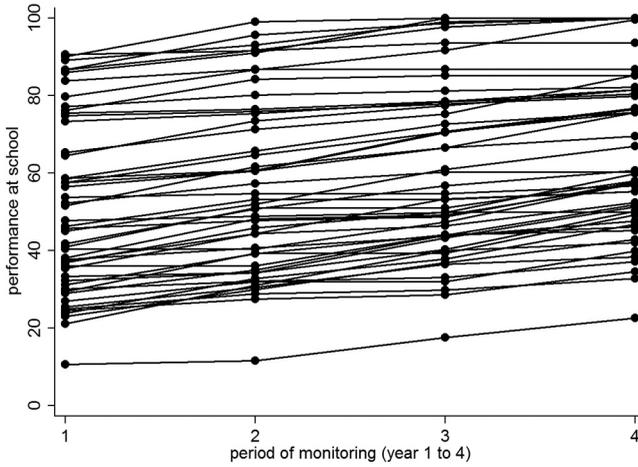


Figure 3.
Temporal evolution
of the school
performance of the
first 50 students in
the sample

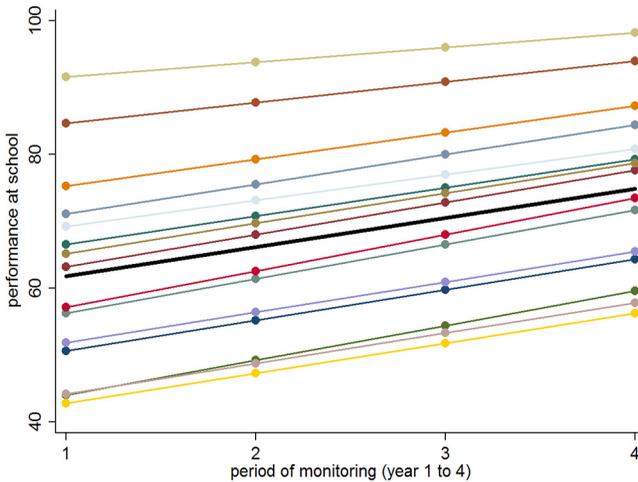


Figure 4.
Temporal evolution
of students' average
school performance at
each school (linear
adjustment through
OLS)

hierarchical model. Figure 4 also shows the linear adjustment through OLS of the school performance behavior over time for each school. In addition, the figure enables us to display the intercept and slope random effects in Level 3 of the models that will be estimated, since the temporal evolutions of the school performance present different intercepts and slopes between the schools.

3.2 Step-up strategic multilevel modeling

Having characterized the temporal nesting of the students from different schools in the data with repeated measures, we can initially estimate a null model (non-conditional model) that enables us to determine if there is variability in the school performance between students from the same school and between those from different schools. No explanatory variable is inserted into the modeling, which only considers the existence of one intercept and of error terms u_{00k} , r_{0jk} and e_{tjk} , with variances, respectively, equal to τ_{u000} , τ_{r000} and σ^2 . The model to be estimated has the following expression:

Null Model:

$$performance_{tjk} = \pi_{0jk} + e_{tjk}$$

$$\pi_{0jk} = b_{00k} + r_{0jk}$$

$$b_{00k} = \gamma_{000} + u_{00k}$$

which results in[1]:

$$performance_{tjk} = \gamma_{000} + u_{00k} + r_{0jk} + e_{tjk}$$

At the top of Figure 5, we can initially demonstrate that we have a balanced longitudinal data structure since for each student we have minimum and maximum quantities of periods of monitoring equal to four, with a mean also equal to four.

About the fixed effects component, we can see that the estimation of parameter γ_{000} is equal to 68.714, which corresponds to the average of students' expected annual school performance of the (horizontal line estimated in the null model, or general intercept). Moreover, at the bottom of Figure 5, the estimations of the variances of error terms $\tau_{u000} = 180.194$, $\tau_{r000} = 325.799$ and $\sigma^2 = 41.649$ are presented.

We can, therefore, define two intraclass correlations, given the existence of two variance proportions. The first one refers to the correlation between the data of variable *performance* in t and in t' ($t \neq t'$) of a certain student j from a certain school k (Level-2 intraclass correlation). The other one refers to the correlation between the data of variable *performance* in t and in t' ($t \neq t'$) of different students j and j' ($j \neq j'$) from a certain school k (Level-3 intraclass correlation). Therefore, we have:

- Level-2 intraclass correlation:

$$rho_{Student|School} = \frac{\tau_{u000} + \tau_{r000}}{\tau_{u000} + \tau_{r000} + \sigma^2} = \frac{180.194 + 325.799}{180.194 + 325.799 + 41.649} = 0.924$$

- Level-3 intraclass correlation:

Mixed-effects REML regression		Number of obs		=		2,440		
Group Variable	No. of Groups	Observations per Group						
		Minimum	Average	Maximum				
school	15	80	162.7	248				
student	610	4	4.0	4				
Log restricted-likelihood = -9092.1387				Wald chi2(0)	=	.		
				Prob > chi2	=	.		
performance	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]			
_cons	68.71395	3.553167	19.34	0.000	61.74987	75.67803		
Random-effects Parameters		Estimate	Std. Err.	[95% Conf. Interval]				
school: Identity								
	var(_cons)	180.1941	71.60437	82.69809	392.6319			
student: Identity								
	var(_cons)	325.7989	19.49574	289.7436	366.3408			
	var(Residual)	41.6494	1.376887	39.03632	44.43739			
LR test vs linear model: chi2(2) = 4036.13				Prob > chi2 = 0.0000				

Figure 5.
Null model

$$\rho_{school} = \frac{\tau_{u000}}{\tau_{u000} + \tau_{r000} + \sigma^2} = \frac{180.194}{180.194 + 325.799 + 41.649} = 0.329$$

Hence, the correlation between annual school performances is equal to 32.9 per cent (ρ_{school}) for the same school, and the correlation between annual school performances is equal to 92.4 per cent ($\rho_{student|school}$) for the same student of a certain school. Therefore, for the model without explanatory variables, while the annual school performance is slightly correlated between schools, the same becomes strongly correlated when the calculation is carried out for the same student from a certain school. In this last case, we estimate that students and schools random effects representing approximately 92 per cent of the total variance of the residuals[2].

Regarding the statistical significance of these variances, the fact that the estimated values of τ_{u000} , τ_{r000} and σ^2 are considerably higher than their respective standard errors suggests that there is significant variation in the annual school performance between students and between schools.

This information is essential to underpin the choice of the multilevel modeling, instead of a simple and traditional regression model through OLS. At the bottom of Figure 5, we can verify this fact by analyzing the result of the likelihood-ratio test. Given $Sig. \chi^2 = 0.000$, we can reject the null hypothesis that the random intercepts are equal to zero ($H_0: u_{00k} = r_{0jk} = 0$), which makes the estimation of a traditional linear regression model be ruled out for the data with repeated measures.

Even though researchers frequently ignore the estimation of null models, analyzing the results may help to reject the research hypotheses or not. It may even provide adjustments in relation to the constructs proposed. For our data, the results of the null model allow us to state that there is significant variability in the school performance throughout the four years under analysis. Furthermore, there is significant variability in the school performance over time between students of the same school, and there is significant variability in the school performance over time between students from different schools.

H1 can be supported, i.e. there is significant variance in student's school performance over time within both students and schools. Since our main objective is to verify if there are student and school characteristics that would explain the variability in the school performance between students from the same school and between those from different schools, we will continue with the next modeling steps, respecting the step-up strategic multilevel analysis.

Let us insert level-1 variable *year* into the analysis, aiming at investigating if the temporal variable has a relationship to students' school performance behavior and, more than this if the school performance has a linear behavior throughout time.

Linear Trend Model with Random Intercepts:

$$performance_{tjk} = \pi_{0jk} + \pi_{1jk} \cdot year_{jk} + e_{tjk}$$

$$\pi_{0jk} = b_{00k} + r_{0jk}$$

$$\pi_{1jk} = b_{10k}$$

$$b_{00k} = \gamma_{000} + u_{00k}$$

$$b_{10k} = \gamma_{100}$$

which results in the following expression[1]:

$$performance_{tjk} = \gamma_{000} + \gamma_{100} \cdot year_{jk} + u_{00k} + r_{0jk} + e_{tjk}$$

First, we can see that the mean of the annual increase in school performance is statistically significant, with an estimated parameter of $\gamma_{100} = 4.348$, *ceteris paribus*. Thus, we can also support *H2*, since student's school performance over time statistically follows a linear trend.

Regarding the random effects components, we have also verified that there is statistical significance in the variances of u_{00k} , r_{0jk} and e_{tjk} , because the estimations of τ_{u000} , τ_{r000} and σ^2 are considerably higher than the respective standard errors. Therefore, new intraclass correlations can be calculated, as follows:

- Level-2 intraclass correlation:

$$rho_{student|school} = \frac{\tau_{u000} + \tau_{r000}}{\tau_{u000} + \tau_{r000} + \sigma^2} = \frac{180.196 + 333.675}{180.196 + 333.675 + 10.146} = 0.981$$

- Level-3 intraclass correlation:

$$rho_{school} = \frac{\tau_{u000}}{\tau_{u000} + \tau_{r000} + \sigma^2} = \frac{180.196}{180.196 + 333.675 + 10.146} = 0.344$$

Both variance proportions are higher than the ones obtained in the estimation of the null model, which shows the importance of including the variable that corresponds to the repeated measure in level 1. Besides, the result of the likelihood-ratio test at the bottom of Figure 6 allows us to prove that the estimation of a simple traditional linear regression model (*performance* based on *year*) only with fixed effects must be ruled out (*H3* supported).

Therefore, now, our model starts to have the following specification:

$$performance_{tjk} = 57.844 + 4.348 \cdot year_{jk} + u_{00k} + r_{0jk} + e_{tjk}$$

Figures 7 and 8 provide better visualization of the random intercepts per school and per student.

Therefore, we are able to conclude that students' school performance follows a linear trend throughout time. In addition, there is a significant variance of intercepts between those who study at the same school and between those who study at different schools.

Thus, we also need to verify if there is a significant variance of the school performance slopes throughout time between the different students. Therefore, let us insert slope random effects into Levels 2 and 3 of our multilevel model that, by maintaining the intercept random effects, will have the following expression:

Linear Trend Model with Random Intercepts and Slopes:

$$performance_{tjk} = \pi_{0jk} + \pi_{1jk} \cdot year_{jk} + e_{tjk}$$

$$\pi_{0jk} = b_{00k} + r_{0jk}$$

Log restricted-likelihood = -7801.4202				Wald chi2(1) = 5683.02		
				Prob > chi2 = 0.0000		
performance	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
year	4.348016	0.0576768	75.39	0.000	4.234972	4.461061
_cons	57.84391	3.556109	16.27	0.000	50.87407	64.81376
Random-effects Parameters		Estimate	Std. Err.	[95% Conf. Interval]		
school: Identity						
	var(_cons)	180.1959	71.60532	82.69876	392.6368	
student: Identity						
	var(_cons)	333.6753	19.49293	297.5759	374.1539	
	var(Residual)	10.14618	.3355141	9.509446	10.82556	
LR test vs linear model: chi2(2) = 6505.83				Prob > chi2 = 0.0000		

Figure 6.
Linear trend model
with random
intercepts



Figure 7.
Random intercepts per school

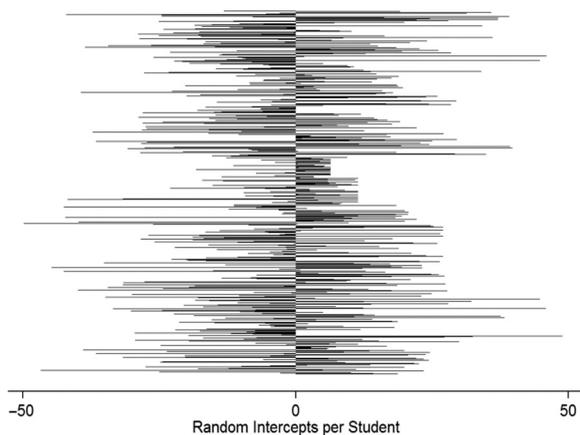


Figure 8.
Random intercepts per student

$$\pi_{1jk} = b_{10k} + r_{1jk}$$

$$b_{00k} = \gamma_{000} + u_{00k}$$

$$b_{10k} = \gamma_{100} + u_{10k}$$

which results in[1]:

$$performance_{tjk} = \gamma_{000} + \gamma_{100} \cdot year_{jk} + u_{00k} + u_{10k} \cdot year_{jk} + r_{0jk} + r_{1jk} \cdot year_{jk} + e_{tjk}$$

Note that the variable *year* is present in the fixed effects component and in the Level-3 random effects components (by multiplying error term u_{10k}), and in the Level-2 ones (by multiplying error term r_{1jk}). [Figure 9](#) shows the results obtained through this estimation.

We can see that, even though the fixed-effects parameter estimations do not change considerably in relation to the previous model, the variance estimations are different, which generates new intraclass correlations, as follows:

- Level-2 intraclass correlation:

$$\begin{aligned} rho_{student|school} &= \frac{\tau_{u000} + \tau_{u100} + \tau_{r000} + \tau_{r100}}{\tau_{u000} + \tau_{u100} + \tau_{r000} + \tau_{r100} + \sigma^2} \\ &= \frac{224.343 + 0.560 + 374.285 + 3.157}{224.343 + 0.560 + 374.285 + 3.157 + 3.868} = 0.994 \end{aligned}$$

- Level-3 intraclass correlation:

$$\begin{aligned} rho_{school} &= \frac{\tau_{u000} + \tau_{u100}}{\tau_{u000} + \tau_{u100} + \tau_{r000} + \tau_{r100} + \sigma^2} \\ &= \frac{224.343 + 0.560}{224.343 + 0.560 + 374.285 + 3.157 + 3.868} = 0.371 \end{aligned}$$

Therefore, for this model, we estimate that the students and schools random effects represent approximately 99 per cent of the total variance of the residuals.

[Figure 10](#) shows a likelihood-ratio test applied to compare the estimations of the linear trend models with random intercepts and with random intercepts and slopes.

Log restricted-likelihood = -7464.819		Wald chi2(1) = 424.89	Prob > chi2 = 0.0000		
performance	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
year	4.343297	0.2107073	20.61	0.000	3.930318 4.756276
_cons	57.85776	3.955816	14.63	0.000	50.1045 65.61102
Random-effects Parameters		Estimate	Std. Err.	[95% Conf. Interval]	
school: Independent					
	var(year)	0.5600495	0.2519118	0.2319283	1.352381
	var(_cons)	224.3434	88.72199	103.344	487.014
student: Independent					
	var(year)	3.157275	.2305444	2.736261	3.643067
	var(_cons)	374.2847	22.00905	333.5408	420.0058
	var(Residual)	3.867725	.1595253	3.567365	4.193374
LR test vs linear model: chi2(4) = 7179.03				Prob > chi2 = 0.0000	

Figure 9.
Linear trend model
with random
intercepts and slopes

By using the values of the restricted likelihood functions obtained in Figures 5 and 9, we arrive at the χ^2 statistic for the test, with 2 degrees of freedom:

$\chi^2_2 = (-2 \cdot LL_{r-randomintercept} - (-2 \cdot LL_{r-randomslope})) = \{-2 \cdot (-7,801.420) - [-2 \cdot (-7,464.819)]\} = 673.20$ which results in a Sig. $\chi^2_2 = 0.000 < 0.05$ and ends up favoring the linear trend model with and random intercepts and slopes. It is important to mention that this likelihood-ratio test is only valid when a comparison of the estimations obtained through REML of two models is carried out with identical fixed effects specification. Given that in our case, both models, that were estimated through REML, present the same fixed effects specification, $\gamma_{000} + \gamma_{100} \cdot year_{jk}$, the test is considered valid.

Hence, our model starts to have the following specification:

$$performance_{ijk} = 57.858 + 4.343 \cdot year_{jk} + u_{00k} + u_{10k} \cdot year_{jk} + r_{0jk} + r_{1jk} \cdot year_{jk} + e_{ijk}$$

In the current situation, we are able to state that students' school performance follows a linear trend throughout time. In addition, there is a significant variance of intercepts and slopes between those students who study at the same school and between those who study at different schools. Therefore, let us insert Level-2 variable *gender* into the analysis to determine if this characteristic explains the variation in the annual school performance between students.

Linear Trend Model with Random Intercepts and Slopes and with Level-2 Variable gender.

$$performance_{ijk} = \pi_{0jk} + \pi_{1jk} \cdot year_{jk} + e_{ijk}$$

$$\pi_{0jk} = b_{00k} + b_{01k} \cdot gender_{jk} + r_{0jk}$$

$$\pi_{1jk} = b_{10k} + b_{11k} \cdot gender_{jk} + r_{1jk}$$

$$b_{00k} = \gamma_{000} + u_{00k}$$

$$b_{01k} = \gamma_{010}$$

$$b_{10k} = \gamma_{100} + u_{10k}$$

$$b_{11k} = \gamma_{110}$$

which results in the following expression[1]:

$$performance_{ijk} = \gamma_{000} + \gamma_{100} \cdot year_{jk} + \gamma_{010} \cdot gender_{jk} + \gamma_{110} \cdot gender_{jk} \cdot year_{jk} + u_{00k} + u_{10k} \cdot year_{jk} + r_{0jk} + r_{1jk} \cdot year_{jk} + e_{ijk}$$

Likelihood-ratio test	LR chi2 (2) =	673.20
(Assumption: randomintercept nested in randomslope)	Prob > chi2 =	0.0000

Figure 10. Likelihood-ratio test

This model shows significant estimations for the fixed effects parameters, as well as for the variances of the random effects terms, at a significance level of < 0.05 (Figure 11). Moreover, at this moment of the modeling, we are able to state that students' school performance follows a linear trend throughout time, and there is a significant variance of intercepts and slopes between those who study at the same school and between those who study at different schools. Additionally, the fact that a certain student is female or male is part of the reason why there is this variation in school performance (*H4* supported).

The model then has the following specification:

$$performance_{tjk} = 64.498 + 4.029 \cdot year_{jk} - 15.033 \cdot gender_{jk} + 0.705 \cdot gender_{jk} \cdot year_{jk} + u_{00k} + u_{10k} \cdot year_{jk} + r_{0jk} + r_{1jk} \cdot year_{jk} + e_{tjk}$$

and, from which we can see that male students (dummy *gender* = 1) have worse performance than female students, on average and *ceteris paribus*.

Finally, let us investigate if Level-3 variable *texp* (professors' years of teaching experience), also explains the variation in the annual school performance between the students. After some intermediate analyses, let us move on to estimate the three-level hierarchical model with the following specification:

Linear Trend Model with Random Intercepts and Slopes, Level-2 Variable gender and Level-3 Variable texp (Complete Model):

$$performance_{tjk} = \pi_{0jk} + \pi_{1jk} \cdot year_{jk} + e_{tjk}$$

$$\pi_{0jk} = b_{00k} + b_{01k} \cdot gender_{jk} + r_{0jk}$$

Log restricted-likelihood = -7424.2732		Wald chi2(3) = 633.54	Prob > chi2 = 0.0000		
performance	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
year	4.028844	0.2024281	19.90	0.000	3.632092 4.425595
gender	-15.03265	1.766749	-8.51	0.000	-18.49542 -11.56989
genderyear	0.7050945	0.1827647	3.86	0.000	0.3468824 1.063307
_cons	64.49828	3.465572	18.61	0.000	57.70589 71.29068
Random-effects Parameters		Estimate	Std. Err.	[95% Conf. Interval]	
school: Independent					
	var(year)	0.4113062	0.1977923	0.1602627	1.055597
	var(_cons)	161.6346	64.79808	73.67059	354.6293
student: Independent					
	var(year)	3.096463	0.2272074	2.681685	3.575395
	var(_cons)	337.7062	19.90229	300.867	379.0562
	var(Residual)	3.867745	0.1594995	3.567432	4.193339
LR test vs linear model: chi2(4) = 6850.06				Prob > chi2 = 0.0000	

Figure 11.
Linear trend model with random intercepts and slopes and Level-2 variable *gender*

$$\pi_{1jk} = b_{10k} + b_{11k} \cdot gender_{jk} + r_{1jk}$$

$$b_{00k} = \gamma_{000} + \gamma_{001} \cdot texp_k + u_{00k}$$

$$b_{01k} = \gamma_{010}$$

$$b_{10k} = \gamma_{100} + \gamma_{101} \cdot texp_k + u_{10k}$$

$$b_{11k} = \gamma_{110}$$

which results in the following expression[1]:

$$\begin{aligned} performance_{ijk} = & \gamma_{000} + \gamma_{100} \cdot year_{jk} + \gamma_{010} \cdot gender_{jk} + \gamma_{001} \cdot texp_k \\ & + \gamma_{110} \cdot gender_{jk} \cdot year_{jk} + \gamma_{101} \cdot texp_k \cdot year_{jk} + u_{00k} \\ & + u_{10k} \cdot year_{jk} + r_{0jk} + r_{1jk} \cdot year_{jk} + e_{ijk} \end{aligned}$$

Even though the estimations of the fixed effects parameters and random effects variances are significant, at a significance level of 0.05, it is necessary to study the structure of the random effects (u_{00k} , u_{10k} and r_{0jk} , r_{1jk}) variance-covariance matrix. Based on the outputs found in Figure 12, we have:

Log restricted-likelihood = -7419.6785		Wald chi2 (5)	=	883.26	
		Prob > chi2	=	0.0000	
performance	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
year	4.528292	.2586443	17.51	0.000	4.021359 5.035226
gender	-14.69529	1.762759	-8.34	0.000	-18.15024 -11.24035
texp	1.179424	0.343969	3.43	0.001	0.5052567 1.85359
genderyear	0.6485018	0.1828469	3.55	0.000	0.2901286 1.006875
texpyear	-0.0570213	0.0211086	-2.70	0.007	-0.0983934 -0.0156491
_cons	54.72215	3.925206	13.94	0.000	47.02889 62.41541
Random-effects Parameters		Estimate	Std. Err.	[95% Conf. Interval]	
school: Independent					
var(year)		0.262667	0.1394859	0.0927653	0.7437469
var(_cons)		87.99372	37.97699	37.7645	205.031
student: Independent					
var(year)		3.092474	0.2267586	2.678496	3.570436
var(_cons)		337.6269	19.89377	300.8031	378.9587
var(Residual)		3.867764	0.1595005	3.567449	4.19336
LR test vs linear model: chi2 (4) = 6557.63				Prob > chi2 = 0.0000	

Figure 12. Linear trend model with random intercepts and slopes and Level-2 variable *gender* and Level-3 variable *texp*

- Random effects variance-covariance matrix for level *school*:

$$\text{var} \begin{bmatrix} u_{00k} \\ u_{10k} \end{bmatrix} = \begin{bmatrix} 87.994 & 0 \\ 0 & 0.263 \end{bmatrix}$$

- Random effects variance-covariance matrix for level *student*:

$$\text{var} \begin{bmatrix} r_{0jk} \\ r_{1jk} \end{bmatrix} = \begin{bmatrix} 337.627 & 0 \\ 0 & 3.092 \end{bmatrix}$$

Since we did not specify any covariance structure for these error terms, we are assuming that this structure is independent, that is, both $\text{cov}(u_{00k}, u_{10k}) = 0$ and $\text{cov}(r_{0jk}, r_{1jk}) = 0$. Nevertheless, we can generalize the structure of these matrices by allowing u_{00k} and u_{10k} to be correlated, and r_{0jk} and r_{1jk} to be correlated too. Thus, following [Short et al. \(2006\)](#), one additional contribution of this paper is the testing of alternative covariance structures to account for heteroskedasticity. In our sample, we found significant differences based on the model assumptions and covariance structure specified in our empirical test.

Log restricted-likelihood = -7376.7147				Wald chi2(5)	=	868.08
				Prob > chi2	=	0.0000
performance	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
year	4.515641	0.2583749	17.48	0.000	4.009236	5.022047
gender	-14.70213	1.795536	-8.19	0.000	-18.22131	-11.18294
texp	1.178656	0.3459065	3.41	0.001	0.5006918	1.856621
genderyear	0.6518855	0.1847166	3.53	0.000	0.2898477	1.013923
texpyear	-0.0566496	0.0209988	-2.70	0.007	-0.0978065	-0.0154928
_cons	54.73435	3.951437	13.85	0.000	46.98968	62.47902
Random-effects Parameters		Estimate	Std. Err.	[95% Conf. Interval]		
school: Unstructured						
var(year)		0.2554224	0.1378072	0.0887183	0.7353682	
var(_cons)		88.7366	38.40337	37.99447	207.2456	
cov(year,_cons)		-3.185306	1.904226	-6.91752	0.5469079	
student: Unstructured						
var(year)		3.2575	0.2350138	2.827965	3.752276	
var(_cons)		350.9127	20.68884	312.6185	393.8978	
cov(year,_cons)		-13.25089	1.673704	-16.53129	-9.970493	
var(Residual)		3.795043	0.1536567	3.505521	4.108476	
LR test vs linear model: chi2(6) = 6643.55				Prob > chi2 = 0.0000		

Figure 13. Outputs of the linear trend model with random intercepts and slopes and Level-2 variable *gender* and Level-3 variable *texp*, with correlated random effects (u_{00k} , u_{10k}) and (r_{0jk} , r_{1jk})

The results obtained through the estimation considering correlated random effects (u_{00k} , u_{10k}) and (r_{0jk} , r_{1jk}) are shown in [Figure 13](#).

The fixed effects parameter estimations are close to those obtained when estimating the model that considers the existence of a structure that is independent from the random-effects variance-covariance matrices ([Figure 12](#)). Regarding the random-effects parameters, except for the estimations of u_{10k} and $\text{cov}(u_{00k}, u_{10k})$, which are statistically significant at a significance level of 0.10, all the other estimations are significant at a significance level of 0.05. Considering that $\text{cov}(u_{00k}, u_{10k})$ and $\text{cov}(r_{0jk}, r_{1jk})$ are statistically different from zero, based on the outputs in [Figure 13](#) we can write that:

- Random effects variance-covariance matrix for level *school*:

$$\text{var} \begin{bmatrix} u_{00k} \\ u_{10k} \end{bmatrix} = \begin{bmatrix} 88.737 & -3.185 \\ -3.185 & 0.255 \end{bmatrix}$$

- Random effects variance-covariance matrix for level *student*:

$$\text{var} \begin{bmatrix} r_{0jk} \\ r_{1jk} \end{bmatrix} = \begin{bmatrix} 350.913 & -13.251 \\ -13.251 & 3.258 \end{bmatrix}$$

Even statistically different from zero, the estimations of the random effects covariances in both levels of the analysis, if researchers wish to prove the better suitability of this last model over the one that considers the matrix with independent error terms, they just need to run a likelihood-ratio test to compare both estimations ([Figure 14](#)).

[Figure 15](#) presents the result of the likelihood-ratio test applied to compare the estimations of the complete models with independent and correlated random effects (u_{00k} , u_{10k}) and (r_{0jk} , r_{1jk}).

Random-effects covariance matrix for level school		
	year	_cons
year	.2554224	
_cons	-3.185306	88.7366

Random-effects covariance matrix for level student		
	year	_cons
year	3.2575	
_cons	-13.25089	350.9127

Figure 14. Variance-covariance matrices with correlated random effects (u_{00k} , u_{10k}) and (r_{0jk} , r_{1jk})

Likelihood-ratio test	LR chi2(2) =	85.93
(Assumption: finalindependent nested in finalunstructured)	Prob > chi2 =	0.0000

Figure 15. Likelihood-ratio test

As *Sig.* $\chi^2_2 = 0.000 < 0.05$, we can state that the structure of the random effects variance-covariance matrices can be considered unstructured. That is, we can consider that error terms u_{00k} and u_{10k} are correlated ($\text{cov}(u_{00k}, u_{10k}) \neq 0$) and that error terms r_{0jk} and r_{1jk} are correlated too ($\text{cov}(r_{0jk}, r_{1jk}) \neq 0$).

We have arrived at our final model, with the following specification:

$$\begin{aligned} performance_{tjk} = & 54.734 + 4.516 \cdot year_{jk} - 14.702 \cdot gender_{jk} + 1.179 \cdot texp_k + 0.652 \\ & \cdot gender_{jk} \cdot year_{jk} - 0.057 \cdot texp_k \cdot year_{jk} + u_{00k} + u_{10k} \cdot year_{jk} + r_{0jk} \\ & + r_{1jk} \cdot year_{jk} + e_{tjk} \end{aligned}$$

Besides, the expected values of each student's school performance in each of the periods monitored are given by:

$$\begin{aligned} performance_student_{tjk} = & 54.734 + 4.516 \cdot year_{jk} - 14.702 \cdot gender_{jk} + 1.179 \cdot texp_k \\ & + 0.652 \cdot gender_{jk} \cdot year_{jk} - 0.057 \cdot texp_k \cdot year_{jk} + u_{00k} + u_{10k} \\ & \cdot year_{jk} + r_{0jk} + r_{1jk} \cdot year_{jk} \end{aligned}$$

On the other hand, the expected values of each student's school performance in each of the periods monitored, without considering the random effects in the level *student*, can be given by:

$$\begin{aligned} performance_school_k = & 54.734 + 4.516 \cdot year_{jk} - 14.702 \cdot gender_{jk} + 1.179 \cdot texp_k \\ & + 0.652 \cdot gender_{jk} \cdot year_{jk} - 0.057 \cdot texp_k \cdot year_{jk} + u_{00k} + u_{10k} \\ & \cdot year_{jk} \end{aligned}$$

We have seen that students' school performance follows a linear trend over time. Moreover, there is a significant variance of intercepts and slopes between those who study at the same school and between those who study at different schools, and students' gender is significant to explain part of this variation. *H5* can be supported since professors' years of teaching experience at each school (Level-3 variable) itself also explains part of the discrepancies in the annual school performance between students from different schools.

Figure 16 shows the predicted values of school performance throughout time for the first 50 students in the sample (*yhatstudent*) and, through which, we can see different intercepts and slopes throughout time for different students.

Finally, a more inquisitive researcher, aiming at questioning the superiority of multilevel models in relation to traditional regression models estimated through OLS, whenever there are datasets with nested structures, decides to construct a graph (Figure 17). Through this graph, it is possible to compare the predicted school performance values generated by this three-level hierarchical modeling (HLM3) to those generated by an estimation through OLS, for all the students in the sample, in each of the periods analyzed, using the same explanatory variables *year*, *gender*, *texp*, *genderyear* and *texpyear*. Obviously, there are only fixed effects in the estimation through OLS.

The gray line at 45° shows the observed school performance values of each one of the students in the sample in each of the periods analyzed (*performance x performance*). By

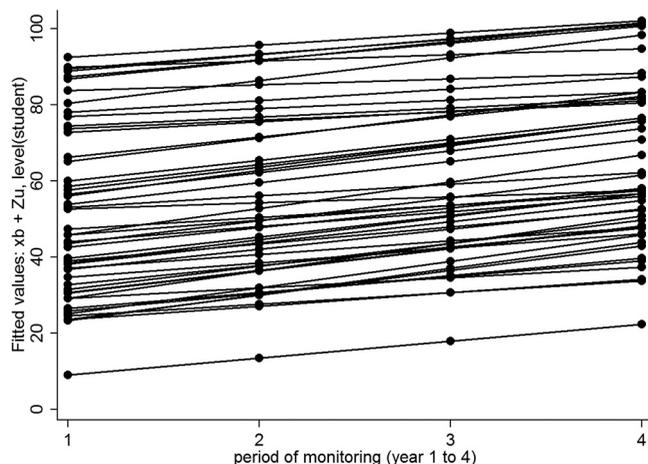


Figure 16. Predicted school performance values throughout time for the first 50 students in the sample

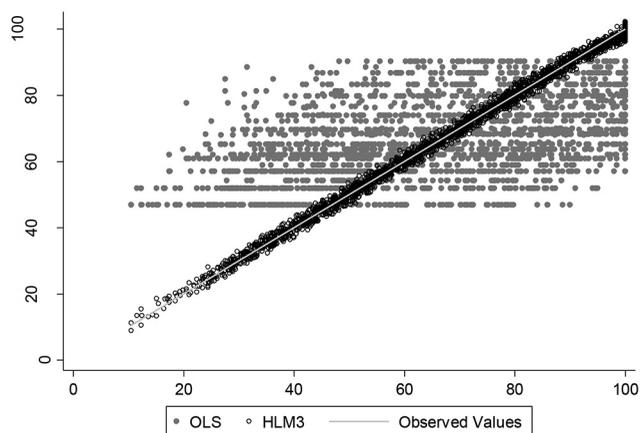


Figure 17. Values predicted through OLS and through HLM3 \times observed school performance values

using [Figure 17](#), we can clearly see the superiority of the linear trend model with explanatory variables and random intercepts and slopes in levels 2 and 3 (complete HLM3 model) over the multiple linear regression model estimated through OLS with the same explanatory variables in this case. The absence of random terms for each contextual effect in traditional regression models, such as OLS, prevents greater adherence between predicted and observed values of the outcome variable in phenomena where it is possible to directly or indirectly identify hierarchies, or levels, in the data structure ([Raudenbush & Bryk, 2002](#)). This demonstrates the importance of considering the random effects components whenever there are nested data structures.

4. Final remarks

This paper provides a brief discussion about the concepts, processes, stages, tasks, and the types of methods and techniques it can employ. It enables researchers and managers to assess the relationship between a certain performance variable and one or more predictor variables, which characterize different levels of analysis.

Moreover, as well as contributed and discussed by [Short *et al.* \(2006\)](#), this study offers progress toward resolving the ambiguity related to the structure of the random effects variance-covariance matrices by applying a multilevel model with random intercepts and slopes, by explicitly accounting for and modeling heteroskedasticity in the data analysis, and by articulating the importance of and offering interpretations for different specifications to testing the effects of time on performance. Time studies in business management research have not reflected enough attention to these issues, and multilevel modeling provides a tool to ameliorate such issues when using a longitudinal design.

Researchers can also estimate hierarchical cross-classified models (HCM) and multilevel multiple membership classification models (MMMM) in situations where lower level observations are nested within multiple higher-level units from the same classification. According to [Durrant, Vassallo, and Smith \(2018\)](#), not accounting correctly for such multiple membership structures leads to biased results. Researchers can study HCM and MMMC in-depth in [Browne, Goldstein, and Rasbash \(2001\)](#), [Meyers and Beretvas \(2006\)](#), [Uyar and Brown \(2007\)](#), [Fávero \(2011\)](#), [Rabe-Hesketh and Skrondal \(2012a, 2012b\)](#), [Chung and Beretvas \(2012\)](#), [Brunton-Smith, Sturgis, and Leckie \(2017\)](#), and [Fávero, Serra, Santos, and Brunaldi \(2018b\)](#).

Multilevel modeling is a broad theme that can often be explored in depth in the field of business management. Each level is formed by individuals or groups nested into other groups and so on. Since variables from a certain group are invariable between groups or individuals that correspond to lower levels that are nested into that group, it is natural for many researchers and constructs to use such models ([Zhang, Li, & Song, 2014](#)).

Many can be the characteristics of the datasets with nested data structures. The most common are those with absolute nesting, in which there are clustered data or data with repeated measures. In this paper, we chose to use a dataset to estimate three-level hierarchical linear models with repeated measures. Nonetheless, from which, we believe researchers will have the conditions to estimate, for example, for three-level models with clustered data or even to consider a higher number of analysis levels, resulting from more complex nesting structures.

Multilevel models enable us to identify and analyze individual heterogeneities and the heterogeneities between the groups to which these individuals belong, making it possible to specify random components in each analysis level. This fact represents the main difference of the traditional regression models estimated through OLS, which cannot consider the natural nesting of data and, consequently, generate biased parameter estimators ([Fávero & Belfiore, 2017](#); [Lazega & Snijders, 2016](#); [Pinheiro & Bates, 2009](#)).

Although many papers use multilevel models only to estimate null models to investigate the variance decomposition of the phenomenon being studied in the different analysis levels, the possibility of including explanatory variables that correspond to the different levels in the fixed and random effects components enables us to investigate possible relationships between these variables and the dependent variable. This makes it possible to establish and examine new research objectives and interesting constructs ([Gelman, 2006](#)).

The models we studied in this paper are part of what we call generalized linear latent and mixed models (GLLAMM), which encompass the hierarchical linear models (HLM). In this sense, as discussed by [Fávero, Santos, and Serra \(2018a\)](#), a researcher can also propose hierarchical nonlinear models, that refer to the situations in which, in a nested data structure, the dependent variable is a categorical variable or a variable with count data. Thus, logistic, Poisson or negative binomial multilevel models could also be estimated.

Discovering implicit and contextual standards from larger and larger volumes of data becomes an essential condition for organizations to become successful in competitive environments, and multilevel modeling contributes in a considerable way to our ability to understand phenomena.

Notes

1. All parameters are defined in the [Appendix 1](#).
2. One might notice that the sum of variance percentages is not equal to 1. This is the default for displaying Stata outputs. In order to obtain the variance percentages of each level on the outcome variable, we must proceed with the following calculations:

Level-1 intraclass correlation:

$$\rho_{time} = \frac{41.649}{180.194 + 325.799 + 41.649} = 0.076$$

Level-2 intraclass correlation:

$$\rho_{student} = \frac{325.799}{180.194 + 325.799 + 41.649} = 0.595$$

Level-3 intraclass correlation:

$$\rho_{school} = \frac{180.194}{180.194 + 325.799 + 41.649} = 0.329$$

whose sum is 1.

References

- Aitkin, M., & Longford, N. (1986). Statistical modelling in school effectiveness studies (with discussion). *Journal of the Royal Statistical Society (Series A)*, *149*, 1–43.
- Arceneaux, K., & Nickerson, D. W. (2009). Modeling certainty with clustered data: A comparison of methods. *Political Analysis*, *17*, 177–190.
- Bliese, P. D., & Ployhart, R. E. (2002). Growth modeling using random coefficient models: Model building, testing, and illustrations. *Organizational Research Methods*, *5*, 362–387.
- Bowman, E. H., & Helfat, C. E. (2001). Does corporate strategy matter? *Strategic Management Journal*, *22*, 1–23.
- Browne, W. J., Goldstein, H., & Rasbash, J. (2001). Multiple membership multiple classification (MMMC) models. *Statistical Modelling*, *1*, 103–124.
- Brunton-Smith, I., Sturgis, P., & Leckie, G. (2017). Detecting and understanding interviewer effects on survey data by using a cross-classified mixed effects location-scale model. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *180*, 551–568.
- Brush, T. H., & Bromiley, P. (1997). What does a small corporate effect mean? A variance components simulation of corporate and business effects. *Strategic Management Journal*, *18*, 825–835.
- Brush, T. H., Bromiley, P., & Hendrickx, M. (1999). The relative influence of industry and corporation on business segment performance: An alternative estimate. *Strategic Management Journal*, *20*, 519–547.
- Chang, S., & Singh, H. (2000). Corporate and industry effects on business unit competitive position. *Strategic Management Journal*, *21*, 739–752.
- Christmann, P., Day, D. L., & Yip, G. S. (1999). The relative influence of country conditions, industry structure and business strategy on MNC subsidiary performance. *Journal of International Management*, *5*, 241–265.
- Chung, H., & Beretvas, S. N. (2012). The impact of ignoring multiple membership data structures in multilevel models. *British Journal of Mathematical and Statistical Psychology*, *65*, 185–200.
- Courgeau, D. (2003). *Methodology and epistemology of multilevel analysis*, London, United Kingdom: Kluwer Academic Publishers.
- Durrant, G. B., Vassallo, R., & Smith, P. W. F. (2018). Assessment of multiple membership multilevel models: An application to interviewer effects on nonresponse. *Multivariate Behavioral Research*, *53*, 595–611.
- Fávero, L. P. (2008). Time, firm and country effects on performance: An analysis under the perspective of hierarchical modeling with repeated measures. *Brazilian Business Review*, *5*, 163–180.

- Fávero, L. P. (2010). Modelagem hierárquica com medidas repetidas [Hierarchical modeling with repeated measures]. Tese (Livre-Docência) – Thesis – Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo, p. 202.
- Fávero, L. P. (2011). Preços hedônicos no mercado imobiliário comercial de São Paulo: A abordagem da modelagem multinível com classificação cruzada [Hedonic prices in the commercial real estate market of São Paulo: a cross-classified multilevel modeling approach]. *Estudos Econômicos (São Paulo)*, 41, 777–810.
- Fávero, L. P., & Belfiore, P. (2017). *Manual de análise de dados: Estatística e modelagem multivariada com excel®, SPSS® e stata®*, Rio de Janeiro, Brazil: Elsevier.
- Fávero, L. P., & Belfiore, P. (2019). *Data science for business and decision making*, Cambridge, MA: Academic Press.
- Fávero, L. P., Santos, M. A., & Serra, R. G. (2018a). Cross-border branching in the Latin American banking sector. *International Journal of Bank Marketing*, 36, 496–528.
- Fávero, L. P., Serra, R. G., Santos, M. A., & Brunaldi, E. (2018b). Cross-classified multilevel determinants of firm's sales growth in Latin America. *International Journal of Emerging Markets*, 13, 902–924.
- Garner, C. L., & Raudenbush, S. (1991). Neighborhood effects on educational attainment: a multilevel analysis. *Sociology of Education*, 64, 251–262.
- Gelman, A. (2006). Multilevel (hierarchical) modeling: what it can and cannot do. *Technometrics*, 48, 432–435.
- Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*, New York, NY: Cambridge University Press.
- Goldstein, H. (2011). *Multilevel statistical models*, 4th ed., Chichester, United Kingdom: John Wiley & Sons.
- Goldszmidt, R. G. B., Brito, L. A. L., & Vasconcelos, F. C. (2007). O efeito país sobre o desempenho da firma: uma abordagem multinível [Country effect on firm performance: a multilevel approach]. *Revista de Administração de Empresas*, 47, 12-25.
- Hawawini, G., Subramanian, V., & Verdin, P. (2004). The home country in the age of globalization: How much does it matter for firm performance?. *Journal of World Business*, 39, 121–135.
- Heck, R. H., & Thomas, S. L. (2009). *An introduction to multilevel modeling techniques*, 2nd ed., New York, NY: Routledge.
- Hitt, M. A., Beamish, P. W., Jackson, S. E., & Mathieu, J. E. (2007). Building theoretical and empirical bridges across levels: multilevel research in management. *Academy of Management Journal*, 50, 1385–1399.
- Holcomb, T. R., Combs, J. G., Sirmon, D. G., & Sexton, J. (2010). Modeling levels and time in entrepreneurship research: An illustration with growth strategies and post-IPO performance. *Organizational Research Methods*, 13, 348–389.
- Hough, J. R. (2006). Business segment performance redux: A multilevel approach. *Strategic Management Journal*, 27, 45–61.
- Kreft, I., & De Leeuw, J. (1998). *Introducing multilevel modeling*, London, United Kingdom: Sage Publications.
- Lazega, E., & Snijders, T. A. B. (2016). *Multilevel network analysis for the social sciences: Theory, methods and applications*, New York, NY: Springer.
- Lee, B. H. (2003). Using hierarchical linear modeling to illustrate industry and group effects on organizational commitment in a sales context. *Journal of Managerial Issues*, 15, 353–368.
- Littell, R. C., Milliken, G. A., Stroup, W. W., & Wolfinger, R. D. (2004). *SAS system for mixed models*, Cary, NC: SAS Institute.
- Makino, S., Beamish, P. W., & Zhao, N. B. (2004a). The characteristics and performance of Japanese FDI in less developed and developed countries. *Journal of World Business*, 39, 377–392.
- Makino, S., Isobe, T., & Chan, C. M. (2004b). Does country matter? *Strategic Management Journal*, 25, 1027–1043.

-
- Martins, H. C., & Terra, P. R. S. (2015). Maturidade do endividamento, desenvolvimento financeiro e instituições legais: Análise multinível em empresas latino-americanas [Debt maturity, financial development and legal institutions: a multilevel analysis in Latin American companies]. *Revista de Administração, 50*, 381–394.
- Mathieu, J. E., & Chen, G. (2011). The etiology of the multilevel paradigm in management research. *Journal of Management, 37*, 610–641.
- Mauri, A. J., & Michaels, M. P. (1998). Firm and industry effects within strategic management: An empirical examination. *Strategic Management Journal, 19*, 211–219.
- McGahan, A. M., & Porter, M. E. (1997). How much does industry matter, really? *Strategic Management Journal, 18*, 15–30.
- McGahan, A. M., & Porter, M. E. (2002). What do we know about variance in accounting profitability? *Management Science, 48*, 834–851.
- McNeish, D. M., & Stapleton, L. M. (2016). The effect of small sample size on two-level model estimates: a review and illustration. *Educational Psychology Review, 28*, 295–314.
- Meyers, J. L., & Beretvas, S. N. (2006). The impact of inappropriate modeling of crossclassified data structures. *Multivariate Behavioral Research, 41*, 473–497.
- Misangyi, F., Lepine, J. A., Algina, J., & Goeddeke, F. Jr. (2006). The adequacy of repeated-measures regression for multilevel research. *Organizational Research Methods, 9*, 5–28.
- O'Connell, A. A. & McCoach, D. B. (Ed.) (2008). *Multilevel modeling of educational data*, Charlotte, NC: Information Age Publishing.
- Pinheiro, J. C., & Bates, D. M. (2009). *Mixed-effects models in S and S-PLUS*, New York, NY: Springer.
- Rabe-Hesketh, S., & Skrondal, A. (2012a). *Multilevel and longitudinal modeling using Stata: Continuous responses* (3rd ed., Vol. 1). College Station, TX: Stata Press.
- Rabe-Hesketh, S., & Skrondal, A. (2012b). *Multilevel and longitudinal modeling using Stata: Categorical responses, counts, and survival* (3rd ed., Vol. 2). College Station, TX: Stata Press.
- Raudenbush, S. (1993). A crossed random effects model for unbalanced data with applications in cross-sectional and longitudinal research. *Journal of Educational Statistics, 18*, 321–349.
- Raudenbush, S., & Bryk, A. (1986). A hierarchical model for studying school effects. *Sociology of Education, 59*, 1–17.
- Raudenbush, S., & Bryk, A. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Raudenbush, S., Bryk, A., Cheong, Y. F., Congdon, R., & Du Toit, M. (2004). *HLM 6: Hierarchical linear and nonlinear modeling*, Lincolnwood, IL: Scientific Software International.
- Raudenbush, S. W., Rowan, B., & Kang, S. J. (1991). A multilevel, multivariate model of studying school climate with estimation via the EM algorithm and application to U.S. high-school data. *Journal of Educational Statistics, 16*, 295–330.
- Ruefli, T. W., & Wiggins, R. R. (2003). Industry, corporate and segment effects and business performance: A non-parametric approach. *Strategic Management Journal, 24*, 861–879.
- Rumberger, R. W., & Thomas, S. L. (1993). The economic returns to college major, quality and performance: A multilevel analysis of recent graduates. *Economics of Education Review, 12*, 1–19.
- Santos, M. A., Fávero, L. P., & Distadio, L. F. (2016). Adoption of the international financial reporting standards (IFRS) on companies' financing structure in emerging economies. *Finance Research Letters, 16*, 179–189.
- Schmalensee, R. (1985). Do markets differ much?. *The American Economic Review, 75*, 341–351.
- Serra, R. G., & Fávero, L. P. (2018). Multiples' valuation: the selection of cross-border comparable firms. *Emerging Markets Finance and Trade, 54*, 1973–1992.

- Short, J. C., Ketchen, D. J., Bennett, N., & Du Toit, M. (2006). An examination of firm, industry, and time effects on performance using random coefficients modeling. *Organizational Research Methods, 9*, 259–284.
- Short, J. C., Ketchen, D. J., Palmer, T. B., & Hult, G. T. M. (2007). Firm, strategic group, and industry influences on performance. *Strategic Management Journal, 28*, 147–167.
- Short, J. C., McKelvie, A., Ketchen, D. J., & Chandler, G. N. (2009). Firm and industry effects on firm performance: A generalization and extension for new ventures. *Strategic Entrepreneurship Journal, 3*, 47–65.
- Singer, J. D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics, 23*, 323–355.
- Snijders, T. A. B., & Bosker, R. J. (2011). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). London, United Kingdom: Sage Publications.
- Steenbergen, M. R., & Jones, B. S. (2002). Modeling multilevel data structures. *American Journal of Political Science, 46*, 218–237.
- Uyar, B., & Brown, K. H. (2007). Neighborhood affluence, school-achievement scores, and housing prices: cross-classified hierarchies and HLM. *Journal of Housing Research, 16*, 97–116.
- West, B. T., Welch, K. B., & Galecki, A. T. (2015). *Linear mixed models: a practical guide using statistical software* (2nd ed.). Boca Raton, FL: Chapman & Hall/CRC Press.
- Zhang, Y., Li, X., & Song, W. (2014). Determinants of cropland abandonment at the parcel, household and village levels in Mountain areas of China: A multilevel analysis. *Land Use Policy, 41*, 186–192.

Appendix 1. Econometrics of the Three-Level hierarchical linear models with repeated measures (HLM3)

In general, according to Raudenbush *et al.* (2004), a three-level hierarchical model has three sub-models, one for each analysis level of the nested data structure. Therefore, we can define a general model with three analysis levels and nested data. The first level presents explanatory variables Z_1, \dots, Z_P that refer to level-1 units i ($i = 1, \dots, n$). The second level, explanatory variables X_1, \dots, X_Q that refer to level-2 units j ($j = 1, \dots, J$). Whereas the third level presents explanatory variables W_1, \dots, W_S that refer to level-3 units k ($k = 1, \dots, K$), as follows:

$$\text{Level 1: } Y_{ijk} = \pi_{0jk} + \sum_{p=1}^P \pi_{pj k} \cdot Z_{pj k} + e_{ijk} \tag{A1}$$

where $\pi_{pj k}$ ($p = 0, 1, \dots, P$) refer to the level-1 coefficients, $Z_{pj k}$ is the p -th level-1 explanatory variable for observation i in the level-2 unit j and in the level-3 unit k , and e_{ijk} refers to the level-1 error terms that follow a normal distribution, with mean equal to zero and variance equal to σ^2 .

$$\text{Level 2: } \pi_{pj k} = b_{p0k} + \sum_{q=1}^{Q_p} b_{pqk} \cdot X_{qjk} + r_{pj k} \tag{A2}$$

where b_{pqk} ($q = 0, 1, \dots, Q_p$) refer to the level-2 coefficients, X_{qjk} is the q -th Level-2 explanatory variable for unit j in the level-3 unit k , and $r_{pj k}$ are the level-2 random effects, assuming, for each unit j , that the vector $(r_{0jk}, r_{1jk}, \dots, r_{Pjk})'$ follows a multivariate normal distribution with each element having mean zero and variance $\tau_{r\pi p p}$.

$$\text{Level 3: } b_{pqk} = \gamma_{pq0} + \sum_{s=1}^{S_{pq}} \gamma_{pqs} \cdot W_{sk} + u_{pqk} \tag{A3}$$

where γ_{pqs} ($s = 0, 1, \dots, S_{pq}$) refer to the level-3 coefficients, W_{sk} is the s -th level-3 explanatory variable for unit k , and u_{pqk} are the level-3 random effects, assuming that, for each unit k , the vector formed by terms u_{pqk} follows a multivariate normal distribution with each element having mean zero and variance $\tau_{u\pi pp}$, which results in a variance-covariance matrix \mathbf{T}_b with a maximum dimension equal to:

$$\text{Dim}_{\max} T_b = \sum_{p=0}^P (Q_p + 1) \cdot \sum_{p=0}^P (Q_p + 1) \tag{A4}$$

which depends on the number of Level-3 coefficients specified with random effects.

Let's imagine a single Level-1 explanatory variable that corresponds to the periods in which the data of the dependent variable are monitored. In other words, Level-2 units j nested into Level-3 units k are monitored for a period t ($t = 1, \dots, T_j$), which makes the dataset have j time series, as shown in Table II. The main objective is to verify if there are discrepancies in the temporal evolution of the data of the dependent variable and, if yes, if these occur due to characteristics of the Level-2 and Level-3 units. This temporal evolution is what characterizes the term repeated measures.

In this regard, Expression (A1) can be rewritten as follows, in which subscripts i become subscripts t :

$$Y_{tjk} = \pi_{0jk} + \pi_{1jk} \cdot \text{period}_{jk} + e_{tjk} \tag{A5}$$

where π_{0jk} represents the intercept of the model that corresponds to the temporal evolution of the dependent variable of Level-2 unit j nested into Level-3 unit k , and π_{1jk} corresponds to the average evolution (slope) of the dependent variable for the same unit throughout the period analyzed. The substructures that correspond to Levels 2 and 3 remain with the same specifications as those respectively presented in Expressions (A2) and (A3).

Figure A1 shows the plotting of the set of models represented by Expression (A5) in a conceptual way. Through the plotting of the models, we can see that the individual models that represent Level-2 units j can present different intercepts and slopes throughout period t . The fact that this may occur is due to certain characteristics of the Level-2 units j themselves or due to characteristics of the Level-3 units k .

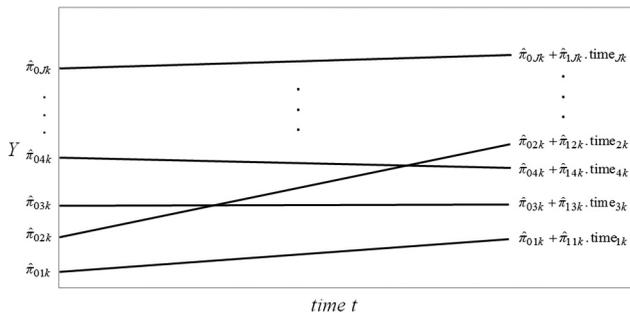


Figure A1. Individual models that represent the temporal evolution of the dependent variable for each of the J Level-2 units

Thus, there must be characteristics of level-2 units j , temporally invariable, and of Level-3 units k , invariable also for Level-2 units j nested into each Level-3 unit k (as shown in Table II), that can explain the differences in the model intercepts and slopes $\hat{Y}_{tjk} = \hat{\pi}_{0jk} + \hat{\pi}_{1jk} \cdot period_{jk}$ represented in Figure A1.

Hence, assuming there is a single explanatory variable X that represents a characteristic of Level-2 units j , and a single explanatory variable W that represents a characteristic of Level-3 units k , from Expression (A5) and based on Expressions (A2) and (A3), we can define the following model with three analysis levels. In this model, the first level refers to the measure repeated and only contains the temporal variable:

$$\text{Level 1: } Y_{tjk} = \pi_{0jk} + \pi_{1jk} \cdot period_{jk} + e_{tjk} \tag{A6}$$

$$\text{Level 2: } \pi_{0jk} = b_{00k} + b_{01k} \cdot X_{jk} + r_{0jk} \tag{A7}$$

$$\pi_{1jk} = b_{10k} + b_{11k} \cdot X_{jk} + r_{1jk} \tag{A8}$$

$$\text{Level 3: } b_{00k} = \gamma_{000} + \gamma_{001} \cdot W_k + u_{00k} \tag{A9}$$

$$b_{01k} = \gamma_{010} + \gamma_{011} \cdot W_k + u_{01k} \tag{A10}$$

$$b_{10k} = \gamma_{100} + \gamma_{101} \cdot W_k + u_{10k} \tag{A11}$$

$$b_{11k} = \gamma_{110} + \gamma_{111} \cdot W_k + u_{11k} \tag{A12}$$

By combining Expressions (A6) to (A12), we obtain the following expression:

$$\begin{aligned}
 Y_{tjk} = & \underbrace{(\gamma_{000} + \gamma_{001} \cdot W_k + \gamma_{010} \cdot X_{jk} + \gamma_{011} \cdot W_k \cdot X_{jk} + u_{00k} + u_{01k} \cdot X_{jk} + r_{0jk})}_{\text{random effects intercept}} \\
 & + \underbrace{(\gamma_{100} + \gamma_{101} \cdot W_k + \gamma_{110} \cdot X_{jk} + \gamma_{111} \cdot W_k \cdot X_{jk} + u_{10k} + u_{11k} \cdot X_{jk} + r_{1jk}) \cdot period_{jk}}_{\text{random effects slope}} \\
 & + e_{tjk} \tag{A13}
 \end{aligned}$$

where γ_{000} represents the expected value of the dependent variable at the initial moment and when $X = W = 0$ (general intercept), γ_{001} represents the increase in the expected value of the dependent variable at the initial moment (alteration in the intercept) for a certain Level-2 unit j that belongs to a Level-3 unit k when there is a unit alteration in the characteristic W of k , *ceteris paribus*.

Besides that, γ_{010} represents the increase in the expected value of the dependent variable at the initial moment for a certain unit jk when there is a unit alteration in the characteristic X of j , *ceteris paribus*, and γ_{011} represents the increase in the expected value of the dependent variable at the initial moment for a certain unit jk when there is a unit alteration in $W \cdot X$, *ceteris paribus*. Moreover, u_{00k} and u_{01k} represent the error terms that indicate there is randomness in the intercepts, and the last one impacts the alterations in variable X .

Additionally, γ_{100} represents the alteration in the expected value of the dependent variable when there is a unit alteration in the analysis period (change in the slope due to a unit temporal evolution), *ceteris paribus*, γ_{101} represents the alteration in the expected value of the dependent variable due to a unit temporal evolution for a certain unit jk when there is a unit alteration in the characteristic W , *ceteris paribus*.

Finally, γ_{110} represents the alteration in the expected value of the dependent variable due to a unit temporal evolution for a certain unit jk when there is a unit alteration in the characteristic X , *ceteris paribus*, and γ_{111} represents the alteration in the expected value of the dependent variable due to a unit temporal evolution for a certain unit jk when there is a unit alteration in $W \cdot X$, *ceteris paribus*. Terms u_{10k} and u_{11k} represent errors that indicate there is randomness in the slopes, and the last one impacts the alterations in variable X .

Expression (A13) facilitates the visualization that the intercept and slope can be influenced by random effects resulting from different behaviors of the dependent variable throughout time for each of the level-2 units (different time series), and this phenomenon can be a result of these units' characteristics, as well as of characteristics of the groups to which such units belong.

If researchers wish to elaborate an analysis about the fixed and random effects components that can influence the behavior of the dependent variable, given that this procedure even facilitates the insertion of the commands to estimate multilevel models in Stata, as we will see below we just need to rearrange the terms of Expression (A13) as follows:

$$\begin{aligned}
 Y_{ijk} = & \underbrace{\gamma_{000} + \gamma_{001} \cdot W_k + \gamma_{010} \cdot X_{jk} + \gamma_{011} \cdot W_k \cdot X_{jk}}_{\text{Fixed Effects}} \\
 & + \underbrace{\gamma_{100} \cdot \text{period}_{jk} + \gamma_{101} \cdot W_k \cdot \text{period}_{jk} + \gamma_{110} \cdot X_{jk} \cdot \text{period}_{jk} + \gamma_{111} \cdot W_k \cdot X_{jk} \cdot \text{period}_{jk}}_{\text{Fixed Effects}} \\
 & + \underbrace{u_{00k} + u_{01k} \cdot X_{jk} + u_{10k} \cdot \text{period}_{jk} + u_{11k} \cdot X_{jk} \cdot \text{period}_{jk} + \Gamma_{0jk} + \Gamma_{1jk} \cdot \text{period}_{jk} + e_{ijk}}_{\text{Random Effects}}
 \end{aligned}
 \tag{A14}$$

As multilevel models allow interactions between variables in the fixed effects component and, more than that, allow interactions between error terms and variables in the random effects component, the model can be naturally heteroskedastic. Regarding Expression (A14), if the variances of the random terms u_{10k} , u_{11k} , r_{0jk} and r_{1jk} are statistically different from zero, traditional parameter estimations, such as OLS, will not be adequate.

In three-level hierarchical models, we can define two intraclass correlations given the existence of two variance proportions. One corresponds to the behavior of the data that belong to the same Level-2 units j and the same Level-3 units k (Level-2 intraclass correlation), and the other corresponds to the behavior of the data that belong to the same Level-3 units k . But the data are from different Level-2 units j (Level-3 intraclass correlation).

While fixed effects parameters are estimated traditionally through maximum likelihood, the variance components of the error terms can be estimated through maximum likelihood or through REML.

Appendix 2. Commands in stata

We now present all Stata commands (Version 15) used throughout the paper to generate tables and graphs and to estimate models:

Figure 3:

```
graph twoway connected performance year if student <= 50,
connect(L)
```

Figure 4:

```
statsby intercept=_b[_cons] slope=_b[year], by(school) saving
(ols, replace): reg performance year
sort school
merge school using ols
drop _merge
gen yhat_ols = intercept + slope*year
sort school year
separate performance, by(school)
separate yhat_ols, by(school)
graph twoway connected yhat_ols1-yhat_ols15 year || lfit
performance year, clwidth(thick) clcolor(black) legend(off)
ytitle(performance at school)
```

Figure 5:

```
xtmixed performance || school: || student:, var nolog reml
```

This command shows two random effects components, one that corresponds to level 3 (school) and another to level 2 (student). The order in which the random effects components are inserted into the command `xtmixed` is decreasing when there are more than two levels. That is, we must begin with the highest data nesting level and continue until the lowest level (level 2):

Obtention of intraclass correlation:

```
estat icc
```

right after the estimation of the corresponding model.

Figure 6:

```
xtmixed performance year || school: || student:, var nolog reml
estimates store randomintercept
predict u00 r0, reffects
```

Figures 7 and 8:

```
graph hbar (mean) u00, over(school) ytitle("Random Intercepts
per School")
graph hbar (mean) r0, over(student) ytitle("Random Intercepts
per Student")
```

Figure 9:

```
xtmixed performance year || school: year || student: year, var
nolog reml
estimates store randomslope
```

Figure 10:

```
lrtest randomslope randomintercept
```

Figure 11:

```
gen genderyear = gender*year
xtmixed performance year gender genderyear || school: year ||
student: year, var nolog reml
```

If the researcher wants to type directly the command without the necessity of previously calculate the product term `gender * year`, the following command can be used alternatively:

```
xtmixed performance year gender i.gender##c.year || school:
year || student: year, var nolog reml
```

Figure 12:

```
gen texpyear = texp*year
xtmixed performance year gender texp genderyear texpyear
|| school: year || student: year, var nolog reml estimates store
finalindependent
```

Figure 13:

```
xtmixed performance year gender texp genderyear texpyear ||
school: year, covariance(unstructured) || student: year, covari-
ance(unstructured) var nolog reml
estimates store finalunstructured predict yhatstudent, fitted
level(student)
```

which defines the variable *yhatstudent*, which can also be obtained through the following command:

```
gen yhatstudent = 54.73435 + 4.515641*year - 14.70213*gender +
1.178656*texp + 0.6518855*genderyear - 0.0566496*texpyear +
u0final + u10final*year + r0final + r1final*year
```

Figure 14:

```
estat recovariance
```

Figure 15:

```
lrtest finalunstructured finalindependent
```

Figure 16:

```
sort student year graph twoway connected yhatstudent year if stu-
dent <= 50, connect(L)
```

Figure 17:

```
quietly reg performance year gender texp genderyear texpyear
predict yhatreg
graph twoway scatter yhatreg performance || scatter yhatstudent
performance || lfit performance performance ||, legend(label(1
"OLS") label(2 "HLM3") label(3 "Observed Values"))
```

Corresponding author

Luiz Paulo Fávero can be contacted at: lpfaver@usp.br

For instructions on how to order reprints of this article, please visit our website:

www.emeraldgrouppublishing.com/licensing/reprints.htm

Or contact us for further details: permissions@emeraldinsight.com