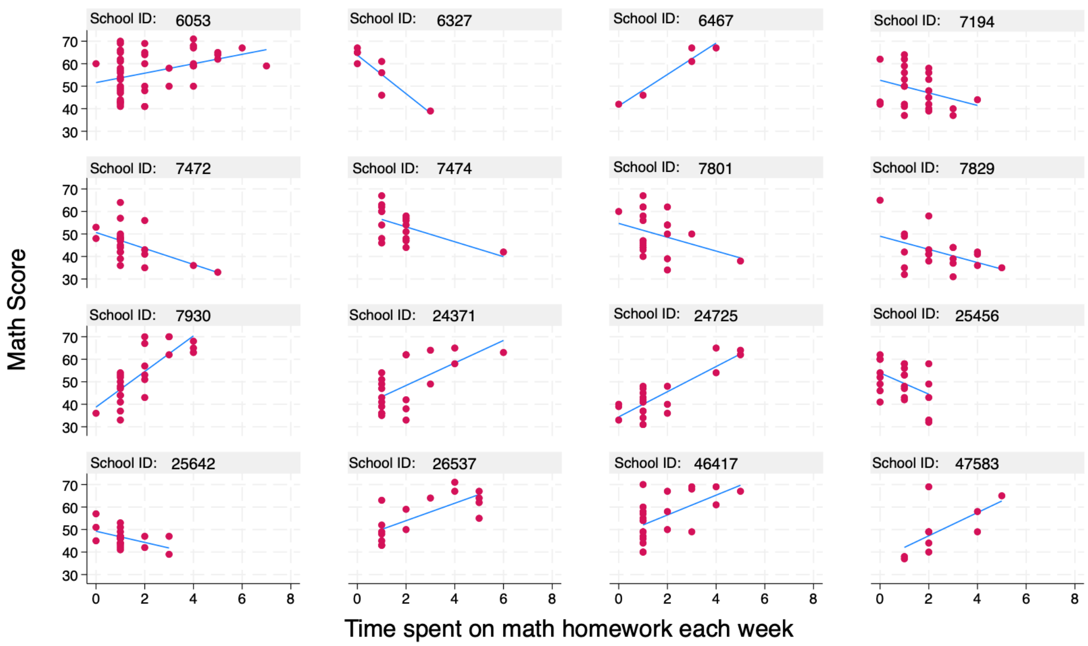
Introduction to multilevel modeling



Meghan K. Cain, Ph.D.

Assistant Director, Educational Services

StataCorp LLC

mcain@stata.com

Society for Prevention Research

September 30, 2025

Copyright ©2025 StataCorp LLC

Multilevel/mixed models using Stata

Contents

List of Figures

Copyright ©2025 StataCorp LLC

1 Course introduction

1.1 What are multilevel/mixed models (MLMMs)?

1.1.1 Other names

- Multilevel models
- Hierarchical linear models (HLM)
- Mixed-effects models
- Mixed models
- Random-effects models
- Variance-component models

1.1.2 What are MLMMs used for?

- MLMMs are used on data with observations that are clustered or grouped. In other words, the data have a natural nesting structure.

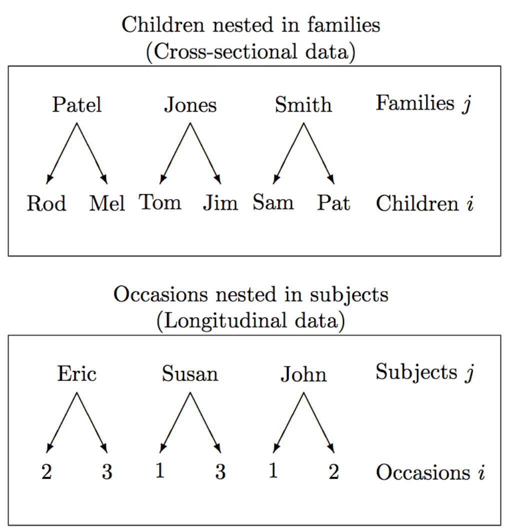


Figure 1: Nesting/clustering data structure (Rabe-Hesketh and Skrondal 2022)

- Some examples of nesting include students within classrooms, patients within doctors, doctors within hospitals, clients within therapists, or workers within companies or teams. Longitudinal or panel data can also be considered a type of nesting where measurement occasions are nested in individuals or other units.
- MLMMs are models that account for the correlation within clusters through random effects (we'll learn about these later).

1.2 Course goals

- What are multilevel models?
- When are they useful?
- How are they fit in Stata?

Copyright ©2025 StataCorp LLC

1.3 Introduction to Stata

1.3.1 Typography

- There are various fonts that will be used when describing commands.
  - `Items that must be typed as shown` will be in a `monospaced font`

```
. or on a separate line with a . in front
```

- When commands have an `underline`, you only need to type the underlined portion.
- *Items for which a substitution is needed* will be in *italics.*
- [Optional items] will be [in square brackets], though the brackets do not get typed.

1.3.2 Downloading the course files

- To download all the materials we will use for this workshop, visit the following webpage: www.stata.com/news/conferences/spr/
- Click the link "Download here".

1.3.3 Handy navigation commands

- To print the current working directory, type `pwd`
- To change the working directory, type `cd` *folder*
- To go up one folder, type `cd ..`
- To list all files in a directory, type `ls`
- To list all Stata datasets in the working directory, type `ls *.dta`

2 Taking a step back: Linear regression

2.1 An example: Math scores

- We will start by using a dataset on math scores. This is adapted from Rabe-Hesketh and Skrondal (2022, Ex.4.3). Kreft and De Leeuw (1998) consider a subsample of students in eighth grade from the National Education Longitudinal Study of 1988 collected by the National Center for Educational Statistics of the U.S. Department of Education. The data are given in `math.dta`.

```
. use math, clear
```

- Try using the `codebook` command to see details about the dataset.

```
. codebook
--------------------------------------------------------------------------------
schid                                                                  School ID
--------------------------------------------------------------------------------

                  Type: Numeric (float)

                 Range: [6053,72991]                 Units: 1
         Unique values: 23                       Missing .: 0/519

                  Mean: 35488.7
             Std. dev.: 26056.7

           Percentiles:     10%       25%       50%       75%       90%
                           6327      7801     25642     62821     72080
--------------------------------------------------------------------------------
ses                                                         Socioeconomic status
--------------------------------------------------------------------------------

                  Type: Numeric (float)
```

```
         Range: [-2.41,1.85]                 Units: .01
 Unique values: 248                      Missing .: 0/519

          Mean: -.001272
      Std. dev.:  .880821

   Percentiles:      10%       25%       50%       75%       90%
                    -1.1      -.62      -.12       .73      1.23
--------------------------------------------------------------------------------
homework                                   Time spent on math homework each week
--------------------------------------------------------------------------------

          Type: Numeric (float)

         Range: [0,7]                        Units: 1
 Unique values: 8                        Missing .: 0/519

    Tabulation: Freq.  Value
                   42  0
                  225  1
                  111  2
                   47  3
                   47  4
                   38  5
                    6  6
                    3  7
--------------------------------------------------------------------------------
schtype                                                              School type
--------------------------------------------------------------------------------

          Type: Numeric (float)
         Label: sch

         Range: [0,1]                        Units: 1
 Unique values: 2                        Missing .: 0/519

    Tabulation: Freq.   Numeric  Label
                  202         0  Private
                  317         1  Public
--------------------------------------------------------------------------------
ratio                                                      Student-teacher ratio
--------------------------------------------------------------------------------

          Type: Numeric (float)

         Range: [10,28]                      Units: 1
 Unique values: 14                       Missing .: 0/519

          Mean: 16.7572
      Std. dev.: 4.93068

   Percentiles:     10%       25%       50%       75%       90%
                     10        13        18        20        22
--------------------------------------------------------------------------------
math                                                                  Math score
--------------------------------------------------------------------------------

          Type: Numeric (float)

         Range: [30,71]                      Units: 1
 Unique values: 42                       Missing .: 0/519

          Mean: 51.7225
      Std. dev.: 10.7092

   Percentiles:     10%       25%       50%       75%       90%
                     38        43        51        62        67
```

2.2 The linear regression model

2.2.1 Simple linear regression

- The first model we will fit is

$$\texttt{math}_{\texttt{i}} = \beta_0 + \beta_1 \texttt{schtype}_{\texttt{i}} + \varepsilon_i \tag{1}$$

  where subscript i refers to the ith student.

- We can fit this in Stata using the `regress` command. It has the following syntax:
  `regress` *depvar* [*indvars*] [, *options*]

Copyright ©2025 StataCorp LLC

  - *depvar* is the outcome variable (dependent variable)
  - *indvars* are the predictors (independent variables)

- We can fit a simple linear regression model to our data using the following command:

```
. regress math i.schtype, base
      Source |       SS           df       MS      Number of obs   =       519
-------------+----------------------------------   F(1, 517)       =     74.89
       Model |  7517.11179         1  7517.11179   Prob > F        =    0.0000
    Residual |  51890.9345       517  100.369312   R-squared       =    0.1265
-------------+----------------------------------   Adj R-squared   =    0.1248
       Total |  59408.0462       518  114.687348   Root MSE        =    10.018

------------------------------------------------------------------------------
        math | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
     schtype |
    Private  |          0  (base)
     Public  |  -7.805556   .9019425    -8.65   0.000    -9.577479   -6.033634
             |
       _cons |    56.4901   .7048956    80.14   0.000     55.10529    57.87491
------------------------------------------------------------------------------
```

- Adding a main effect of a categorical variable to a regression in Stata using `i.`*var* includes a dummy variable with 0 as the reference category corresponding to the lowest value of *var* and 1 for the target category. If a variable has k levels, it will include $k - 1$ dummy variables.
- To specify a different category as the base, it can be included in the `i.` notation. For example, `ib1.`*var* would include dummy variables that specify $var = 1$ as the base.
- In this case, `schtype` is already coded as {`0 1`} so it makes no difference in the results, but using the `i.` notation will allow for more interesting postestimation.
- When the `i.` notation is used, it is highly recommended to use the `base` option so that the output shows which level is the reference category and which level is the target category for each comparison. Instead of doing this for each model, you can set this option on permanently using

```
. set showbaselevels on, permanently
(set showbaselevels preference recorded)
```

- The regression coefficients are interpreted as
  - `_cons` (β_0): The expected (average) math score for students who attend private school
  - `Public` (β_1): The expected difference in math score of students in public school versus private school
- Alternatively, we could use the `mixed` command to fit a linear regression model. This is the command we'll be using to fit multilevel/mixed models, so let's start using it now.

```
. mixed math i.schtype
Mixed-effects ML regression                          Number of obs =    519
                                                     Wald chi2(1)  =  75.18
Log likelihood = -1931.4254                          Prob > chi2   = 0.0000

------------------------------------------------------------------------------
        math | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
     schtype |
    Private  |          0  (base)
     Public  |  -7.805556   .9002029    -8.67   0.000    -9.569922   -6.041191
             |
       _cons |    56.4901   .7035361    80.29   0.000     55.11119      57.869
------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects parameters  |   Estimate   Std. err.     [95% conf. interval]
-----------------------------+------------------------------------------------
               var(Residual) |   99.98253   6.206624      88.52869    112.9183
------------------------------------------------------------------------------
```

- The results are slightly different due to differences in estimation. In addition, we get an estimate of the residual variance, that is, $var(\varepsilon_i)$. The regression coefficients can be interpreted in the same way: The expected math score for students in private school is 56.5 [55.1, 57.9], and the expected math score for students in public schools is 7.8 [6.0, 9.6] points lower.

2.2.2 Multiple linear regression

- The next model we will fit is

$$\texttt{math}_\texttt{i} = \beta_0 + \beta_1 \texttt{schtype}_\texttt{i} + \beta_2 \texttt{homework}_\texttt{i} + \varepsilon_i \tag{2}$$

- We could fit this model in Stata using:

```
. mixed math i.schtype homework
Mixed-effects ML regression                          Number of obs =    519
                                                     Wald chi2(2)  = 168.81
Log likelihood = -1893.4542                          Prob > chi2   = 0.0000

        math   Coefficient  Std. err.      z    P>|z|     [95% conf. interval]

     schtype
     Private            0  (base)
      Public     -5.51161   .8743008    -6.30   0.000    -7.225208   -3.798012

    homework     2.600431   .2875559     9.04   0.000     2.036832     3.16403
       _cons     49.96327   .9739051    51.30   0.000     48.05446    51.87209

  Random-effects parameters    Estimate   Std. err.     [95% conf. interval]

               var(Residual)    86.37263   5.361761      76.47791    97.54752
```

- The regression coefficients are interpreted as
  - `_cons` (β_0): The expected (average) math score for students who attend private school
  - `Public` (β_1): The expected difference in math score of students in public school versus private school
  - `homework` (β_2): The expected change in math score with each additional hour spent on homework each week (assumed to be the same effect for students in private and public school)
- The random effects are
  - `var(Residual)`: The variance of the residual, ε_i

2.2.3 Full factorial multiple linear regression

- The next model we will fit is

$$\texttt{math}_\texttt{i} = \beta_0 + \beta_1 \texttt{schtype}_\texttt{i} + \beta_2 \texttt{homework}_\texttt{i} + \beta_3 \texttt{schtype}_\texttt{i}\texttt{homework}_\texttt{i} + \varepsilon_i \tag{3}$$

- We could fit this model in Stata using:

```
. mixed math schtype##c.homework
Mixed-effects ML regression                          Number of obs =    519
                                                     Wald chi2(3)  = 169.47
Log likelihood = -1893.2039                          Prob > chi2   = 0.0000

               math   Coefficient  Std. err.      z    P>|z|     [95% conf. interval]

            schtype
            Private            0  (base)
             Public     -4.68586   1.457863    -3.21   0.001    -7.543218   -1.828502

           homework     2.780179   .3835765     7.25   0.000     2.028383    3.531975
```

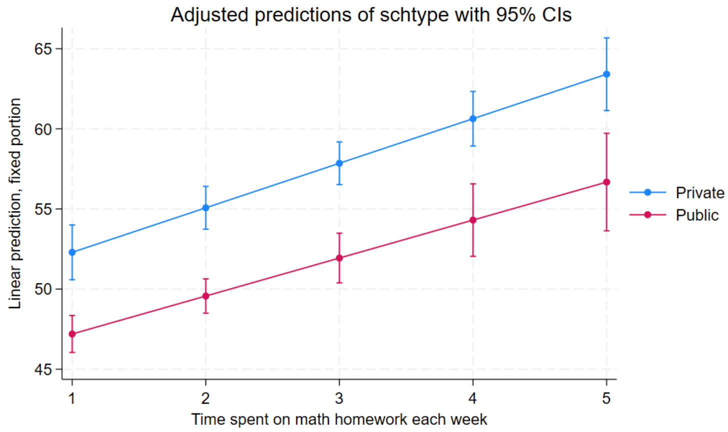
```
schtype#c.homework |
            Public |  -.4098806   .5792275    -0.71   0.479    -1.545146    .7253845
                   |
             _cons |   49.51213   1.163633    42.55   0.000     47.23145     51.7928
------------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects parameters  |   Estimate   Std. err.     [95% conf. interval]
-----------------------------+------------------------------------------------
               var(Residual) |   86.28937   5.356593       76.4042     97.4535
------------------------------------------------------------------------------
```

- Adding an interaction effect to a regression in Stata automatically assumes each variable is a factor variable and will include dummy variables for them as explained previously. Each continuous variable in an interaction term needs to be preceded by a `c.`, i.e., `c.`*var*. Using `c.var#var2` will just include the interaction term while `c.var##var2` will include the interaction and both main-effects terms.
- The regression coefficients are interpreted as
  - `_cons` (β_0): The expected (average) math score for students who attend private school and don't do any homework
  - `Public` (β_1): The expected difference in math score of students who don't do any homework in public school versus private school
  - `homework` (β_2): The expected change in math score for students in private school with each additional hour spent on homework each week
  - `schtype#c.homework` (β_3): The expected difference in the effect of homework on math score of students in public schools versus private school
- The random effects are
  - `var(Residual)`: The variance of the residual, ε_i
- To understand the interaction effect, we can plot marginal means.

```
. margins schtype, at(homework=(1/5))

Adjusted predictions                                    Number of obs = 519

Expression: Linear prediction, fixed portion, predict()
1._at: homework = 1
2._at: homework = 2
3._at: homework = 3
4._at: homework = 4
5._at: homework = 5

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
 _at#schtype |
   1#Private |    52.2923   .8732722    59.88   0.000     50.58072    54.00389
    1#Public |   47.19656   .5885924    80.19   0.000     46.04294    48.35018
   2#Private |   55.07248   .6822236    80.72   0.000     53.73535    56.40962
    2#Public |   49.56686   .5461755    90.75   0.000     48.49638    50.63735
   3#Private |   57.85266   .6800849    85.07   0.000     56.51972     59.1856
    3#Public |   51.93716   .7917838    65.60   0.000     50.38529    53.48903
   4#Private |   60.63284   .8682534    69.83   0.000      58.9311    62.33459
    4#Public |   54.30746   1.154246    47.05   0.000     52.04518    56.56974
   5#Private |   63.41302   1.157357    54.79   0.000     61.14464     65.6814
    5#Public |   56.67776   1.553831    36.48   0.000      53.6323    59.72321
------------------------------------------------------------------------------
```

- This table reports the average predicted math scores of students who spend 1–5 hours on math homework each week, in public and private schools. We can visualize these results using `marginsplot`.

```
. marginsplot

Variables that uniquely identify margins: homework schtype
```



2.2.4 Predictions

- The postestimation command `predict` can be used after any estimation command to get predictions based on the fitted model. It uses the following syntax:

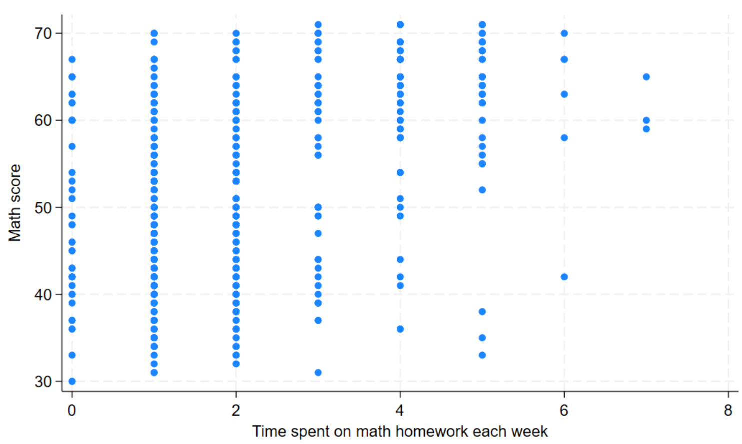
  `predict` *newvar* [, *statistic*]

- After `mixed`, there are several *statistic* options. For now we'll just consider fitted values, the default statistic.

```
. predict yhat1
(option xb assumed)
```

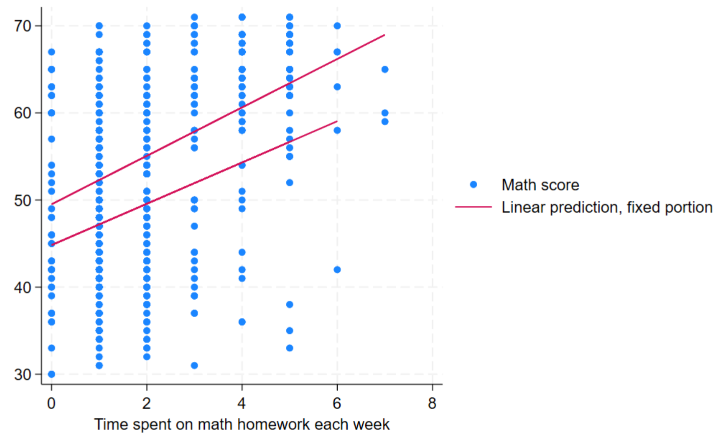
- We have a new variable we named `yhat1`, which contains the predicted math score for each student based on the time they spent on math homework and the type of school they attend. We can visualize these predictions using the `twoway` command. Let's start with a scatterplot of the observed variables.

```
. twoway (scatter math homework)
```



- Now, let's add the predicted math scores for each observation. We'll do this by adding a line plot with the `connect(ascending)` option. This will continue drawing the same line until the X-variable decreases, so we first need to sort by homework within school.

```
. sort schid homework
. twoway (scatter math homework) (line yhat1 homework, connect(ascending))
```



- We see blue dots representing the observed math score for each student. We also see two red lines: one for the predicted math scores for students in public schools and one for the predicted math scores for students in private schools. These are the same lines we saw with `marginsplot`.
- Notice that school type and hours spent on homework are the only variables that determine predicted math score, as our model implies.

2.3 Assumptions

2.3.1 Assumptions of the linear regression model

- There are four assumptions we need to consider when fitting a linear model.

1. Linearity: $E(y_i) = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip}$ or $E(e_i) = 0$
   (a) Also implies y_i depends only on the x_i's
   (b) Note that this assumption applies to the regression coefficients, not the predictors
2. Homoskedasticity: $var(\varepsilon_i) = \sigma^2$
3. Independence: $cov(\varepsilon_i, \varepsilon_j) = 0$
4. Normality: $\varepsilon_i \sim N(0, \sigma^2)$
   - Not required by ordinary least squares (OLS) to estimate coefficients, but required by maximum likelihood (ML) to test hypotheses on the estimated parameters

- Linearity, homoskedasticity and normality can be evaluated using diagnostic plots.
- Independence cannot be evaluated through plots. Rather, the study design should be carefully thought over. Could we consider the observations in our dataset independent of each other? Probably not. It is natural to think that students who attend the same school are likely to have more similar math scores than students from different schools.

2.3.2 The nested data problem

- Let's take a look at observations from two of the schools in our dataset.

```
. list schid math schtype homework if (schid==6327 | schid==72991), sepby(schid) noobs
```

schid	math	schtype	homework
6327	65	Private	0
6327	67	Private	0
6327	60	Private	0
6327	65	Private	0
6327	46	Private	1

6327	56	Private	1
6327	61	Private	1
6327	39	Private	3
72991	30	Public	0
72991	40	Public	0
72991	63	Public	1
72991	56	Public	1
72991	43	Public	1
72991	41	Public	1
72991	53	Public	1
72991	55	Public	1
72991	65	Public	2
72991	53	Public	2
72991	69	Public	4
72991	67	Public	4
72991	71	Public	5
72991	63	Public	5

- We can see that we have observed multiple students from the same school. This indicates that our data has nesting; specifically, students are nested within school.
- Nesting is when observations share a natural common environment or grouping. This creates statistical dependencies among observations of the same cluster.
- Linear regression (OLS) assumes the error terms are independent. In other words, observations are independent. When this assumption has not been met:
  - The parameter estimates are generally unbiased.
  - The standard error estimates are biased. This can result in
    - Inflated test statistics
    - Inflated type I error

3 The multilevel/mixed model

- Multilevel models include both a fixed-effects portion (regression coefficients) and a random-effects portion (variability of the regression coefficients across clusters).

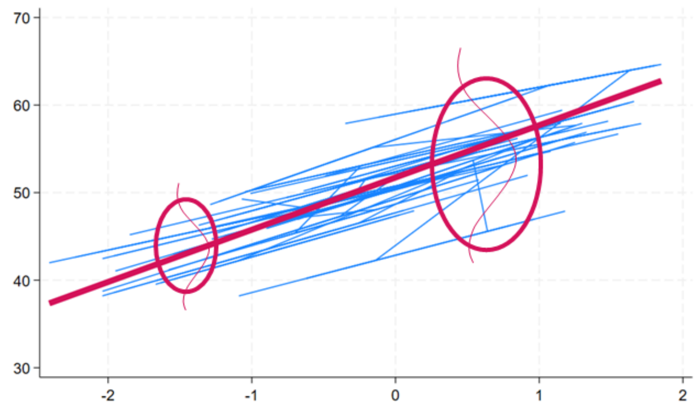
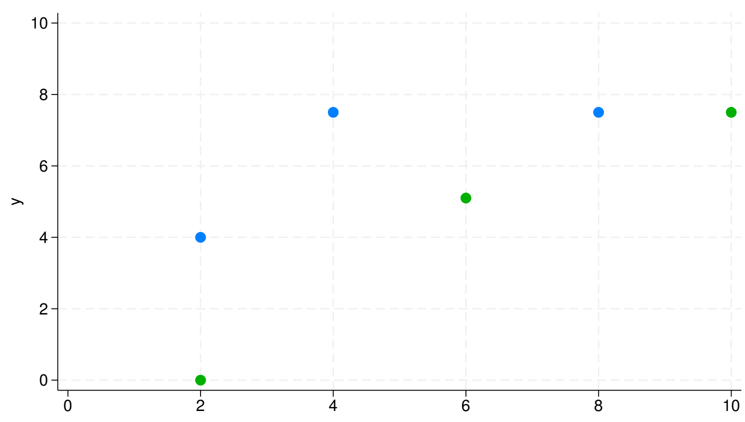


Figure 2: MLMMs estimate average coefficients and their variability across clusters.

- Adding random effects to the model allows
  - Dependencies/heteroskedasticity among the observations
  - More generalizable inferences, because we assume the observed clusters are a random sample of the population of clusters.
  - Investigation of more complex research questions
    - Estimating heterogeneity in effects across clusters
    - Evaluating the effect of a predictor both within and between clusters
    - Studying interindividual differences in intraindividual change
- Let's get to know this model in the next few sections.

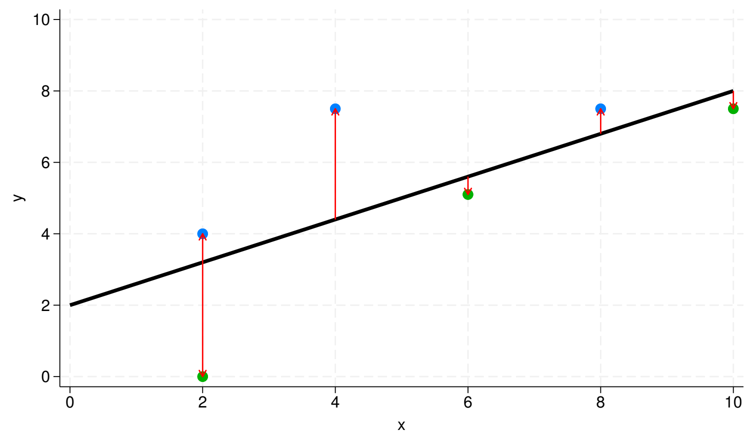
3.1 Understanding the MLMM

- Let's start by spending some time understanding the MLMM model conceptually by considering the following scatterplot, where different colored dots represent individuals from different clusters.

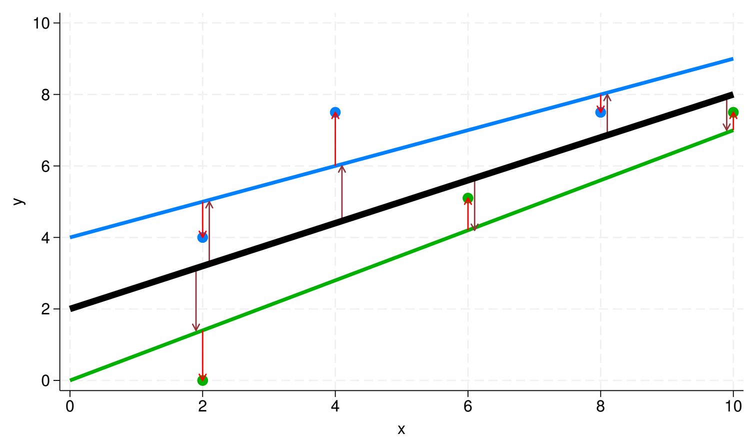


- In linear regression, the clustering is ignored and we fit one line to the entire dataset (see Figure 3a). The residuals (red arrows) are the deviations from each observation to their predicted outcome.

- In multilevel models, it is like we are fitting a separate regression model within each cluster (see Figure 3b). We can make inferences both about the averages of these lines, called fixed effects, and the variances of these lines, called random effects.
  - The fixed-effects coefficients (black line) are the average of the cluster-specific regression coefficients.
  - The level-2 random effects (maroon arrows), are the deviations from the cluster-specific regression line to the average regression line.
  - The level-1 residuals (red arrows) are deviations from each observation to their cluster's predicted outcome.



(a) Linear regression



(b) Multilevel model

Figure 3: Multilevel models disaggregate random variation into two components

- Essentially, what an MLMM is doing is splitting up the OLS residuals between observation-level residuals and cluster-level random effects. Hence, variance-component model!
- An important caveat of this demonstration is that when fitting a MLMM, the cluster-specific intercepts and slopes are not actually being estimated (they are in Bayesian estimation). They can be predicted from the model after it has been fit, but the model only provides estimates of the variances/covariances of the random effects, not the random effects themselves.

3.2 Fitting MLMMs with `mixed`

- MLMMs can be fit in Stata using the `mixed` command by adding a random-effects clause to our command after a double bar (||). Here is the general syntax:

  `mixed` *depvar fe_equation* [, *fe_options*] [|| *clustvar* : *re_equation*] [, *re_options*]

- *depvar* is the outcome variable.
- The variables in *fe_equation* are the variables in the fixed-effects portion of the model. The model will estimate regression coefficients for these variables.
- *clustvar* is the cluster identification variable. Adding *clustvar*: adds a random intercept to the model.
- The variables in *re_equation* are the variables in the random-effects portion of the model. The model will estimate variances for these variables' coefficients.
- The level-1 residual is always estimated.
- You can specify additional options in either the fixed-effects or the random-effects portion of the model.

3.2.1 Random-intercept model

- We start by fitting a random-intercept model to our data. Now that we are fitting more complicated models, we will start using the `nolog` option to save space in the output.

```
. mixed math schtype##c.homework || schid: , nolog
Mixed-effects ML regression                         Number of obs    =    519
Group variable: schid                               Number of groups =     23
                                                    Obs per group:
                                                                 min =      5
                                                                 avg =   22.6
                                                                 max =     67
                                                    Wald chi2(3)     =  82.26
Log likelihood = -1862.6296                         Prob > chi2      = 0.0000

------------------------------------------------------------------------------
            math | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-----------------+------------------------------------------------------------
         schtype |
          Public |  -5.200103    2.23247   -2.33   0.020    -9.575664   -.8245421
        homework |   2.200494   .3753395    5.86   0.000     1.464842    2.936146
                 |
schtype#c.homework |
          Public |   .3507493   .5556287    0.63   0.528     -.738263    1.439762
                 |
           _cons |   49.77747   1.800483   27.65   0.000     46.24859    53.30635
------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects parameters  |   Estimate   Std. err.     [95% conf. interval]
-----------------------------+------------------------------------------------
schid: Identity              |
                  var(_cons) |   15.94038   5.793795      7.818351    32.49989
-----------------------------+------------------------------------------------
               var(Residual) |   71.04279   4.510533      62.73023    80.45686
------------------------------------------------------------------------------
LR test vs. linear model: chibar2(01) = 61.15         Prob >= chibar2 = 0.0000
```

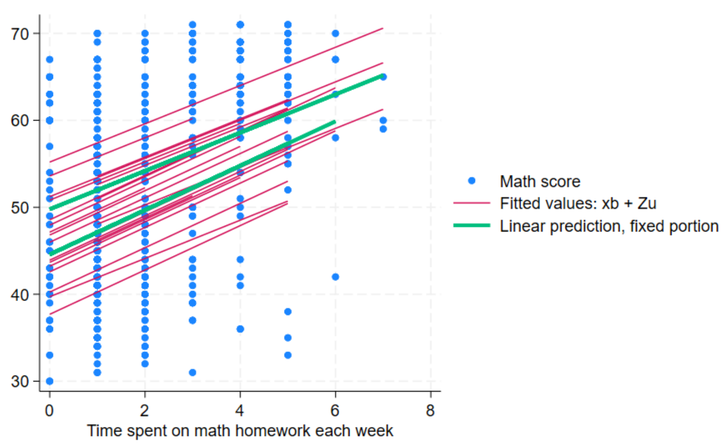
- The coefficients in the fixed-effects portion of the model are still
  - `_cons`: The expected (average) math score for students who attend private school and don't do any homework
  - `Public`: The expected difference in math score of students who don't do any homework in public school versus private school
  - `homework`: The expected change in math score for students in private school with each additional hour spent on homework each week
  - `schtype#c.homework`: The expected difference in the effect of homework on math score of students in public schools versus private school
- The random effects are
  - `var(_cons)`: The variance of the school-specific intercepts around the overall intercept conditioning on school type.
  - `var(Residual)`: The variance of the level-1 residual

- The random intercept is telling us about how different schools' average math scores are from each other, while the residual tell us about how different students' math scores are from each other within the same school.
- Let's see how the addition of the random intercept influences our predictions.
- Now that we have added a random effect to our model, we have two options to predict fitted values:
  - `xb` (default): Linear prediction of the outcome based on the fixed-effects portion of the model only
  - `fitted`: Linear prediction of the outcome based on the fixed-effects and random-effects (school-level) portions of the model
- Let's predict both and plot them.

```
. predict yhat2_fe
(option xb assumed)
. predict yhat2, fitted
```

- We'll use the same `twoway` syntax we used previously to overlay our predictions on top of a scatterplot of the observed math scores.

```
. twoway (scatter math homework) (line yhat2 homework, connect(ascending)) ///
>         (line yhat2_fe homework, connect(ascending) lwidth(thick))
```



- The predictions based on both fixed and random effects are in red, showing us the predicted math scores for each school. The predictions based on the fixed effects are in green, mimicking the linear regression model.
- Our model is assuming there is a normal distribution of intercepts around each of the green lines. This means that, even after accounting for school type, there are still school-level differences in expected math score, as captured by the random intercept. These are random effects because we're not estimating an intercept for each school, rather, we're estimating a distribution of schools.
- The only slope differences in our predictions right now are due to the school type by homework interaction in our model. In other words, we've estimated a difference in slope by school type. In the next section, we'll allow a random distribution of slopes by adding a random effect to the homework coefficient.

3.2.2 Random-slope model

- It's probable that the effect of homework on math score is not the same in every school. Perhaps some schools tend to give very good homework, while others tend to give busy work. In this case, the effect of time spent doing this work will impact math scores differently.
- If we wanted to let the effect of homework vary by school, we could additionally add a random effect for homework. We can fit random-coefficient models in `mixed` by adding those variables after *clustvar*`:` in the model statement.

```
. mixed math schtype##c.homework || schid: homework , nolog
Mixed-effects ML regression                     Number of obs     =        519
Group variable: schid                           Number of groups  =         23
                                                Obs per group:
                                                              min =          5
                                                              avg =       22.6
                                                              max =         67
                                                Wald chi2(3)      =       7.62
Log likelihood = -1830.0249                     Prob > chi2       =     0.0546
```

math	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
schtype						
Public	-3.583506	3.197013	-1.12	0.262	-9.849536	2.682525
homework	2.230744	1.343726	1.66	0.097	-.4029109	4.8644
schtype#c.homework						
Public	-.3770907	1.667344	-0.23	0.821	-3.645025	2.890844
_cons	48.83767	2.597976	18.80	0.000	43.74573	53.92961

Random-effects parameters	Estimate	Std. err.	[95% conf.	interval]
schid: Independent				
var(homework)	12.42178	4.591824	6.019067	25.63529
var(_cons)	43.93973	15.68365	21.82897	88.44667
var(Residual)	54.11022	3.579197	47.53082	61.60035

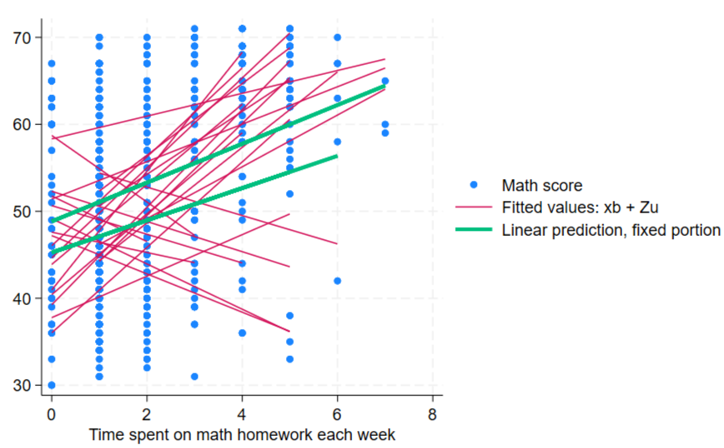
```
LR test vs. linear model: chi2(2) = 126.36            Prob > chi2 = 0.0000
Note: LR test is conservative and provided only for reference.
```

- Now, we have a new school-level random-effects parameter.
  - `var(homework)` : The variance of the school-specific coefficients for homework around the overall coefficient conditioning on school type
- `var(_cons)` still captures the school-level variability in average math scores and `var(Residual)` captures the student-level variability.
- Let's see how the additional random effect influences our predictions.

```
. predict yhat3_fe
(option xb assumed)
. predict yhat3, fitted
```

- We can use the same `twoway` plot we created earlier with these updated predictions.

```
. twoway (scatter math homework) (line yhat3 homework, connect(ascending)) ///
>         (line yhat3_fe homework, connect(ascending) lwidth(thick))
```



- Notice that our fixed-effects (green lines) haven't changed, but now our school-level prediction lines have all different slopes. Again, we aren't estimating a separate slope for each school. Rather, we are estimating a distribution in slopes around their fixed effects.
- Usually, when we have two or more cluster-level random effects we would like to estimate their covariance rather than constraining their effects to be independent. We can do this in Stata by adding the `covariance(unstructured)` option to specify an unstructured covariance matrix on the random effects.

```
. mixed math schtype##c.homework || schid: homework , covariance(unstructured) nolog
Mixed-effects ML regression                          Number of obs    =    519
Group variable: schid                                Number of groups =     23
                                                     Obs per group:
                                                                  min =      5
                                                                  avg =   22.6
                                                                  max =     67
                                                     Wald chi2(3)     =   9.72
Log likelihood = -1817.3855                          Prob > chi2      = 0.0211

------------------------------------------------------------------------------------
              math | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------------+----------------------------------------------------------------
           schtype |
            Public |  -3.289684   3.545698    -0.93   0.354    -10.23913    3.659758
          homework |   2.306169   1.511835     1.53   0.127    -.6569722    5.269311
                   |
schtype#c.homework |
            Public |    -.49695   1.874733    -0.27   0.791     -4.17136     3.17746
                   |
             _cons |   48.54675   2.879905    16.86   0.000     42.90224    54.19126
------------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects parameters  |   Estimate   Std. err.     [95% conf. interval]
-----------------------------+------------------------------------------------
schid: Unstructured          |
               var(homework) |   16.30338   5.675551      8.240556    32.25514
                  var(_cons) |    56.2109   19.20075      28.77819    109.7937
         cov(homework,_cons) |  -25.92574    9.61518     -44.77114   -7.080332
-----------------------------+------------------------------------------------
               var(Residual) |   53.34396   3.471673       46.9557    60.60134
------------------------------------------------------------------------------
LR test vs. linear model: chi2(3) = 151.64               Prob > chi2 = 0.0000
Note: LR test is conservative and provided only for reference.
```

- It can be easier to understand random effects in terms of standard deviations and correlations rather than variances and covariances. We can see these after fitting a `mixed` model using postestimation command `estat sd`, or by adding the `stddev` option to `mixed`.

```
. estat sd

------------------------------------------------------------------------------
  Random-effects parameters  |   Estimate   Std. err.     [95% conf. interval]
-----------------------------+------------------------------------------------
schid: Unstructured          |
                sd(homework) |   4.037745   .7028121      2.870637    5.679361
                   sd(_cons) |   7.497393   1.280495      5.364531    10.47825
        corr(homework,_cons) |  -.8564106   .0640335     -.9414375   -.6689814
-----------------------------+------------------------------------------------
                sd(Residual) |   7.303695   .2376655      6.852423    7.784686
------------------------------------------------------------------------------
```

- The correlation is highly negative, meaning that schools with higher average math scores tend to have lower than average homework slopes.
- One way to interpret a standard deviation is the "typical" distance from the mean. Here, we can see that while the average effect of homework is 2.31, a particular school's effect is typicallly 4.03 away from that.

3.3 Model formulation

- Let's take a look at our final model formulation.

$$\begin{aligned}\texttt{math}_{\texttt{ij}} &= \beta_{0j} + \beta_{1j}\texttt{homework}_{\texttt{ij}} + e_{ij} \\ \beta_{0j} &= \gamma_{00} + \gamma_{01}\texttt{schtype}_{\texttt{j}} + u_{0j} \\ \beta_{1j} &= \gamma_{10} + \gamma_{11}\texttt{schtype}_{\texttt{j}} + u_{1j}\end{aligned} \tag{4}$$

or, equivalently,

$$\begin{aligned}\texttt{math}_{\texttt{ij}} = {} & \gamma_{00} + \gamma_{10}\texttt{homework}_{\texttt{ij}} + \gamma_{01}\texttt{schtype}_{\texttt{j}} + \gamma_{11}\texttt{homework}_{\texttt{ij}}\texttt{schtype}_{\texttt{j}} \\ & + u_{0j} + u_{1j}\texttt{homework}_{\texttt{ij}} + e_{ij}\end{aligned} \tag{5}$$

- where $e_{ij} \sim N(0, \sigma_e^2)$ and $\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim MVN\left(\mathbf{0}, \boldsymbol{\tau}\right)$.
- The subscripts refer to the ith student of the jth school. Accordingly, the outcome, $\texttt{math}_{\texttt{ij}}$, varies at the student level, $\texttt{homework}_{\texttt{ij}}$ is a student-level predictor, and $\texttt{schtype}_{\texttt{j}}$ is a school-level predictor.

- The first line (level 1) of Eq. 4 is the observation-level model. The second and third lines (level 2) are the cluster-level model. Hence the name, multilevel model!
- The first line of Eq. 5 contains the fixed-effects portion of the model: $\gamma_{00}, \gamma_{01}, \gamma_{10}, \gamma_{11}$. The second line contains the random-effects portion of the model: e_{ij}, u_{0j}, u_{1j}. Hence the name, mixed-effects model!
- This is a conditional random-coefficient model. Here are some definitions:
  - Random-intercept/Random-coefficient model:
    - Models without any random effects, u_j terms, are OLS linear regression models.
    - Models with only one u_{0j} for the intercept are random-intercept models.
    - Models with more than one level-2 random effect are random-coefficient or random-slope models.
  - Conditional/Unconditional model:
    - This refers to whether the random effects are conditional on a cluster-level predictor.
    - Conditional models have at least one level-2 predictor. In our example, we have `schtype`.
    - A multilevel model without any level-2 predictors is called an unconditional model.
- For a more detailed introduction to MLMM, see Bryk and Raudenbush (1992).
- The `mixed` command is written out similarly to the condensed version of the model, with all predictors (observation and cluster-level) listed on the left side of the || and the predictors with random effects listed on the right.
- The `covariance()` option refers to the structure of the $\boldsymbol{\tau}$ matrix. Here are all the `cov()` suboptions:
  - `independent`: (default) Each random effect has its own variance; all are independent. This structure is most commonly used when the random effect is an intervention.

$$\begin{bmatrix} u_{0i} \\ u_{1i} \end{bmatrix} \sim MVN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \tau_0 & 0 \\ 0 & \tau_1 \end{bmatrix}\right)$$

  - `unstructured`: All variances and covariances uniquely estimated. Typical for random-slope modes, and what we use in this course.

$$\begin{bmatrix} u_{0i} \\ u_{1i} \end{bmatrix} \sim MVN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \tau_{00} & \tau_{01} \\ \tau_{01} & \tau_{11} \end{bmatrix}\right)$$

  - `exchangeable`: All random effects share a single variance and covariance.

$$\begin{bmatrix} u_{0i} \\ u_{1i} \end{bmatrix} \sim MVN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \tau & \rho\tau \\ \rho\tau & \tau \end{bmatrix}\right)$$

– `identity`: All random effects share a single variance; all are independent.

$$\begin{bmatrix} u_{0i} \\ u_{1i} \end{bmatrix} \sim MVN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \tau & 0 \\ 0 & \tau \end{bmatrix}\right)$$

- Putting it all together, this gives us our `mixed` command we've seen before.

```
. mixed math schtype##c.homework || schid: homework , covariance(unstructured) nolog

Mixed-effects ML regression                         Number of obs    =    519
Group variable: schid                               Number of groups =     23
                                                    Obs per group:
                                                                 min =      5
                                                                 avg =   22.6
                                                                 max =     67
                                                    Wald chi2(3)     =   9.72
Log likelihood = -1817.3855                         Prob > chi2      = 0.0211

------------------------------------------------------------------------------------
                 math | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
----------------------+-------------------------------------------------------------
              schtype |
               Public |  -3.289684   3.545698    -0.93   0.354    -10.23913    3.659758
             homework |   2.306169   1.511835     1.53   0.127    -.6569722    5.269311
                      |
   schtype#c.homework |
               Public |    -.49695   1.874733    -0.27   0.791     -4.17136     3.17746
                      |
                _cons |   48.54675   2.879905    16.86   0.000     42.90224    54.19126
------------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects parameters  |   Estimate   Std. err.     [95% conf. interval]
-----------------------------+------------------------------------------------
schid: Unstructured          |
               var(homework) |   16.30338   5.675551      8.240556    32.25514
                  var(_cons) |    56.2109   19.20075      28.77819    109.7937
         cov(homework,_cons) |  -25.92574    9.61518     -44.77114   -7.080332
-----------------------------+------------------------------------------------
               var(Residual) |   53.34396   3.471673       46.9557    60.60134
------------------------------------------------------------------------------
LR test vs. linear model: chi2(3) = 151.64                Prob > chi2 = 0.0000
Note: LR test is conservative and provided only for reference.
```

- Can you match each estimate to the parameters in Equation 4 and 5?
- The coefficients in the fixed-effects portion of the model are
  – `_cons` (γ_{00}): The expected (average) math score for students who attend private school and don't do any homework
  – `homework` (γ_{10}): The expected change in math score for students in private school with each additional hour spent on homework each week
  – `Public` (γ_{01}): The expected difference in math score of students who don't do any homework in public school versus private school
  – `schtype#c.homework` (γ_{11}): The expected difference in the effect of homework on math score of students in public schools versus private school
- The random effects are
  – `var(_cons)` (τ_{00}): The variance of the school-specific intercepts around γ_{00} conditioning on school type
  – `var(homework)` (τ_{11}): The variance of the school-specific coefficients for homework around γ_{10} conditioning on school type
  – `var(homework,_cons)` (τ_{01}): The covariance between the random intercept and the random coefficient for homework
  – `var(Residual)` (σ_e^2): The variance of the level-1 residual

3.4 Intraclass correlation (ICC)

- The ICC is the proportion of the total variance that occurs at the cluster-level:

$$\rho = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\varepsilon^2} \tag{6}$$

where σ_u^2 is the cluster-level variance around the random intercept, `var(_cons)`, and σ_ε^2 is the level-1 residual variance, `var(Residual)`.

- It is called the intraclass correlation because it is also the correlation between level-1 units within the same level-2 unit (i.e., between students in the same school).
- The ICC can be computed after any `mixed` model using `estat icc`, but it is often most useful to calculate the ICC on an intercept-only random-intercept model. Otherwise, it is reporting the decomposition of variance conditional on the predictors in the model. When the model includes random coefficients, it is calculated by setting the random coefficients to zero.

```
. mixed math || schid: , nolog

Mixed-effects ML regression                         Number of obs    =  519
Group variable: schid                               Number of groups =   23
                                                    Obs per group:
                                                                 min =    5
                                                                 avg = 22.6
                                                                 max =   67
                                                    Wald chi2(0)     =    .
Log likelihood = -1900.3882                         Prob > chi2      =    .

------------------------------------------------------------------------------
        math | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
       _cons |   50.75742   1.126718    45.05   0.000     48.54909    52.96575
------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects parameters  |   Estimate   Std. err.     [95% conf. interval]
-----------------------------+------------------------------------------------
schid: Identity              |
                  var(_cons) |    24.8507    8.40812      12.80379    48.23238
-----------------------------+------------------------------------------------
               var(Residual) |   81.23738   5.152984       71.7403     91.9917
------------------------------------------------------------------------------
LR test vs. linear model: chibar2(01) = 132.29        Prob >= chibar2 = 0.0000

. estat icc

Intraclass correlation

------------------------------------------------------------------------------
                       Level |        ICC   Std. err.     [95% conf. interval]
-----------------------------+------------------------------------------------
                       schid |   .2342459   .0619988      .1344784    .3758852
------------------------------------------------------------------------------
```

- About 23% of the variance in math score varies between schools, and the remaining 77% varies within schools between students.
- Additionally, the average correlation in math scores between students at the same school is 0.23 [0.13, 0.38].
- In model building, it is important to start here to see where the variance lies. Adding level-1 predictors will account for both level-1 and level-2 variance:

```
. mixed math homework || schid: , nolog

Mixed-effects ML regression                       Number of obs    =     519
Group variable: schid                             Number of groups =      23
                                                  Obs per group:
                                                               min =       5
                                                               avg =    22.6
                                                               max =      67
                                                  Wald chi2(1)     =   75.32
Log likelihood =  -1865.247                       Prob > chi2      =  0.0000
```

```
------------------------------------------------------------------------------
        math | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
    homework |   2.401972   .2767745     8.68   0.000     1.859504     2.94444
       _cons |   46.34945   1.141154    40.62   0.000     44.11283    48.58607
------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects parameters  |   Estimate   Std. err.     [95% conf. interval]
-----------------------------+------------------------------------------------
schid: Identity              |
                  var(_cons) |   20.22587   7.071243      10.19331    40.13276
-----------------------------+------------------------------------------------
               var(Residual) |   71.14391   4.517152      62.81918    80.57183
------------------------------------------------------------------------------
LR test vs. linear model: chibar2(01) = 94.71         Prob >= chibar2 = 0.0000
```

- When we recalculate ICC, it will be conditional on this predictor.

```
. estat icc
Residual intraclass correlation

------------------------------------------------------------------------------
                       Level |        ICC   Std. err.     [95% conf. interval]
-----------------------------+------------------------------------------------
                       schid |   .2213628   .0616008      .1236598    .3641807
------------------------------------------------------------------------------
```

- After accounting for time spent on homework, 22% of the remaining variability in math score is at the school level.
- Adding level-2 predictors will only account for level-2 variance:

```
. mixed math homework i.schtype || schid: , nolog
Mixed-effects ML regression                     Number of obs     =        519
Group variable: schid                           Number of groups  =         23
                                                Obs per group:
                                                              min =          5
                                                              avg =       22.6
                                                              max =         67
                                                Wald chi2(2)      =      81.88
Log likelihood = -1862.8283                     Prob > chi2       =     0.0000

------------------------------------------------------------------------------
        math | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
    homework |   2.360639    .276887     8.53   0.000      1.81795    2.903328
             |
     schtype |
     Public  |  -4.531238   1.956746    -2.32   0.021    -8.366389   -.6960857
       _cons |   49.42096   1.703702    29.01   0.000     46.08177    52.76016
------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects parameters  |   Estimate   Std. err.     [95% conf. interval]
-----------------------------+------------------------------------------------
schid: Identity              |
                  var(_cons) |   15.76732   5.747039      7.717968    32.21163
-----------------------------+------------------------------------------------
               var(Residual) |   71.12805   4.516064      62.80532    80.55369
------------------------------------------------------------------------------
LR test vs. linear model: chibar2(01) = 61.25         Prob >= chibar2 = 0.0000
. estat icc
Residual intraclass correlation

------------------------------------------------------------------------------
                       Level |        ICC   Std. err.     [95% conf. interval]
-----------------------------+------------------------------------------------
                       schid |   .1814518   .0553255        .09651     .315082
------------------------------------------------------------------------------
```

- After accounting for time spent on homework and school type, 18% of the remaining variability in math score is at the school level.

Copyright ©2025 StataCorp LLC

3.5 The disaggregated model

- In some applications, it is expected that a predictor may have a different effect on the outcome within clusters than it does between clusters.
- For example, the effect of anxiety on performance often has a different impact between individuals than it does within an individual.
  - Within: When an individual is more anxious about a test than usual, their performance is worse than usual.
  - Between: Individuals with higher average levels of anxiety perform better than individuals with lower average anxiety.

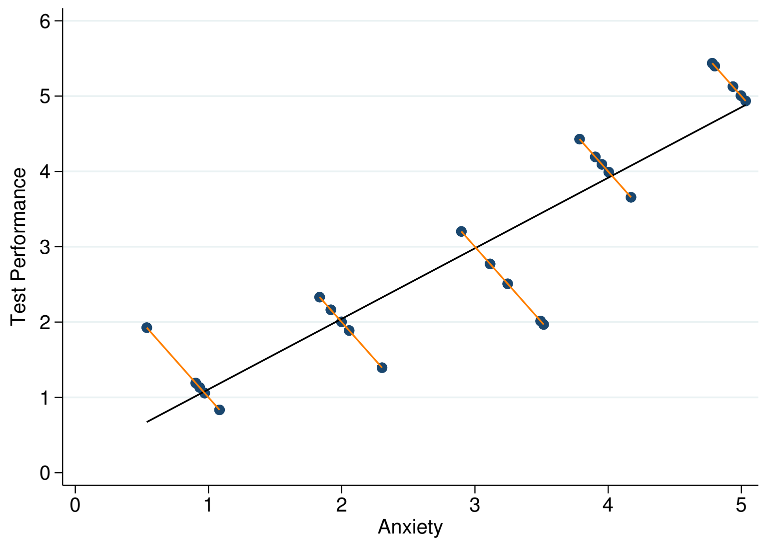


Figure 4: Between (black line) vs. within (red line) participant effects

- When it's expected that the between effect may differ from the within effect, that variable can be group-mean centered on level 1 and its group mean can be added to level 2.
- Model with conflated within-between effect

$$
\begin{aligned}
y_{ij} &= \beta_{0j} + \beta_{1j}x_{ij} + e_{ij} \\
\beta_{0j} &= \gamma_{00} + u_{0j} \\
\beta_{1j} &= \gamma_{10}
\end{aligned} \tag{7}
$$

or,

$$y_{ij} = \gamma_{00} + \gamma_{10}x_{ij} + u_{0j} + e_{ij}$$

  - γ_{10}: composite effect (blend of within and between effect)

- Model with disaggregated within-between effect

$$
\begin{aligned}
y_{ij} &= \beta_{0j} + \beta_{1j}(x_{ij} - \bar{x_j}) + e_{ij} \\
\beta_{0j} &= \gamma_{00} + \gamma_{01}\bar{x_j} + u_{0j} \\
\beta_{1j} &= \gamma_{10}
\end{aligned} \tag{8}
$$

Copyright ©2025 StataCorp LLC

or,

$$y_{ij} = \gamma_{00} + \gamma_{01}\bar{x_j} + \gamma_{10}(x_{ij} - \bar{x_j}) + u_{0j} + e_{ij}$$

where $\bar{x_j}$ is the mean of x in cluster j.

- γ_{01}: between-cluster effect
- γ_{10}: within-cluster effect

- Returning to the math dataset, let's compare the within-school impact of homework on math achievement with the between-school impact. Let's start by fitting the conflated model.

```
. mixed  math homework || schid: , nolog
Mixed-effects ML regression                          Number of obs    =    519
Group variable: schid                                Number of groups =     23
                                                     Obs per group:
                                                                  min =      5
                                                                  avg =   22.6
                                                                  max =     67
                                                     Wald chi2(1)     =  75.32
Log likelihood =  -1865.247                          Prob > chi2      = 0.0000

------------------------------------------------------------------------------
        math | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
    homework |   2.401972   .2767745     8.68   0.000     1.859504     2.94444
       _cons |   46.34945   1.141154    40.62   0.000     44.11283    48.58607
------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects parameters  |   Estimate   Std. err.     [95% conf. interval]
-----------------------------+------------------------------------------------
schid: Identity              |
                  var(_cons) |   20.22587   7.071243      10.19331    40.13276
-----------------------------+------------------------------------------------
               var(Residual) |   71.14391   4.517152      62.81918    80.57183
------------------------------------------------------------------------------
LR test vs. linear model: chibar2(01) = 94.71         Prob >= chibar2 = 0.0000
```

- We can store these results using `estimates store`.

```
. estimates store conflated
```

- To fit the disaggregated model, we need to create two new predictors: one for the between effect and one for the within effect.
  - To get the between effect, we generate a group mean for `homework`: the average number of hours spent on homework within each school.

    ```
    . bysort schid: egen HWb = mean(homework)
    ```

  - To get the within effect, we generate a variable of deviations of each student's homework hours from their school's average homework hours.

    ```
    . generate HWw = homework - HWb
    ```

- We fit the disaggregated model with both of these variables.

```
. mixed  math HWb HWw || schid: , nolog
Mixed-effects ML regression                          Number of obs    =    519
Group variable: schid                                Number of groups =     23
                                                     Obs per group:
                                                                  min =      5
                                                                  avg =   22.6
                                                                  max =     67
                                                     Wald chi2(2)     =  76.48
Log likelihood = -1864.8012                          Prob > chi2      = 0.0000
```

```
        math   Coefficient  Std. err.      z    P>|z|     [95% conf. interval]

         HWb      4.023327    1.70176    2.36   0.018     .6879397    7.358715
         HWw      2.360971   .2804099    8.42   0.000     1.811378    2.910564
       _cons      43.37124   3.276835   13.24   0.000     36.94876    49.79372

  Random-effects parameters    Estimate   Std. err.     [95% conf. interval]

schid: Identity
                  var(_cons)   18.90948   6.809269      9.336057    38.29973

               var(Residual)   71.20156    4.52476      62.86325    80.64588

LR test vs. linear model: chibar2(01) = 67.15         Prob >= chibar2 = 0.0000
```

- We would also like to store these for comparison later.

```
. estimates store wb
```

- Within schools, students who spent one more hour on homework each week than the typical student at their school had a 2.36-point [1.81, 2.91] higher predicted math achievement scores.
- Between schools, schools with students who spent an hour longer on average on homework each week had a 4.02-point [0.69, 7.36] higher average predicted math achievement score.
- This difference in coefficients indicates that school-wide changes to homework policies would have a stronger impact than student-level changes.
- If the cluster mean did not significantly contribute to the model, we could return to our original model and have higher power to detect that effect (Wang and Maxwell 2015).
- We can compare these results with the conflated model using `etable`.

```
. etable, estimates(conflated wb) column(estimates)

                                        conflated    wb

Time spent on math homework each week     2.402
                                         (0.277)
HWb                                                 4.023
                                                  (1.702)
HWw                                                 2.361
                                                  (0.280)
Intercept                                46.349    43.371
                                         (1.141)  (3.277)
var(_cons)                               20.226    18.909
                                         (7.071)  (6.809)
var(e)                                   71.144    71.202
                                         (4.517)  (4.525)
Number of observations                      519       519
```

- The conflated effect is always a weighted average of the within- and between-effect where the weight is based on the inverse of the variance of each estimate (Baltagi, 2008).

4 Exercise

1. Open the `patient.dta` dataset and run `codebook` to view each of the variables.
2. What is the nesting structure of these data?
   (a) Which variable specifies cluster?
   (b) Which variables vary at level 1? Which vary at level 2?
3. Fit an intercept-only random-intercept model for `satisfaction`. Calculate and interpret the intraclass correlation.
4. Fit the following multilevel model using `mixed`. Interpret the parameter estimates. Use `estat sd` to interpret the random effects.

$$\begin{aligned} \texttt{satisfaction}_{\texttt{ij}} &= \beta_{0j} + \beta_{1j}\texttt{minutes}_{\texttt{ij}} + \varepsilon_i \\ \beta_{0j} &= \gamma_{00} + \gamma_{01}\texttt{female}_{\texttt{i}} + u_{0j} \\ \beta_{1j} &= \gamma_{10} + \gamma_{11}\texttt{female}_{\texttt{i}} + u_{1j} \end{aligned}$$

5. Disaggregate the effect of `minutes` by instead including its cluster mean and cluster-mean centered variable in the model. Is the effect of spending more time with a patient larger between doctors or within doctor?

Copyright ©2025 StataCorp LLC

References

Baltagi, Badi (2008). *Econometric analysis of panel data*. John Wiley & Sons.

Bryk, Anthony S and Stephen W Raudenbush (1992). *Hierarchical linear models: Applications and data analysis methods.* Sage Publications, Inc.

Kreft, Ita GG and Jan De Leeuw (1998). *Introducing multilevel modeling.* Sage.

Rabe-Hesketh, Sophia and Anders Skrondal (2022). *Multilevel and longitudinal modeling using Stata*. 4th. College Station, TX: Stata Press.

Wang, Lijuan Peggy and Scott E Maxwell (2015). "On disaggregating between-person and within-person effects with longitudinal data using multilevel models." In: *Psychological methods* 20.1, p. 63.

Copyright ©2025 StataCorp LLC