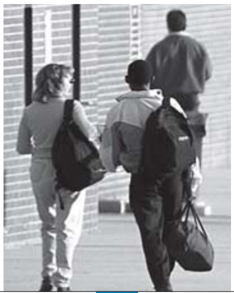


15

Multilevel Analyses

PISA DATA ANALYSIS MANUAL: SAS® SECOND EDITION – ISBN 978-92-64-05624-4 – © OECD 2009

INTRODUCTION

Over the last 20 years, education survey data have been increasingly analysed with multilevel models. Indeed, since simple linear regression models without taking into account the potential effects that may arise from the way in which students are assigned to schools or to classes within schools,[1] they may provide an incomplete or misleading representation of efficiency in education systems. In some countries, for instance, the socio-economic background of a student may partly determine the type of school that he or she attends and there may be little variation in the socio-economic background of students within each school. In other countries or systems, schools may draw on students from a wide range of socio-economic backgrounds, but within the school, the socio-economic background of the student impacts the type of class he or she is allocated to and, as a result, the within-school variance is affected. A linear regression model that does not take into account the hierarchical structure of the data will thus not differentiate between these two systems.

The use of multilevel models (Goldstein, 1995), also called hierarchical linear models (Bryk and Raudenbush, 1992), acknowledges the fact that students are nested within classes and schools. The relative variation in the outcome measures, between students within the same school and between schools can therefore be evaluated.

Figure 15.1 shows four graphs that highlight the distinction between a simple linear regression and a multilevel linear regression model. These four graphs represent the relationship between student socio-economic backgrounds and performance estimates in different countries; let's say for mathematics.

The thick black line represents the simple regression line when the hierarchical structure of the data is **not** taken into account. The thin blue lines represent the relationship between these two variables within particular schools. For each school, there is a regression line (the blue line in this example). The larger black dot on the simple linear regression lines (black) represents the point with the mean of *X* and *Y* as coordinates, $(\ _{x},\ _{y})$, and the blue point on the multilevel regression lines represents the point with the school mean of *X* and *Y* as coordinates, $(\ _{xi},\ _{yi})$.

The simple linear regression analysis, graphically represented by the black lines, shows that the expected score of a student from a higher socio-economic background is considerably higher than the expected score of a student from a lower socio-economic background. The comparison between the black lines on these four graphs shows the similarity of the relationship between the student's socio-economic background and student performance between countries. Based on simple linear regression analyses, therefore, the conclusion could be that the relationship between socio-economic background and student performance is identical in different countries.

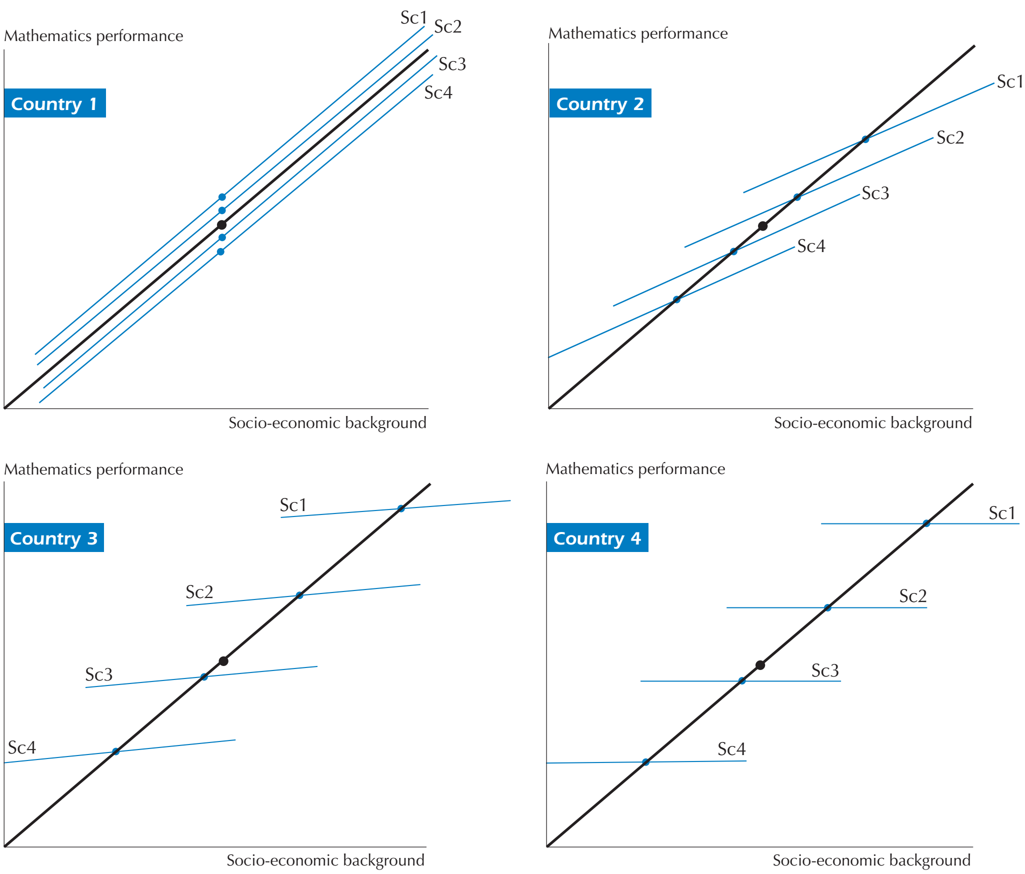
However, the multilevel regression analyses clearly distinguish the relationship between students' socio-economic backgrounds and their performance in the four countries.

In Country 1, the multilevel regression lines are similar and close to the simple linear regression line. This means that:

- Regarding the socio-economic background of the student (*X* axis):
  - The different schools are attended by students from a wide range of socio-economic backgrounds. All the within-school regression lines cover the whole range of values on the *X* axis.
  - The schools have the same socio-economic intake, *i.e.* the mean of the student socio-economic background. Indeed, the projections of the blue dots on the *X* axis are very close to each other.
  - In summary, there is no social segregation.

- Regarding the student performance in mathematics (*Y* axis):
  - In each school, there are low, medium, and high achievers. All the within-school regression lines cover the *Y* axis.
  - On average, the schools have a similar level of performance. Indeed, the projections of the blue dots on the *Y* axis are very close to each other. It also means that the school variance is quite small.
  - In summary, there is no academic segregation.
- Regarding the relationship between the socio-economic background and mathematics performance:
  - In each school, there is a positive relationship between the socio-economic background and achievement.
  - Within all schools, disadvantaged socio-economic background students perform well below students with advantaged socio-economic background students. The steep slope of the within-school regression line indicates that there is a relationship between students' socio-economic background and their performance.

Figure 15.1
Simple linear regression analysis versus multilevel regression analysis



Each school in Country 1 can therefore be considered as a simple random sample of the population and each school reflects the relationships that exist at the population level. Northern European countries tend generally to behave as the fictitious case of Country 1.

The opposite case of Country 1 is graphically represented by Country 4. The multilevel regression lines differ considerably from the simple linear regression line. In this case, it means that:

- Regarding the socio-economic background of the student (*X* axis):
  - The schools do not cover the range of socio-economic backgrounds that exist at the population level. School 1 is mainly attended by advantaged socio-economic background students while School 4 is mainly attended by disadvantaged socio-economic background students.
  - The schools have therefore different socio-economic intakes as the projections of the blue dots on the *X* axis show.
  - In summary, there is a significant social segregation at the school level.
- Regarding the student performance in mathematics (*Y* axis):
  - The schools do not cover the range of the student performance that exists at the population level. School 1 is mainly attended by high achievers and School 4 is mainly attended by low achievers.
  - Schools largely differ in their average performance level, as the projections of the blue dots on the *Y* axis show. In Country 4, the school performance variance is therefore very important.
  - In summary, there is a high academic segregation.
- Regarding the relationship between the socio-economic background and mathematics performance:
  - In each school, there is no relationship between socio-economic background and achievement.
  - What does matter is the school the student will attend knowing that the socio-economic background of the student will determine this school.

Countries 2 and 3 present intermediate situations between these two extreme examples.

TWO-LEVEL MODELLING WITH SAS®

Usually, two types of indices are relevant in multilevel analyses: *(i)* the regression coefficients, usually denoted as the fixed parameters of the model; and *(ii)* the variance estimates, usually denoted as the random parameters of the model. Any multilevel regression analysis should always begin with the computation of the Level 1 and Level 2 variance estimates for the dependent variable.

Decomposition of the variance in the empty model

The first recommended step in multilevel regression analysis consists of a decomposition of the variance of the dependent variable into the different levels. Here, as an example, the variance of the student performance in science will be decomposed into two components: the within-school variance and the between-school variance.

These two variance components can be obtained with an Mixed ANOVA (analysis of variance) model, as well as with a multilevel regression. The multilevel regression equation is equal to:

$$Y_{ij} = {}_{0j} + {}_{ij}$$

$${}_{0j} = \gamma_{00} + U_{0j}$$

with Y_{ij} representing the reading performance of student i in school j, β_{0j} the intercept for school j, ε_{ij} the student residual, γ_{00} the overall intercept and U_{0j} the school departure from the overall intercept. This model simply predicts the student performance by the average performance of his/her school and the school performance is predicted by the grand mean. Indeed, as the regression model has no predictors, the school intercepts, *i.e.* β_{0j} will therefore be equal or close to the school means. The variance of U_{0j}, usually denoted τ_{00} or τ_0^2, will be equal to the between-school variance. As each student will be assigned his/her school mean as predicted score, the variance of ε_{ij}, usually denoted σ^2, will be equal to the within-school variance.

The SAS® PROC MIXED procedure is devoted to multilevel regressions. However, it requires the normalisation of the weights, *i.e.* the sum of the weights is equal to the number of students in the dataset.[2] If the BY statement is used, then the normalisation will be done by category of the breakdown variable.

Box 15.1 provides the SAS® syntax for this normalisation, as well as a short checking procedure.

Box 15.1 **Normalisation of the final student weights (*e.g.* PISA 2006)**

```
libname PISA2006  "c:\pisa\2006\data\";
options nofmterr notes;
run;
data temp1;
        set pisa2006.stu;
        keep cnt schoolid stidstd w_fstuwt pv1scie;
run;
proc sort data=temp1;
        by cnt;
run;
proc univariate data=temp1 noprint;
        var w_fstuwt;
        by cnt;
        output out=temp2 sum=wgt N=nbre;
run;
data temp3;
        merge temp1 temp2;
        by cnt;
        std_wgt=(w_fstuwt*nbre)/wgt;
run;

/* VERIFICATION */

proc means data=temp3 noprint;
        var std_wgt;
        by cnt;
        output out=cnt N=nbstud sum=wgtsum;
run;
proc print data=cnt;
        var nbstud wgtsum;
run;
```

Box 15.2 provides the SAS® syntax for a multilevel regression model as well as the SAS® syntax for the computation of the intraclass correlation.

Box 15.2 **SAS® syntax for the decomposition of the variance in student performance in science (*e.g.* PISA 2006)**

```
proc mixed data= temp3 method=ml;
        class schoolid;
        model pv1scie = /solution;
        random intercept/subject=schoolid solution;
        weight std_wgt;
        by cnt;
        ods output covparms=decompvar solutionf=fixparm solutionr=ranparm;
run;
proc transpose data=decompvar out=rho;
        var estimate;
        by cnt;
        id covparm;
run;
data rho;
        set rho;
        rho=intercept/(intercept+residual);
        keep cnt intercept residual rho;
run;
proc print data=rho;
run;
```

The **`class`** statement defines the second level of the analyses. Similar to all linear models, the **`model`** statement specifies the dependent and independent variables. In this particular example, there is no predictor. Therefore the between-school and within-school residual variances will be equal to the between-school and within-school variance estimates. The **`random`** statement distinguishes between fixed and random predictors, as explained in the previous section. It should be noted that **`"intercept"`** always needs to be mentioned. The **`weight`** and the **`by`** statements are self-explanatory. Finally, the **`ods`** statement will save the results in three data files. The variance estimates will be saved in the file **`"decompvar"`**, the fixed parameters will be saved in the file **`"fixparm"`** and the random parameters will be saved in the file **`"ranparm"`**.

Table 15.1 provides the between-school and within-school variance estimates and the intraclass correlation. These variance estimates were saved in the file **`"decompvar"`**. As shown in Box 15.2, the intraclass correlation is equal to:

$$\rho = \frac{\sigma^2_{between\text{-}school}}{\sigma^2_{between\text{-}school} + \sigma^2_{within\text{-}school}} = \frac{\tau_0^2}{\tau_0^2 + \sigma^2}$$

with $\sigma^2_{between\text{-}school}$ or τ_0^2 the between-school variance and $\sigma^2_{within\text{-}school}$ or ^{2} the within-school variance.

In Australia, the between-school variance is equal to 1 793 and the within-school variance is equal to 8 263. In Australia, the intraclass correlation is therefore equal to 1 793/(1 793 + 8 263) = 0.18. The intraclass correlation is the percentage of the total variance that is accounted for by the school. It reflects how schools differ in their student average performance. The estimate of the intraclass correlation ranges among countries from 0.06 in Finland to 0.61 in Hungary.

If the Level 2 variance is equal to 0, a multilevel regression would be mathematically equal to a linear regression. As the between-school variance becomes larger, the differences between these two regression models increase. Knowing the intraclass correlation will therefore help the researcher correctly interpret the results.

PISA DATA ANALYSIS MANUAL: SAS® SECOND EDITION – ISBN 978-92-64-05624-4 – © OECD 2009

Table 15.1

Between- and within-school variance estimates and intraclass correlation (PISA 2006)

Country	Between-school variance	Within-school variance	rho (intraclass correlation)
AUS	1 793.90	8 263.15	0.18
AUT	5 417.72	4 487.38	0.55
BEL	5 128.06	4 776.88	0.52
CAN	1 659.45	7 121.52	0.19
CHE	3 341.69	5 900.62	0.36
CZE	5 576.30	5 068.80	0.52
DEU	5 979.48	4 483.83	0.57
DNK	1 411.05	7 313.88	0.16
ESP	1 131.29	6 663.92	0.15
FIN	424.32	6 958.82	0.06
FRA	5 547.85	4 711.89	0.54
GBR	2 169.93	8 925.47	0.20
GRC	4 467.86	5 054.07	0.47
HUN	5 450.09	3 461.37	0.61
IRL	1 496.87	7 551.06	0.17
ISL	887.96	8 641.69	0.09
ITA	4 803.95	4 657.73	0.51
JPN	4 769.06	5 326.91	0.47
KOR	2 881.59	5 353.91	0.35
LUX	2 752.47	6 584.74	0.29
MEX	2 281.54	3 462.17	0.40
NLD	5 343.29	3 525.86	0.60
NOR	947.58	8 338.64	0.10
NZL	1 913.37	9 702.39	0.16
POL	1 113.91	7 107.97	0.14
PRT	2 480.24	5 234.37	0.32
SVK	3 644.47	5 059.48	0.42
SWE	1 034.22	7 863.48	0.12
TUR	3 702.19	3 199.89	0.54
USA	2 610.97	8 529.74	0.23

Models with only random intercepts

The following examples are based on the PISA 2006 data in Belgium.

In the PISA databases, there are no missing data for the final weight and for the student performance estimates. However, contextual variables that might be used as predictors in a multilevel regression model usually have missing data. These missing data generate two major issues:

- The sum of the weights will slightly differ from the number of cases that will be used by the regression models. Note that cases with missing values are usually dropped from regression models.[3]
- The school and student variances from different models cannot be compared as missing values are not always random. For instance, disadvantaged socio-economic background students are usually less likely to provide answers about their mother's and/or father's occupations.

To avoid these two problems, it is recommended to delete any cases with missing data for the different predictors that will be used in the regression models before normalising the weights.

At the student level, different variables were included in the temporary file:

- The variable of ST01Q01 indicates students' grades.

- The variable of GENDER indicates students' gender derived from ST04Q01:
  - Value 0 is assigned to males.
  - Value 1 is assigned to females.
- The variable of IMIG indicates students' immigrant status derived from ST11Q01 to 03:
  - Value 1 is assigned to a student whose parents were born in a country other than Belgium.
  - Value 0 if the student was born in Belgium and at least one of the parents was also born in Belgium.
- The variable of ESCS indicates the PISA index of economic, social and cultural status for students.
- The variable of VOCATION indicates students' programme orientation derived from ISCEDO:
  - Value 0 is assigned to students enrolled in academic programmes.
  - Value 1 is assigned to students enrolled in pre-vocational and vocational programmes.

At the school level, three variables were derived:

- The variable of MU_ESCS indicates schools' socio-economic intake measured by the school average ESCS.
- The variable of PPCT_IM indicates the proportion of students with an immigrant background in the school.
- The variable of TYPE indicates the school type:
  - Value 1 is assigned to schools that propose only academic programmes.
  - Value 0 is assigned to schools that propose pre-vocational or vocational programs.

Box 15.3 presents the SAS® syntax for the preparation of the data file.

As mentioned earlier in this chapter, the first step in multilevel modelling consists of running a regression without any independent variables. This model will return the estimate of the between-school and within-school variances. In Belgium, the between-school variance is equal to 5 010 and the within-school variance is equal to 4 656, as saved in the **"decompvar1"** file. It should be noted that the variance estimates have to be computed after the deletion of cases with missing data. Indeed, as residual variances will be compared between different regressions, it is of prime importance that the different models be computed exactly on the same dataset.

The **"fixparm1"** file contains the fixed parameters. With an empty model, it presents γ_{00}, *i.e.* 510.78 for the data in Belgium.

The **"ranparm1"** file lists the random parameters. With an empty model, only the school departure U_{0j} will be listed. Table 15.2 is a printout of the **"ranparm1"** file for the first ten cases. It contains:

- the breakdown variables used in the model, *i.e.* CNT;
- the effect, *i.e.* the intercept or as it will be shown later, the random predictor, the estimate;
- the class variable, *i.e.* the SCHOOLID;
- the estimate;
- the standard error on the estimate;
- the number of degrees of freedom (the number of students minus the number of schools);
- the *t* statistic;
- the probability that the estimates differ from 0.

Box 15.3 [1/2] **SAS® syntax for normalising PISA 2006 final student weights with deletion of cases with missing values and syntax for variance decomposition (*e.g.* PISA 2006)**

```
data temp4;
      set  pisa2006.stu;
      if (cnt="BEL");

      if (st01Q01 not in (7,8,9,10,11,12,13,14)) then st01Q01=.;

      gender=.;
      if (st04q01 in (1)) then gender=1;
      if (st04q01 in (2)) then gender=0;

      if (st11q01 in (.,.M,.N,.I)) then st11q01=9;
      if (st11q02 in (.,.M,.N,.I)) then st11q02=9;
      if (st11q03 in (.,.M,.N,.I)) then st11q03=9;
      immig=(100*st11q01)+(10*st11q02)+(st11q03);

      img=.;
      if (immig in (111,121,112)) then img=0;
      if (immig in (122,222)) then img=1;

      vocation=.;
      if (iscedo in (1)) then vocation=0;
      if (iscedo in (2,3)) then vocation=1;

      nbmis=0;
      array vecmis (5) vocation st04q01 st01Q01 escs img;
      do i=1 to 5;
      if (vecmis(i) in (.,.N,.I,.M)) then nbmis=nbmis+1;
      end;

      if (nbmis=0);

      scie1=pv1scie;
      scie2=pv2scie;
      scie3=pv3scie;
      scie4=pv4scie;
      scie5=pv5scie;
      w_fstr0=w_fstuwt;
      keep  CNT SCHOOLID stidstd
            scie1-scie5  w_fstr0-w_fstr80
               vocation gender st01Q01 escs img;
run;

proc sort data=temp4;
      by cnt schoolid stidstd;
run;

proc univariate data=temp4 noprint;
      var w_fstr0;
      by cnt;
      output out=temp5 sum=somwgt n=nbre;
run;
```

For instance, the departure of the school 2 from the overall intercept is 45.27. This departure statistically differs from 0, as shown by the *t* statistic and its associated probability value. In other words, the intercept of school 2 is significantly different from the overall intercept. On the other hand, the intercept of school 1 is not significantly different from the overall intercept.

Box 15.3 [2/2] SAS® syntax for normalising PISA 2006 final student weights with deletion of cases with missing values and syntax for variance decomposition (*e.g.* PISA 2006)

```
data temp6;
      merge temp4 temp5;
      by cnt;
          array wgt (81) w_fstr0-w_fstr80;
          do i=1 to 81;
            wgt(i) =(wgt(i)/somwgt)*nbre;
      end;
run;
proc univariate data=temp6 noprint vardef=wgt;
      weight w_fstr0;
      by cnt schoolid;
      var img escs;
      output out=temp7 mean=pct_im mu_escs;
run;
proc freq data=temp6 noprint;
      table vocation/out=temp8;
      by cnt schoolid;
run;
proc transpose data=temp8 out=temp9;
      var count;
      by cnt schoolid;
      id vocation;
run;
data temp10;
      set temp9;
      if (_0=.)then _0=0;
      if (_1=.)then _1=0;
      type=0;
      if (_1>0) then type=1;
      keep cnt schoolid type;
run;
data temp11;
      merge temp7 temp10;
      by cnt schoolid;
run;
data temp12;
        merge temp6 temp11;
        by cnt schoolid;
run;
/*Variance decomposition*/
proc mixed data=temp12 method=ml;
        class schoolid;
        model scie1=/solution;
        random intercept/subject=schoolid solution;
        weight w_fstr0;
        by cnt;
        ods output covparms=decompvar1 solutionf=fixparm1 solutionr=ranparm1;
run;
```

Table 15.2
Output data file "ranparm1" from Box 15.3

CNT	Effect	SCHOOLID	Estimate	StdErrPred	Degrees of freedom	tValue	Probability
BEL	Intercept	1	-32.3468	27.5127	8113	-1.18	0.2397
BEL	Intercept	2	45.2674	11.9007	8113	3.80	0.0001
BEL	Intercept	3	16.2277	13.9353	8113	1.16	0.2443
BEL	Intercept	4	-13.7326	12.4275	8113	-1.11	0.2692
BEL	Intercept	5	31.4794	11.6458	8113	2.70	0.0069
BEL	Intercept	6	25.2378	13.1742	8113	1.92	0.0554
BEL	Intercept	7	111.2300	12.7806	8113	8.70	<.0001
BEL	Intercept	8	-20.3494	15.1814	8113	-1.34	0.1801
BEL	Intercept	9	69.3355	10.1656	8113	6.82	<.0001
BEL	Intercept	10	16.7966	12.2150	8113	1.38	0.1691

Shrinkage factor

In the case of an empty model, it might be considered that the sum of the overall intercept γ_{00} and a particular school departure U_{0j} should be perfectly equal to the school performance mean.

Multilevel models shrink the school departures. To illustrate this shrinkage process, let's suppose that we have an educational system with 100 schools and that the school performance means are perfectly identical. In other words, the school variance is equal to 0. If 20 students are tested within each school, it is expected that school mean estimates will differ slightly from the school means. Indeed, within particular schools, predominantly high performers or low performers may be sampled so that the school mean is respectively overestimated or underestimated. As the number of sampled students within schools increases, the difference between the school mean and its estimate is likely to decrease. Therefore, the shrinkage factor is inversely proportional to the number of sampled students within schools.

The shrinkage factor is equal to:

$$\frac{n_j \sigma^2_{between\text{-}school}}{n_j \sigma^2_{between\text{-}school} + \sigma^2_{within\text{-}school}}$$

with n_j being the number of students in school j in the sample (Goldstein, 1997).

This shrinkage factor ranges from 0 to 1. As it multiplies the school departure, the shrinkage will:

- depend on the ratio between between-school and within-school variance ;
- be proportional to the school departure, *i.e.* the shrinkage factor mainly affects low and high performing schools;
- be inversely proportional to the number of observed students in the school.

The between-school variance can also be estimated with an ANOVA. Mathematically, the between-school variance will be equal to:

$$\sigma^2_{between\text{-}school} = \frac{MS_{between\text{-}school} - MS_{within\text{-}school}}{n_j}$$

As it can be depicted from the ANOVA formula for estimating the between-school variance, the correction is also proportional to the within unit variance and inversely proportional to the number of cases sampled from each unit.

Models with random intercepts and fixed slopes

With the introduction of the student-level variable ESCS as a fixed effect, the equation can be written as:

$$Y_{ij} = \beta_{0j} + \beta_{1j}\,(ESCS)_{ij} + \varepsilon_{ij}$$

$$\beta_{0j} = \gamma_{00} + U_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

This model has two random components, *i.e. (i)* the variance of ε_{ij}, denoted ^{2}; and *(ii)* the variance of U_{0j}, denoted τ_{00}; and two fixed parameters, *i.e.* γ_{00} and γ_{10}. The SAS® syntax for this model is presented in Box 15.4 and parts of the SAS® output are presented in Box 15.5.

Box 15.4 **SAS® syntax for a multilevel regression model with random intercepts and fixed slopes (*e.g.* PISA 2006)**

```
proc mixed data= temp12 method=ml;
        class schoolid;
        model scie1 = escs/solution;
        random intercept/subject=schoolid solution;
        weight w_fstr0;
        by cnt;
        ods output covparms=decompvar2 solutionf=fixparm2 solutionr=ranparm2;
run;
```

Box 15.5 **SAS® output for the multilevel model in Box 15.4**

```
                         Covariance parameter estimates
                     Cov parm         Subject        Estimate
                     Intercept        SCHOOLID       3 971.20
                     Residual                        4 475.09

                           Solution for fixed effects
      Effect       Estimate    Standard error     DF       t Value    Pr>|t|
      Intercept     507.90         3.9345         268      129.09     <.0001
      ESCS         18.9064         0.9502        8112       19.90     <.0001
```

Only one change has been introduced in comparison with the syntax presented at the end of Box 15.3. The name ESCS has been added to the model statement.

The overall intercept γ_{00} is now equal to 507.90 and the within-school regression coefficient γ_{10} is equal to 18.9064. This means that, within a particular school, an increase of one unit on the ESCS index will be associated with an increase of 18.9064 on the science scale. By comparison, the linear regression coefficient of ESCS on the science performance is equal to 47.38.[4] It appears that the education system in Belgium behaves in a similar manner to fictional Country 3 presented in Figure 15.2.

The between-school and within-school residual variable estimates, respectively denoted τ_{00} and ^{2}, are equal to 3 971 and 4 475. In the empty model, the between-school variance is 5010 and the within-school variance is 4656.

The percentage of variance explained by the ESCS variable can be computed as:

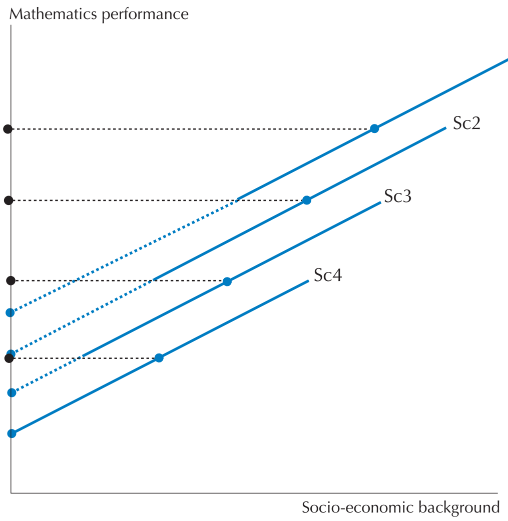
$1-\frac{3971}{5010}=0.21$ at the school level, and

$1-\frac{4475}{4656}=0.04$ at the student level.

How can a student-level variable explain about 21% of the between-school variance and only 4% of the within-school variance? This mainly reflects the school socio-economic background segregation. Some of the schools are mainly attended by advantaged socio-economic background students, while other schools are mainly attended by disadvantaged socio-economic background students.

Figure 15.2 provides a graphical explanation of this phenomenon. In any case, the between-school variance can be graphically represented by the variability of the school intercepts on the *Y* axis. In the case of an empty model, the intercept is close to the orthogonal projection of the school performance average on the *Y* axis, as shown by the black line in Figure 15.2. As explained in the previous section, the difference between the school mean and the intercept results from the application of the shrinkage factor.

Figure 15.2
Graphical representation of the between-school variance reduction



The between-school residual variance can be obtained by the extension of the regression line on the *Y* axis, as shown by the blue discontinuous line in Figure 15.2. As shown, the range of the black intercepts is larger than the range of the blue intercepts.

Broadly speaking, a student-level variable will have an impact on the between-school variance if:

- Schools differ in the mean and range of students with regard to the student-level variable (see Countries 2, 3 and 4 in Figure 15.1).
- The within-school regression coefficient of the student-level variable differs from 0. Country 4 in Figure 15.1 illustrates a case where using the ESCS variable at the student level in the model will not reduce the between-school variance. On the other hand, the introduction of the school socio-economic intake, *i.e.* the school ESCS mean, will have a substantial impact on the between-school variance.

Models with random intercepts and random slopes

In the cases examined so far, the within-school regression lines were all parallel, but multilevel regression analyses also allowed the regression slopes to vary. In the former, the effect, *i.e.* the *X* effect, will be considered as fixed, while in the latter, the effect will be considered as random. Figure 15.3 presents a case with a random effect.

Usually, empirical data should better fit a model that does not force the parallelism of the within-school regression lines. On the other hand, it implies that more parameters have to be estimated and that therefore convergence might not be reached.

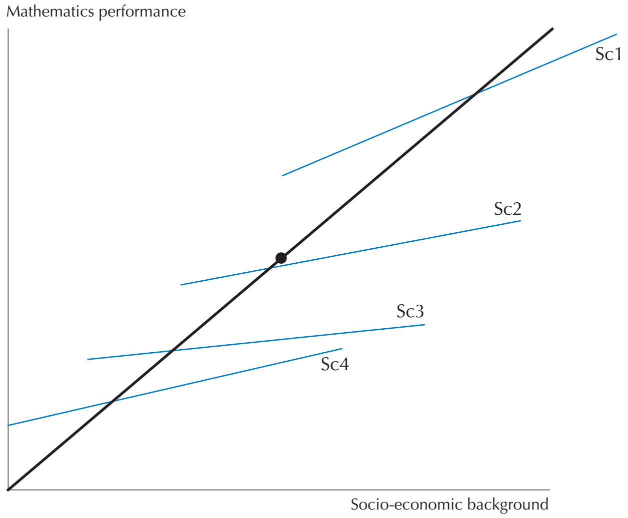
As demonstrated in the previous section, the student performance within a particular school is influenced by his/her socio-economic background. Schools may thus not be considered as equitable as expected by educational policies. One might further investigate if schools differ in terms of inequity. Are there schools that appear to be more equitable than others? This question can be answered by considering the ESCS slope as random and then testing if its variance significantly differs from 0. With the ESCS slope as random, the equation can be written as:

$$Y_{ij} = \alpha_j + \beta_{1j}\,(ESCS)_{ij} + \varepsilon_{ij}$$

$$\alpha_j = \gamma_{00} + U_{0j}$$

$$\beta_{1j} = \gamma_{10} + U_{1j}$$

Figure 15.3
A random multilevel model



Box 15.6 sets out the SAS® syntax for a multilevel regression model.

Box 15.6 **SAS® syntax for a multilevel regression model (*e.g.* PISA 2006)**

```
proc mixed data= temp12 method=ml cl covtest;
   class schoolid;
   model scie1 = escs/solution;
   random intercept escs/subject=schoolid solution ;
   weight w_fstr0;
   by cnt;
   ods output covparms=decompvar3 solutionf=fixparm3 solutionr=ranparm3;
run;
```

The variable ESCS has been added to the random statement. The standard error and a confidence interval for random parameters can be obtained by adding two options in the **proc mixed** statement: **cl** and **covtest**.

The fixed parameter file contains the overall intercept γ_{00} and the ESCS overall regression coefficient γ_{10}. Similar to the school intercepts which are divided into two parts – an overall intercept and a school departure – the within-school regression coefficient is divided into two parts: an overall regression coefficient (the fixed part, denoted γ_{10}) and a school regression coefficient departure (the random part, denoted U_{1j}).

PISA DATA ANALYSIS MANUAL: SAS® SECOND EDITION – ISBN 978-92-64-05624-4 – © OECD 2009

The overall intercept and regression coefficient are presented in Table 15.3. The overall intercept is equal to 508.05 and the overall ESCS regression coefficient is equal to 18.718. As shown by the *t* statistic and its associated probability, both parameters are significantly different from 0.

Table 15.3
Output data file "fixparm3" from Box 15.6

CNT	Effect	Estimate	S.E.	Degrees of freedom	tValue	Probability
BEL	Intercept	508.05	3.9602	268	128.29	<.0001
BEL	ESCS	18.719	1.1499	268	16.28	<.0001

The random parameter file lists the school departures:

- U_{0j} from the intercept γ_{00}, *i.e.* 508.05,
- U_{1j} from the ESCS regression coefficient γ_{10}, *i.e.* 18.72.

Table 15.4 presents the school departure from the overall ESCS regression coefficient for the first ten schools.

Table 15.4
Output data file "ranparm3" from Box 15.6

CNT	Effect	SCHOOL	Estimate	StdErrPred	Degrees of freedom	tValue	Probability
BEL	ESCS	1	0.3412	9.7250	7844	0.04	0.9720
BEL	ESCS	2	0.2670	8.1305	7844	0.03	0.9738
BEL	ESCS	3	-1.6771	9.1097	7844	-0.18	0.8539
BEL	ESCS	4	-8.1808	8.3040	7844	-0.99	0.3246
BEL	ESCS	5	0.6080	7.7785	7844	0.08	0.9377
BEL	ESCS	6	-2.0933	8.7431	7844	-0.24	0.8108
BEL	ESCS	7	-0.2759	8.6122	7844	-0.03	0.9744
BEL	ESCS	8	2.8939	9.0724	7844	0.32	0.7498
BEL	ESCS	9	-0.4817	8.1833	7844	-0.06	0.9531
BEL	ESCS	10	-1.1952	8.3590	7844	-0.14	0.8863

The ESCS regression coefficient for school 1 is equal to 18.718 + 0.267 = 18.985, but it cannot be considered as significantly different from the overall intercept. Of the 269 schools, only 2 schools present a regression coefficient that significantly differs from the overall coefficient.

SAS® now provides three variance estimates. Box 15.7 presents these estimates and their related information.

- the between-school residual variance τ_0^2, *i.e.* 4 009;
- the within-school residual variance ^{2}, *i.e.* 4 411;
- the variance of ESCS regression coefficients τ_1^2, *i.e.* 99.

Box 15.7 **SAS® output for the multilevel model in Box 15.6**

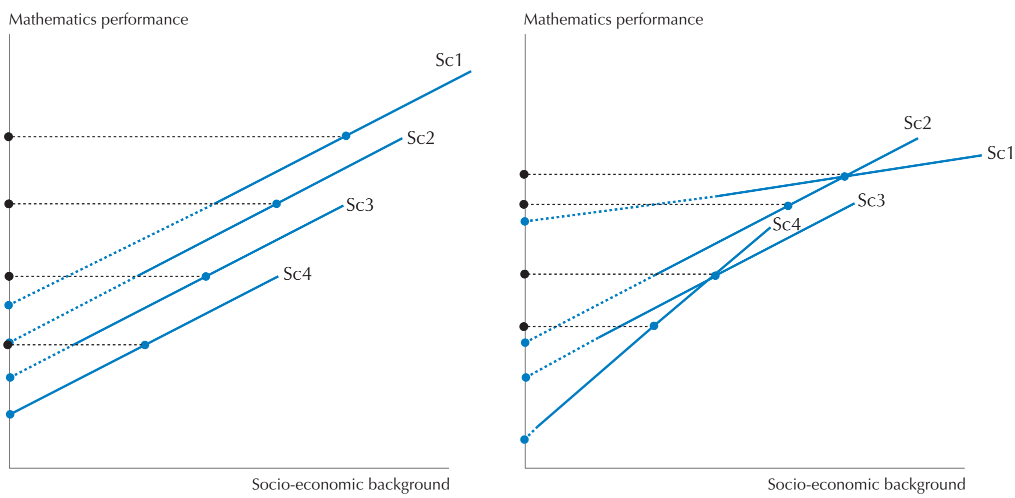
```
                   Covariance parameter estimates

                              Standard
Cov parm  Subject   Estimate   error   Z Value   Pr Z   Alpha   Lower     Upper
Intercept SCHOOLID  4 009.43   366.99   10.93   <.0001   0.05  3 377.22 4 838.43
ESCS      SCHOOLID   99.0653  27.3568    3.62   0.0001   0.05   61.5503   185.47
Residual            4 411.33  70.2516   62.79   <.0001   0.05  4 276.81 4 552.32
```

As variance parameters have a default lower boundary constraint of 0, their confidence intervals are not symmetric. The ESCS confidence interval does not include 0; the null hypothesis can therefore be rejected with a type I error risk of 0.05.

In comparison with previous results, the between-school residual variance has slightly increased (from 3 971 to 4 009) and the within-school residual variance has decreased slightly (from 4 475 to 4 411). The reduction of the within-school variance is not surprising as the random effect can only better fit the data. The increase in the school variance in this particular example is negligible, but in some cases it might be substantial. Figure 15.4 helps to understand and interpret a substantial increase of the between-school residual variance, by showing that the range of the projections of the red lines on the *Y* axis varies more in the random slope model than in the fixed slope model. The school intercepts and the school slopes might correlate. In Figure 15.4, the correlation between intercepts and slopes is negative: lower performing schools have deeper slopes and higher performing schools have flatter slopes. This would mean that higher performing schools are more equitable. The correlation between the intercept and the slope could also be positive: in that case, this would mean that lower performing schools are more equitable.

Figure 15.4
Change in the between-school residual variance for a fixed and a random model



The SAS® option **type=un** in the **random** statement will return estimates of the covariance between random parameters. Box 15.8 presents the SAS® output. Without **type=UN**, the covariance between random parameters is set to 0.

Box 15.8 **SAS® output for the multilevel model with covariance between random parameters**

Covariance parameter estimates

Cov parm	Subject	Estimate	Standard error	Z Value	Pr Z	Alpha	Lower	Upper
UN(1,1)	SCHOOLID	4 005.67	366.39	10.93	<.0001	0.05	3 374.43	4 833.23
UN(2,1)	SCHOOLID	-0.5365	73.8023	-0.01	0.9942	0.05	-145.19	144.11
UN(2,2)	SCHOOLID	99.0195	27.3829	3.62	0.0001	0.05	61.4854	185.57
Residual		4 411.48	70.2542	62.79	<.0001	0.05	4276.96	4 552.48

PISA DATA ANALYSIS MANUAL: SAS® SECOND EDITION – ISBN 978-92-64-05624-4 – © OECD 2009

`UN(1,1)` corresponds to the intercept variance, `UN(2,2)` corresponds to the ESCS regression coefficient slope variance and `UN(2,1)` corresponds to the covariance between the intercepts and the slopes. In this example, 0 is included in the confidence interval for `UN(2,1)` and therefore the null hypothesis cannot be rejected.

Suppose that the regression lines in Figure 15.4 are moved 5 cm on the right. The variance of the intercept will be unchanged for the fixed model, but will increase for the random model. Broadly speaking, as the mean of the independent variable differs from 0, the impact of considering it as random on the school variance will increase. Centring the independent variables on 0 limits the changes in the between-school variance estimates. Table 15.5 presents the variance/covariance estimates on the international socio-economic index of occupational status (HISEI) for Belgium. As a reminder, HISEI averages around 50 and has a standard deviation of 15. The left part of Table 15.5 presents the estimates before centring, the right part, after centring on 0. Three models were implemented: *(i)* HISEI as a fixed factor in Model 1; *(ii)* HISEI as a random factor, but no estimation of the covariance in Model 2; and *(iii)* HISEI as a random factor and estimation of the covariance in Model 3.

As illustrated by Table 15.5, centring the independent variable limits changes in the school variance estimates.

Table 15.5
Variance/covariance estimates before and after centering

Estimate		Before centering			After centering		
		Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Intercept variance	τ_{00}	4 112	3 920	4 653	4 112	4 114	4 109
HISEI regression coefficient slope variance	τ_{11}		0.13	0.22		0.22	0.22
Covariance between the intercepts and the slopes	τ_{10}			-10.80			-0.04
Residual	2	4 506	4 472	4 457	4 506	4 457	4 457

Bryk and Raudenbush (1992) distinguish three main locations for the Level 1 independent variables:

- **The natural X metric**: it can only be meaningful if cases with the value 0 on the *X* variable can be observed. Otherwise, the intercept that represents the score on *Y* for a subject with 0 on the *X* variables will be meaningless.
- **Centring around the grand mean**: it consists of transforming the original variables so that their means will be equal to 0. The intercepts will therefore represent the score on *Y* for a subject whose values on the *X* variable are equal to the grand mean.
- **Centring around the Level 2 mean (group-mean centring)**: it consists of transforming the original variables so that their means will be equal to 0 for each school. With this approach, the introduction of Level 1 variables does not affect the between-school variance. For instance, the introduction of the ESCS variable as a fixed effect in the model decreases the between-school variance by around 20%. It reflects a segregation effect of the students based on their economic, social and cultural status. Such effect cannot be observed if Level 1 independent variables are group-mean centred.

In the following model, the student gender, denoted GENDER with males being 0 and females being 1 in the PISA 2006 database, is added as a random factor to the previous model. The equation can be written as:

$Y_{ij} = \beta_{0j} + \beta_{1j}\,(ESCS)_{ij} + \beta_{2j}\,(GENDER)_{ij}$

$\beta_{0j} = \gamma_{00} + U_{0j}$

$\beta_{1j} = \gamma_{10} + U_{1j}$

$\beta_{2j} = \gamma_{20} + U_{2j}$

The fixed parameters are respectively equal to 514.53 for the overall intercept γ_{00}, 18.06 for the overall ESCS regression coefficient γ_{10} and –13.01 for the overall gender coefficient γ_{20}.

The between-school residual variance τ_{00} is equal to 4 144 and the within-school residual variance ^{2} is equal to 4 344. Finally, the variance of the school ESCS regression coefficient τ_{11} is equal to 102 and the variance of the school GENDER regression coefficient τ_{22} is equal to 140. All these variance estimates differ statistically from 0.

The gender regression coefficient of –13.01 reflects the expected gender difference within any school, after controlling for ESCS.

Box 15.9 **Interpretation of the within-school regression coefficient**

The expected within-school gender difference can differ greatly from the overall gender difference, especially in a highly tracked system. It appears that girls are more likely to attend an academic track while boys are more likely to attend a vocational track. The linear regression coefficient of gender on the student performance provides an estimates of the overall gender difference, while a multilevel regression model estimates gender difference after accounting for the differential attendance to school. Therefore, the gender multilevel regression coefficient will substantially differ from the linear regression coefficient. The table below provides the linear and multilevel regression coefficients for gender on the data from Germany.

At the population level, males outperform females by 6.2 in science while females outperform males by 42.6 in reading. But within a particular school, the expected differences in science and reading are respectively equal to 16.6 and 31.9.

Gender differences in Germany (females – males)

	Science	Reading	Mathematics
Linear regression	–6.2	42.6	–18.8
Multilevel regression	–16.6	31.9	–28.9

Models with Level 2 independent variables

The last equation was $Y_{ij} = \beta_{0j} + \beta_{1j}\,(ESCS)_{ij} + \beta_{2j}\,(GENDER)_{ij} + \varepsilon_{ij}$. This equation mainly models the student performance variability within schools by introducing student-level predictors. However, due to the segregation effect, these student-level predictors can explain some of the between-school variance. It is also possible to introduce school-level predictors.

First, it is important to understand why some schools perform well and others less so. One usually predicts the school intercept by the school socio-economic intake, *i.e.* the mean of the student economical, social and cultural status. The impact of school socio-economic intake is usually denoted *peer effect* or *school composition effect*. A student's social and academic context may influence his/her behaviour. In a school where most of the students spend hours working on homework, it is likely s/he will work hard. Conversely, s/he will probably not work hard in a school where most of the students skip class.

PISA DATA ANALYSIS MANUAL: SAS® SECOND EDITION – ISBN 978-92-64-05624-4 – © OECD 2009

Mathematically, testing the influence of the school socio-economic intake can be written as:

$$Y_{ij} = \beta_{0j} + \beta_{1j}(ESCS)_{ij} + \beta_{2j}(GENDER)_{ij} + \varepsilon_{ij}$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(mu\_ESCS)_j + U_{0j}$$

$$\beta_{1j} = \gamma_{10} + U_{1j}$$

$$\beta_{2j} = \gamma_{20} + U_{2j}$$

with mu_ESCS representing the school average of the PISA index of economic, social and cultural status for a student. The SAS® syntax is presented in Box 15.10.

Box 15.10 **SAS® syntax for a multilevel regression model with a school-level variable (*e.g.* PISA 2006)**

```
proc mixed data= temp12 method=ml cl covtest;
        class schoolid;
        model scie1 = escs gender mu_escs/solution;
        random intercept escs gender/subject=schoolid solution ;
        weight w_fstr0;
        by cnt;
        ods output covparms=decompvar4 solutionf=fixparm4 solutionr=ranparm4;
run;
```

Table 15.6 presents the results for the fixed parameters.

Table 15.6
Output data file of the fixed parameters file

CNT	Effect	Estimate	S.E.	Degrees of freedom	tValue	Probability
BEL	Intercept	498.31	2.7255	267	182.83	<.0001
BEL	ESCS	16.48	1.1612	268	14.19	<.0001
BEL	gender	-12.87	1.8127	244	-7.10	<.0001
BEL	mu_escs	105.89	5.1549	7 599	20.54	<.0001

As shown in Table 15.6, the regression coefficient of the school socio-economic intake is highly significant. For two students with a similar ESCS background but attending two schools that differ by one index point in the average ESCS, their performance will differ by 106 points.

It might be useful to understand why some schools appear to be more equitable, *i.e.* schools with lower ESCS and/or GENDER regression coefficients, and why some other schools appear to be more inequitable, *i.e.* with higher ESCS and/or GENDER regression coefficients.

In this example, the impact of the school socio-economic intake and the school type will be tested. Mathematically, the model can be written as:

$$Y_{ij} = \beta_{0j} + \beta_{1j}(ESCS)_{ij} + \beta_{2j}(ST03Q01)_{ij} + \varepsilon_{ij}$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(TYPE)_j + \gamma_{02}(mu\_ESCS)_j + U_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}(TYPE)_j + \gamma_{12}(mu\_ESCS)_j + U_{1j}$$

$$\beta_{2j} = \gamma_{20} + \gamma_{21}(TYPE)_j + \gamma_{22}(mu\_ESCS)_j + U_{2j}$$

Box 15.11 presents the SAS® syntax for running this model. Testing the influence of the school type on the ESCS regression coefficients requires modelling the interaction between these two variables. Usually, this interaction is denoted a cross-level interaction. Box 15.12 presents the SAS® output.

Box 15.11 **SAS® syntax for a multilevel regression model with interaction (*e.g.* PISA 2006)**

```
proc mixed data= temp12 method=ml cl covtest;
        class schoolid;
        model scie1 = escs gender
                    type mu_escs
                    escs*type escs*mu_escs
                    gender*type gender*mu_escs /solution;
        random intercept escs gender /subject=schoolid   ;
        weight w_fstr0;
        by cnt;
        ods output covparms=decompvar5 solutionf=fixparm5 solutionr=ranparm5;
run;
```

Box 15.12 **SAS® output for the multilevel model in Box 15.11**

Covariance parameter estimates

Cov parm	Subject	Estimate	Standard error	Z Value	Pr Z	Alpha	Lower	Upper
Intercept	SCHOOLID	1 338.93	134.27	9.97	<.0001	0.05	1 110.32	1 646.60
ESCS	SCHOOLID	89.9421	26.0773	3.45	0.0003	0.05	54.7338	174.65
gender	SCHOOLID	87.6427	58.2604	1.50	0.0662	0.05	32.9396	602.44
Residual		4 348.94	70.2643	61.89	<.0001	0.05	4 214.45	4 490.01

Solution for fixed effects

Effect	Estimate	Standard error	DF	t Value	Pr>|t|
Intercept	522.37	6.2466	266	83.62	<.0001
ESCS	10.6204	2.6153	266	4.06	<.0001
gender	-16.1028	3.9816	242	-4.04	<.0001
type	-30.8790	7.0795	7 599	-4.36	<.0001
mu_escs	94.3191	6.6287	7 599	14.23	<.0001
ESCS*type	8.4097	3.0270	7 599	2.78	0.0055
ESCS*mu_escs	0.07536	2.7466	7 599	0.03	0.9781
gender*type	7.4455	4.5563	7 599	1.63	0.1023
gender*mu_escs	-6.5241	4.4320	7 599	-1.47	0.1411

Only one cross-level interaction is significant, *i.e.* ESCS*type. As the value of 0 is assigned to schools that only propose an academic education and the value of 1 is assigned to schools that provide vocational education, the value of 8.41 for γ_{11} means that the ESCS regression coefficient is equal to 10.62 in academic schools and to 10.62 + 8.41, *i.e.* 19.03 in vocational schools. This result is not surprising as the student population is less diverse in terms of socio-economic background and academic performance in academic schools.

It should be noted that any Level 2 predictors used for testing a cross-level interaction should also be included in the regression of the school intercept. Indeed, the estimation of an interaction requires modelling the main effects.

To demonstrate this, let's suppose the following: Students are distributed according to their immigrant status (native versus immigrant) and according to the type of school (academic versus vocational). Table 15.7 presents the performance average, as well as the percentage of students.

Table 15.7
Average performance and percentage of students by student immigrant status and by type of school

		Native students (0)	Students with an immigrant background (1)
Academic schools (0)	Average performance	650	450
	Percentage of students	25%	25%
Vocational schools (1)	Average performance	500	400
	Percentage of students	25%	25%

Table 15.8 presents the variables for the four groups of students.

Table 15.8
Variables for the four groups of students

	School type	Student immigrant status	Interaction
Academic schools – Native students	0	0	0
Vocational schools – Native students	1	0	0
Academic schools – Students with an immigrant background	0	1	0
Vocational schools – Students with an immigrant background	1	1	1

In a regression model, the reference category will be the native students attending an academic school (0,0,0). Therefore, β_0 will be equal to 650. If the two main dichotomous effects and the interaction are included in the model, it would correspond to three dummies differentiating four categories, as illustrated by Table 15.8. The "school type" main effect will compare the native students in academic schools with the native students in a vocational school. The regression coefficient will therefore be equal to –150. The "student immigrant status" main effect will differentiate student according to their immigrant status within academic schools. The regression coefficient will therefore be equal to –200. Finally, the interaction will compare, for the students with an immigrant background in a vocational school, their expected score based on the main effect and the observed score. The expected score will be 650-150-200, *i.e.* 300. As the average performance is 400, the interaction will therefore be equal to 100.

Let's suppose that the school type is not included as a main effect, but only the immigrant status and the interaction. The reference category (0,0) will be the native students, so β_0 will be equal to 575, the average performance of native students across school types. The "student immigrant status" main effect will be computed by comparing the reference categories with the students with an immigrant background in academic schools and, therefore, the regression coefficient will be equal to –125. Finally, the interaction, as previously, will compare the expected and observed means. The interaction regression coefficient will therefore be equal to the difference between 575 – 125 = 450 and 400. The interaction will be equal to -50.

This fictitious example illustrates that any cross-level interaction should not be tested without the inclusion of the main effect. In terms of multilevel modelling, it means that any variable used for explaining the variability of a Level 1 regression coefficient should also be included in the intercept equation.

Computation of final estimates and their respective standard errors

As described in the previous chapters, the final estimates of a multilevel regression analysis should be also computed for each of five plausible values, with the replicates.

Two SAS® macros have been developed for multilevel regression analyses: one for plausible values as a dependent variable, and the other for non-plausible values as a dependent variable.

Two subroutines are embedded in these two macros:

- The first subroutine deletes any cases with at least one missing value.
- The second subroutine normalises the weight and replicates.

Box 15.13 presents the SAS® syntax for running the macro.

Box 15.13 **SAS® syntax for using the multilevel regression macro (*e.g.* PISA 2006)**

```
%include "c:\pisa\macro\proc_mixed_pv.sas";

%BRR_MIXED_PV(INFILE=temp12,
              REPLI_ROOT=w_fstr,
              PV_ROOT=scie,
              FIXEF=escs vocation st01Q01 type mu_escs pct_im,
              RANEF=img,
              BYVAR=cnt,
              LEVEL2=schoolid,
              OUTSCREEN="c:\ml.out",
              OUTFILE=out);
run;
```

The macro devoted to plausible values has nine arguments. The INFILE, REPLI_ROOT, PV_ROOT, BYVAR and OUTFILE arguments have already been largely described in previous chapters. In the FIXEF and RANEF arguments, the variables considered as fixed and random effects are listed respectively. It should be noted that any cross-level interaction needs to be computed in the data file and the resulting variable should be included in the FIXEF argument.

The Level 2 identification variable (SCHOOLID in the PISA databases) has to be listed in the LEVEL2 argument. Finally, as limited control is provided on the SAS® output, the results are sent to an external file to avoid SAS® stopping due to a lack of memory. The name of the file and its location must be between brackets.

The macro will store the results in four different files with the out-file name specified in the OUTFILE argument, *e.g.* out, followed by the extensions presented in the brackets below:

- a file with the regression coefficients and their standard errors (`out_fixe`),
- a file with the variance estimates and their standard errors (`out_variance`),
- a file with the intraclass correlation and its standard error (`out_intraclass`),
- a file with the percentage of subjects deleted due to missing data (`out_deletion`).

Table 15.9
Comparison of the regression coefficient estimates and their standard errors in Belgium (PISA 2006)

	Using macro		Using PROC MIXED	
Effect	**STAT**	**S.E.**	**STAT**	**S.E.**
ESCS	10.6	1.07	10.3	0.88
Intercept	104.0	17.09	99.5	13.80
st01q01	45.5	1.73	45.8	1.33
img	-24.0	2.61	-23.8	3.07
mu_escs	43.5	2.18	43.9	5.30
pct_im	-52.4	4.91	-53.6	11.41
type	-0.3	1.93	0.6	5.19
vocation	-56.8	2.72	-56.3	2.23

PISA DATA ANALYSIS MANUAL: SAS® SECOND EDITION – ISBN 978-92-64-05624-4 – © OECD 2009

Table 15.9 presents the regression coefficient estimates and their standard errors computed with the macro or just with SAS® PROC MIXED on the first plausible value.

Regression coefficient estimates differ slightly depending on how they are computed. But the most important differences concern the standard errors. Multilevel regression models assume schools were selected according to a simple random procedure. In PISA, explicit and implicit stratification variables improve the efficiency of the sampling design and therefore standard errors are usually smaller than the standard errors obtained from a simple random sample of schools.

From an educational policy perspective, these results show that the percentage of immigrants in a school and the school socio-economic intake do impact the school intercept in Belgium.

As shown in Table 15.10, the macro returns a larger standard error for τ_{11} (variance for the regression coefficient of the variable IMG) because for three out of the five plausible values, the variance estimate was equal to 0.

Table 15.10

Comparison of the variance estimates and their respective standard errors in Belgium (PISA 2006)

	Using macro		Using PROC MIXED	
Cov parm	**STAT**	**S.E.**	**STAT**	**S.E.**
Intercept	738.5	65.39	706.6	73.73
Residual	3619.7	85.63	3599.4	56.99
Img	450.8	237.7	517.5	149.82

These results show the importance of using the 5 plausible values and the 80 replicates to compute the final estimates and their respective standard errors. The use of the replicates is particularly recommended in countries that organised a census of their students (*e.g.* Iceland, Luxembourg).

THREE-LEVEL MODELLING

Three-level regression analyses (*i.e.* Level 1 being the student level, Level 2 the school level and Level 3 the country level) can also be implemented with SAS®. However, even a simple model on the PISA data will run for hours. It is therefore recommended to use specialised software packages such as HLM®. This section shows a simple example of a three-level regression. The detailed example of the preparation for the data files and three-level regression analysis with HLM® applied in the Chapter 5 of the PISA 2006 initial report (OECD, 2007) are presented in Appendix 1.

Three-level modelling requires precaution mainly because removing or adding countries might have a substantial impact on the results at Level 3. For instance, Figure 2.12a of the PISA 2006 initial report (OECD, 2007) presents the relationship between student performance in science and national income. The correlation between these two variables is equal to 0.53. Adding the partner countries would certainly strengthen the relationship.

The following example is based solely on data from OECD countries, since it is less likely that including, or not including, one or a subset of OECD countries will substantially change the results.

First of all, the final student weights need to be normalised. The weight transformation makes: *(i)* the sum of the weight across the countries equal to the number of cases in the databases; and *(ii)* the sum of the weights per country is constant and equal to the total number of cases divided by the number of countries. Box 15.14 presents the SAS® syntax for normalising the weights for a three-level model.

Box 15.14 **SAS® syntax for normalising the weights for a three-level model (*e.g.* PISA 2006)**

```
data temp13;
    set pisa2006.stu;
    if (escs in (.,.N,.I,.M)) then delete;
    oecd=1;
    if (cnt in (“AUS”,“AUT”,“BEL”,“CAN”,“CZE”,“DNK”,“FIN”,“FRA”,“DEU”,“GRC”,
            “HUN”,“ISL”,“IRL”,“ITA”,“JPN”,“KOR”,“LUX”,“MEX”,“NLD”,“NZL”,
            “NOR”,“POL”,“PRT”,“SVK”,“ESP”,“SWE”,“CHE”,“TUR”,“GBR”,“USA”));
    keep cnt schoolid stidstd
        w_fstuwt pv1scie escs oecd;
run;
proc sort data=temp13;
    by cnt;
run;
proc univariate data=temp13 noprint;
    var w_fstuwt;
    by oecd cnt;
    output out=temp14 sum=sum_wgt n=cnt_cases;
run;

data temp15;
    merge temp13 temp14;
    by oecd cnt;
    std_wgt=(w_fstuwt/sum_wgt)*cnt_cases;
run;

proc univariate data=temp14 noprint;
    var cnt_cases;
    by oecd;
    output out=temp16 sum=tot_cases n=n_cnt;
run;
data temp17;
    merge temp15 temp16;
    by oecd;
    final_wgt=std_wgt * ((tot_cases/ n_cnt)/cnt_cases);
    keep cnt schoolid stidstd pv1scie escs final_wgt;
run;
```

The dependent variable is the performance of the student in science. At Level 1, *i.e.* the student level, the only independent variable included in the model is the PISA index of economic, social and cultural status for students. At Level 2, *i.e.* the school level, two variables are included in the model: *(i)* the school socio-economic intake (the school average of the student ESCS index); and *(ii)* the school socio-economic mix (the standard deviation of the student ESCS index).

Table 15.11
Three-level regression analyses

	Model 1	Model 2	Model 3	Model 4
	Fixed parameters			
γ_{000}				501.43
γ_{010}			67.15	68.23
γ_{020}				-7.30
γ_{100}		21.96	19.67	12.34
γ_{110}				0.16
γ_{120}				9.14
	Random parameters			
ε_{ijk}	5 777	5 467	5 469	5 470
U_{0jk}	3 464	2 685	1 244	1 241
U_{00k}	1 037	763	399	397
U_{1jk}		37	37	35
U_{10k}		66	86	89
U_{01k}			1 157	1 138

Table 15.11 shows the results from four different three-level models using HLM®. Model 1 is the empty model, without any independent variables. Model 2 has one independent variable (ESCS) at Level 1 as random slopes at Level 2 and Level 3. Model 3 has one independent variable (MU_ESCS) at Level 2 as random slopes at Level 3 in addition to Model 2. Model 4 has one more independent variable (STD_ESCS) at Level 2 as a fixed slope in addition to Model 3.

The equations for Model 4 are presented below. With only one independent variable at Level 1 and two independent variables at Level 2, a three-level regression model becomes quickly complex, especially when random slopes are modelled.

$$SCIE_{ijk} = \beta_{0jk} + \beta_{1jk}(ESCS)_{ijk} + \varepsilon_{ijk}$$

$$\beta_{0jk} = \beta_{00k} + \beta_{01k}(MU\_ESCS)_{jk} + \beta_{02k}(STD\_ESCS)_{jk} + U_{0jk}$$
$$\beta_{1jk} = \beta_{10k} + \beta_{11k}(MU\_ESCS)_{jk} + \beta_{12k}(STD\_ESCS)_{jk} + U_{1jk}$$

$$\beta_{00k} = \gamma_{000} + U_{00k}$$
$$\beta_{01k} = \gamma_{010} + U_{01k}$$
$$\beta_{02k} = \gamma_{020}$$
$$\beta_{10k} = \gamma_{100} + U_{10k}$$
$$\beta_{11k} = \gamma_{110}$$
$$\beta_{12k} = \gamma_{120}$$

The decomposition of the variance, *i.e.* Model 1, shows that 56% of the variance lies within schools, 34% between schools within countries and only 10% between countries. Model 2 indicates that the national variability of the ESCS regression coefficients (37) is about the same as its international counterpart (66). In other words, the differences in the ESCS regression coefficient between the most equitable schools and the most inequitable schools in a country are similar to the difference between the most equitable countries and the most inequitable countries.

Model 3 demonstrates that the impact of the school socio-economic intake is substantially higher than the impact of the PISA index of economic, social and cultural status of students. Further, the variability of the school social intake regressions (1 157) is higher than the variability of the ESCS regression coefficients. It does mean that countries differs more by the impact of the school socio-economic intake than by the impact of the student socio-economic background.

Finally, Model 4 indicates that the ESCS slope of a particular school is positively correlated with the school socio-economic mix (9.14). A school with a greater socio-economic diversity will have a higher ESCS regression coefficient.

This short example illustrates the potential of three-level regression modelling. However, such models become rapidly complex and their results might be sensitive to which countries are included in the analyses.

LIMITATIONS OF THE MULTILEVEL MODEL IN THE PISA CONTEXT

This section aims to alert PISA data users to some limitations of applying multilevel models in the PISA context.

As PISA draws, per participating school, a random sample of an age population across grades and across classes, it allows the decomposition of the variance into two levels: a between-school variance and a within-school variance. Therefore, the overall variance is expected to be larger with an age-based sample than with a grade sample, unless the age population is attending a single grade, as in Iceland or Japan.

To allow for meaningful international comparisons, these types of indicators require a school definition common to each country. While there are no major differences in definition of what a student is, there are, from one country to another, important differences about what a school is.

International surveys in education are primarily interested in the student sample and therefore one might consider the school sample as a necessary step to draw an efficient sample of students that minimises the cost of testing. In this context, the definition of what a school is, does not present any major issues. However, the increasing importance and popularity of multilevel analyses calls for more attention to the definition issue.

PISA's emphasis in the sampling procedures is on developing a list of units that would guarantee full coverage of the enrolled 15-year-old population and that would additionally give acceptable response rates. Once a "school" was selected, it also had to be practical to sample 35 students from that school to assess them. Thus, the school frame was constructed with issues of student coverage and practical implementation of PISA administration in mind, rather than analytic considerations. Therefore, in the PISA databases, it is possible that the school identification represents different educational institutions that may not be comparable without any restriction. For instance, in some PISA countries, schools were defined as administrative units that may consist of several buildings not necessarily located close to each other. Other countries used the "building" as the school sampling unit and finally, a few countries defined a school as a track within a particular building. It is likely that the larger these aggregates are, the smaller the differences between these aggregates will be and the larger the differences within these aggregates will be. In this context, one would expect to observe high intraclass correlations in these countries and a non-significant within-school regression coefficient for the student socio-economic background (OECD, 2002d).

Besides this problem of an international definition of a school, data users should be aware of the following issues:

- The choice of a school definition in a particular country may be dictated by the availability of the data. Indeed, the national centres have to include a measure of size of the 15-year-old population in the school sample frame (see Chapter 3). This information may be available at the administrative unit level, but not at the "building" level. In federal countries that count several educational systems, the available data might differ from one system to another, so that the concept of a school might differ even within a particular country.
- For practical or operational reasons, the concept of schools might differ between two PISA data collections. For instance, some countries used the administrative units in the PISA 2000 school sample frame and the "building" units in the PISA 2003 school sample frame. Such changes were implemented to increase the school participation rate. These conceptual changes will influence the results of any variance decomposition and might also affect the outcomes of multilevel models. Moving from an administrative definition to a "building" definition will increase the intraclass correlation and should decrease the slope of the within-school regression coefficient. The changes in any trends on variance decomposition or multilevel regressions require careful examination and interpretation.

In summary, multilevel analyses and variance decomposition analyses need to be interpreted in light of the structure of the educational systems and the school definition used in the school sample frame.

CONCLUSION

This chapter described the concept of multilevel analyses and how to perform such models with SAS®. It started with the simplest model, denoted the empty model, and then progressively added complexity by

adding variables. This was followed by a description of the SAS® macro to compute standard errors using five plausible values and replicates.

An example of a three-level regression was then presented.

Finally, in the PISA context, important methodological issues that limit the international comparability of the results were discussed.

Notes

1. While simple linear regression models do not recognize hierarchical structure of data, it is possible to account for some hierarchical aspects of the PISA data in the survey regression models. In many software packages it is also straightforward to correct standard errors in the linear regression by using BRR weights or cluster-robust estimators. These models can adjust for clustering of students within schools and other aspects of survey design.

2. PISA has been using normalised student final weights at the student level for multilevel analyses. But, it is important to note that technical discussion is currently under way regarding the use of separate weights at the different levels.

3. A correlation matrix computed with the pairwise deletion option can however be used as input for a linear regression analysis.

4. This is based on PV1SCIE.

References

Beaton, A.E. (1987), *The NAEP 1983-1984 Technical Report*, Educational Testing Service, Princeton.

Beaton, A.E., *et al.* (1996), *Mathematics Achievement in the Middle School Years, IEA's Third International Mathematics and Science Study*, Boston College, Chestnut Hill, MA.

Bloom, B.S. (1979), *Caractéristiques individuelles et apprentissage scolaire*, Éditions Labor, Brussels.

Bressoux, P. (2008), *Modélisation statistique appliquée aux sciences sociales*, De Boek, Brussels.

Bryk, A.S. and **S.W. Raudenbush** (1992), *Hierarchical Linear Models for Social and Behavioural Research: Applications and Data Analysis Methods*, Sage Publications, Newbury Park, CA.

Buchmann, C. (2000), *Family structure, parental perceptions and child labor in Kenya: What factors determine who is enrolled in school?* a*Soc. Forces,* No. 78, pp. 1349-79.

Cochran, W.G. (1977), *Sampling Techniques*, J. Wiley and Sons, Inc., New York.

Dunn, O.J. (1961), "Multilple Comparisons among Menas", *Journal of the American Statistical Association*, Vol. 56, American Statistical Association, Alexandria, pp. 52-64.

Kish, L. (1995), *Survey Sampling*, J. Wiley and Sons, Inc., New York.

Knighton, T. and **P. Bussière** (2006), "Educational Outcomes at Age 19 Associated with Reading Ability at Age 15", Statistics Canada, Ottawa.

Gonzalez, E. and **A. Kennedy** (2003), *PIRLS 2001 User Guide for the International Database*, Boston College, Chestnut Hill, MA.

Ganzeboom, H.B.G., P.M. De Graaf and **D.J. Treiman** (1992), "A Standard International Socio-economic Index of Occupation Status", *Social Science Research* 21(1), Elsevier Ltd, pp 1-56.

Goldstein, H. (1995), *Multilevel Statistical Models,* 2nd Edition, Edward Arnold, London.

Goldstein, H. (1997), "Methods in School Effectiveness Research", *School Effectiveness and School Improvement* 8, Swets and Zeitlinger, Lisse, Netherlands, pp. 369-395.

Hubin, J.P. (ed.) (2007), *Les indicateurs de l'enseignement*, 2nd Edition, Ministère de la Communauté française, Brussels.

Husen, T. (1967), *International Study of Achievement in Mathematics: A Comparison of Twelve Countries,* Almqvist and Wiksells, Uppsala.

International Labour Organisation (ILO) (1990), *International Standard Classification of Occupations: ISCO-88.* Geneva: International Labour Office.

Lafontaine, D. and **C. Monseur** (forthcoming), "Impact of Test Characteristics on Gender Equity Indicators in the Assessment of Reading Comprehension", *European Educational Research Journal*, Special Issue on PISA and Gender.

Lietz, P. (2006), "A Meta-Analysis of Gender Differences in Reading Achievement at the Secondary Level", *Studies in Educational Evaluation* 32, pp. 317-344.

Monseur, C. and **M. Crahay** (forthcoming), "Composition académique et sociale des établissements, efficacité et inégalités scolaires : une comparaison internationale – Analyse secondaire des données PISA 2006", *Revue française de pédagogie*.

OECD (1998), *Education at a Glance – OECD Indicators,* OECD, Paris.

OECD (1999a), *Measuring Student Knowledge and Skills – A New Framework for Assessment,* OECD, Paris.

OECD (1999b), *Classifying Educational Programmes – Manual for ISCED-97 Implementation in OECD Countries,* OECD, Paris.

OECD (2001), *Knowledge and Skills for Life – First Results from PISA 2000,* OECD, Paris.

OECD (2002a), *Programme for International Student Assessment – Manual for the PISA 2000 Database,* OECD, Paris.

OECD (2002b), *Sample Tasks from the PISA 2000 Assessment – Reading, Mathematical and Scientific Literacy,* OECD, Paris.

OECD (2002c), *Programme for International Student Assessment – PISA 2000 Technical Report,* OECD, Paris.

OECD (2002d), *Reading for Change: Performance and Engagement across Countries – Results from PISA 2000,* OECD, Paris.

OECD (2003a), *Literacy Skills for the World of Tomorrow – Further Results from PISA 2000,* OECD, Paris.

OECD (2003b), *The PISA 2003 Assessment Framework – Mathematics, Reading, Science and Problem Solving Knowledge and Skills,* OECD, Paris.

OECD (2004a), *Learning for Tomorrow's World – First Results from PISA 2003,* OECD, Paris.

OECD (2004b), *Problem Solving for Tomorrow's World – First Measures of Cross-Curricular Competencies from PISA 2003,* OECD, Paris.

OECD (2005a), *PISA 2003 Technical Report,* OECD, Paris.

OECD (2005b), *PISA 2003 Data Analysis Manual,* OECD, Paris.

OECD (2006), *Assessing Scientific, Reading and Mathematical Literacy: A Framework for PISA 2006,* OECD, Paris.

OECD (2007), *PISA 2006: Science Competencies for Tomorrow's World*, OECD, Paris.

OECD (2009), *PISA 2006 Technical Report,* OECD, Paris.

Peaker, G.F. (1975), *An Empirical Study of Education in Twenty-One Countries: A Technical report. International Studies in Evaluation VIII*, Wiley, New York and Almqvist and Wiksell, Stockholm.

Rust, K.F. and **J.N.K. Rao** (1996), "Variance Estimation for Complex Surveys Using Replication Techniques", *Statistical Methods in Medical Research,* Vol. 5, Hodder Arnold, London, pp. 283-310.

Rutter, M., *et al.* (2004), "Gender Differences in Reading Difficulties: Findings from Four Epidemiology Studies", *Journal of the American Medical Association* 291, pp. 2007-2012.

Schulz, W. (2006), *Measuring the socio-economic background of students and its effect on achievement in PISA 2000 and PISA 2003*, Paper presented at the Annual Meetings of the American Educational Research Association (AERA) in San Francisco, 7-11 April.

Wagemaker, H. (1996), *Are Girls Better Readers. Gender Differences in Reading Literacy in 32 Countries*, IEA, The Hague.

Warm, T.A. (1989), "Weighted Likelihood Estimation of Ability in Item Response Theory", *Psychometrika,* Vol. 54(3), Psychometric Society, Williamsburg, VA., pp. 427-450.

Wright, B.D. and **M.H. Stone** (1979), *Best Test Design: Rasch Measurement,* MESA Press, Chicago.

Table of contents

LIST OF BOXES

LIST OF FIGURES

LIST OF TABLES

User's Guide

Preparation of data files

All data files (in text format) and the SAS® control files are available on the PISA website (*www.pisa.oecd.org*).

SAS® users

By running the SAS® control files, the PISA data files are created in the SAS® format. Before starting analysis, assigning the folder in which the data files are saved as a SAS® library.

For example, if the PISA 2000 data files are saved in the folder of "c:\pisa2000\data\", the PISA 2003 data files are in "c:\pisa2003\data\", and the PISA 2006 data files are in "c:\pisa2006\data\", the following commands need to be run to create SAS® libraries:

```
libname PISA2000 "c:\pisa2000\data\";
libname PISA2003 "c:\pisa2003\data\";
libname PISA2006 "c:\pisa2006\data\";
run;
```

SAS® syntax and macros

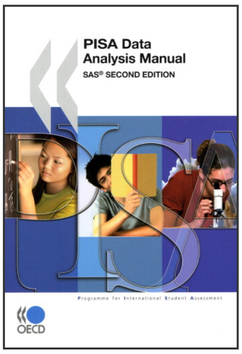
All syntaxes and macros in this manual can be copied from the PISA website (*www.pisa.oecd.org*). The 17 SAS® macros presented in Chapter 17 need to be saved under "c:\pisa\macro\", before staring analysis. Each chapter of the manual contains a complete set of syntaxes, which must be done sequentially, for all of them to run correctly, within the chapter.

Rounding of figures

In the tables and formulas, figures were rounded to a convenient number of decimal places, although calculations were always made with the full number of decimal places.

Country abbreviations used in this manual

AUS	Australia	FRA	France	MEX	Mexico
AUT	Austria	GBR	United Kingdom	NLD	Netherlands
BEL	Belgium	GRC	Greece	NOR	Norway
CAN	Canada	HUN	Hungary	NZL	New Zealand
CHE	Switzerland	IRL	Ireland	POL	Poland
CZE	Czech Republic	ISL	Iceland	PRT	Portugal
DEU	Germany	ITA	Italy	SVK	Slovak Republic
DNK	Denmark	JPN	Japan	SWE	Sweden
ESP	Spain	KOR	Korea	TUR	Turkey
FIN	Finland	LUX	Luxembourg	USA	United States



From:
PISA Data Analysis Manual: SAS, Second Edition

Access the complete publication at:
https://doi.org/10.1787/9789264056251-en

Please cite this chapter as:

OECD (2009), “Multilevel Analyses”, in *PISA Data Analysis Manual: SAS, Second Edition*, OECD Publishing, Paris.

DOI: https://doi.org/10.1787/9789264056251-16-en