



Studying the Relationship between Student Performance and Indices Derived from Contextual Questionnaires

Introduction.....	184
Analyses by quarters.....	184
The concept of relative risk.....	186
▪ Instability of the relative risk.....	187
▪ Computation of the relative risk.....	188
Effect size.....	191
Linear regression and residual analysis.....	193
▪ Independence of errors.....	193
Statistical procedure.....	196
Conclusion.....	197



INTRODUCTION

The PISA initial reports have used the following tools to describe the relationship between student performance and questionnaire indices: (i) dividing the questionnaire indices into quarters and then reporting the mean performance by quarter; (ii) the relative risk; (iii) the effect size; and (iv) the linear regression. This chapter discusses technical issues related to these four tools and presents some SPSS® macros that facilitate their computation.

ANALYSES BY QUARTERS

As described in Chapter 5, the indices derived from questionnaire data were generated with the Rasch Model, and students' estimates were reported with the *weighted likelihood estimates* (WLEs). As previously mentioned, WLE individual estimates constitute a discontinuous variable. Indeed, with the Rasch Model, the raw score is a sufficient statistic.

Table 14.1 presents the distribution of the questionnaire index of **cultural possessions at home** from the PISA 2006 data in Luxembourg. This table shows the discontinuous character of the variable.

Table 14.1
Distribution of the questionnaire index of cultural possession at home
in Luxembourg (PISA 2006)

WLE	%
-1.60	17.52
-1.57	0.13
-1.49	0.10
-0.93	0.13
-0.84	0.02
-0.58	24.53
-0.50	0.27
-0.23	0.02
-0.20	0.15
-0.16	0.38
0.13	0.14
0.21	23.74
0.50	0.07
0.58	0.82
1.11	0.09
1.19	0.05
1.23	31.84

The cultural-possession-at-home scale consists of three dichotomous items. There are therefore four possible scores, ranging from 0 to 3. Thus, 98% of the students ($17.52 + 24.53 + 23.74 + 31.84$) are distributed among four WLEs. All the other WLEs, with negligible percentages, represent response patterns with at least one missing value.

How can this distribution be divided into four quarters, especially given that approximately 32% of the students have the highest WLE?

There are two ways of recoding the data:

1. smaller versus equal and greater;
2. smaller and equal versus greater.

Depending on the procedure adopted, the percentages of students in the bottom quarter, second quarter, third quarter, and top quarter will vary. Further, neither of these two procedures generate four equal quarters of 25% students each. Since the percentages of students in each quarter can vary among countries, no international comparisons can be made.



It is therefore necessary to distribute the students with a WLE equal to one of the three cutpoint percentiles into the two respective adjacent quarters. For instance, if 31.84% of the students get a score equal to percentile 75, it is then necessary to sample 6.84% of these students who will be allocated to the third quarter. The remaining 25% will be allocated to the fourth quarter.

This random subsampling process is implemented by adding a small random variable to the questionnaire index. That random noise will in a sense transform this discontinuous variable into a pseudo-continuous variable. The three new percentiles on this pseudo-continuous variable will divide the index variable into quarters, each including exactly 25% of the students.

This random allocation of some parts of the population to one of the four quarters adds an error component to the standard error. Indeed, in the example, the composition of the 6.84% of the students allocated to the third quarter is likely to differ between two runs of the procedure.

To account for this new error component, the statistical approach adopted for the analyses of plausible values can be implemented. It will therefore consist of:

- computing a set of five plausible quarters for each student;
- computing the required statistic and its respective sampling variance by using the final and 80 replicate weights, per plausible quarter;
- averaging the five estimates and their respective sampling variances;
- computing the imputation variance;
- combining the sampling variance and the imputation variance to obtain the final error variance.

As the dependent variable is exclusively, or nearly exclusively, a set of plausible values, the procedure described in Chapter 8 will be used, except that each plausible value will be analysed with a different plausible quarter.

Box 14.1 SPSS® syntax for the quarter analysis (e.g. PISA 2006)

```
GET FILE="C:\PISA\2006\DATA\INT_STU06_DEC07.SAV" .
SELECT IF CNT="LUX" .
SAVE OUTFILE="C:\TEMP\LUX.SAV" .

INSERT FILE="C:\PISA\MACRO\MCR_SE_PV_WLEQRT.SPS" .

OMSEND.

SET MPRINT=YES.
PVWLEQRT PV=PV1SCIE PV2SCIE PV3SCIE PV4SCIE PV5SCIE /
WLE=CULTPOSS/
STAT=MEAN/
INFILE=C:\TEMP\LUX.SAV/
LIMIT_CRITERIA=100 10 5 1/                /* default=0 */
CNT=CNT/                                    /* is default */
GRP=NOGRP/                                  /* is default */
PSU=SCHOOLID/                               /* is default */
WGT=W_FSTUWT/                               /* is default */
RWGT=W_FSTR/                                /* is default */
NREP=80/                                    /* is default */
CONS=0.05/.                                 /* is default */

/* POSSIBLE STATISTICS ARE: MEAN, SD, MIN, MAX, MEDIAN */.
```



A SPSS® macro was developed to facilitate these computations, as largely used in the three PISA initial reports. Box 14.1 presents the SPSS® syntax and Table 14.2 the output data file.

The arguments of the quarter SPSS® macro have been extensively described in other chapters. The output data file contains: (i) the mean of the contextual index (the independent variable) per quartile and its corresponding standard error; and (ii) the average performance in science (the dependent variable) and its standard errors.

Table 14.2
Output data file from Box 14.1

CNT	Q_CULTPOSS	INDEX_STAT	INDEX_SE	PV_STAT	PV_SE	FLAG_STD	FLAG_SCH	FLAG_PCT
LUX	1	-1.31	0.01	451.40	3.09	0	0	0
LUX	2	-0.35	0.01	471.64	3.37	0	0	0
LUX	3	0.51	0.01	494.24	3.00	0	0	0
LUX	4	1.23	0.00	531.71	3.52	0	0	0

THE CONCEPT OF RELATIVE RISK

The notion of relative risk is a measure of association between an antecedent factor and an outcome factor (Cornfield, 1951). The relative risk is simply the ratio of two risks, *i.e.* the risk of observing the outcome when the antecedent is present, and the risk of observing the outcome when the antecedent is not present.

Table 14.3 presents the notation that will be used.

$P_{..}$ is equal to $\frac{n_{..}}{n_{..}}$, with $n_{..}$ the total number of students and $P_{..}$ is therefore equal to 1; $P_{i.}$, $P_{.j}$ respectively represent the marginal probabilities for each row and for each column. The marginal probabilities are equal to the marginal frequencies divided by the total number of students. Finally, the P_{ij} values represent the probabilities for each cell and are equal to the number of observations in a particular cell divided by the total number of observations.

Table 14.3
Labels used in a two-way table

		Outcome measure		
		Yes	No	Total
Antecedent measure	Yes	P_{11}	P_{12}	$P_{1.}$
	No	P_{21}	P_{22}	$P_{2.}$
	Total	$P_{.1}$	$P_{.2}$	$P_{..}$

In this chapter, the conventions for the two-way table are:

- The rows represent the antecedent factor with:
 - the first row having the antecedent; and
 - the second row not having the antecedent.
- The columns represent the outcome with:
 - the first column having the outcome; and
 - the second column not having the outcome.



With these conditions, the relative risk is equal to:

$$RR = \frac{(p_{11} / p_{1.})}{(p_{21} / p_{2.})}$$

Let's suppose that a psychologist wants to analyse the risk of a student repeating a grade if the parents recently divorced. The psychologist draws a simple random sample of students in grade 10. In this example, the outcome variable is present if the child is repeating grade 10 and the antecedent factor is considered present if the student's parents divorced in the past two years. The results are found in Table 14.4 and Table 14.5.

Table 14.4
Distribution of 100 students by parents' marital status and grade repetition

	Repeat the grade	Does not repeat the grade	Total
Parents divorced	10	10	20
Parents not divorced	5	75	80
Total	15	85	100

Table 14.5
Probabilities by parents' marital status and grade repetition

	Repeat the grade	Does not repeat the grade	Total
Parents divorced	0.10	0.10	0.20
Parents not divorced	0.05	0.75	0.80
Total	0.15	0.85	1.00

The relative risk is therefore equal to:

$$RR = \frac{(p_{11} / p_{1.})}{(p_{21} / p_{2.})} = \frac{(0.10/0.20)}{(0.05/0.80)} = \frac{0.5}{0.0625} = 8$$

This means that the probability of repeating grade 10 is eight times greater if the parents recently divorced than if they had stayed together.

Instability of the relative risk

The relative risk was developed for dichotomous variables. More and more often, this coefficient is extended and is used with continuous variables. However, to apply the coefficient to continuous variables, a cutpoint needs to be set for each variable and the continuous variables need to be dichotomised.

It is important to recognise that when applied to dichotomised variables, the computed values of the relative risk will depend on the value of the chosen cutpoint.

Table 14.6
Relative risk for different cutpoints

Percentile	Relative risk
10	2.64
15	2.32
20	1.90
25	1.73
30	1.63

To demonstrate the influence of the cutpoint on the relative risk, two random variables were generated with a correlation of 0.30. These two variables were then transformed into dichotomous variables by using the 10th, 15th, 20th, 25th and 30th percentiles respectively as cutpoints. Table 14.6 presents the relative risk for a range of choices for the cutpoints.



Table 14.6 shows that the relative risk coefficient is dependent on the setting of the cutpoints; thus, the values should be interpreted in light of this.

Such a comparison of the relative risks was computed for the PISA 2000 data to identify the changes depending on the cutpoint location. The antecedent factor was the mother's educational level and the outcome variable was student performance in reading. Low reading performance was successively defined within countries as being below the 10th, 15th, 20th, 25th, 30th and 35th percentiles.

In PISA 2000, the relative risks for these different cutpoints are on average (across OECD countries) equal to 2.20, 1.92, 1.75, 1.62, 1.53, and 1.46, respectively. In PISA, it was decided to use the 25th percentile as the cutpoint for continuous variables when calculating the relative risk.

Computation of the relative risk

Depending on the variables involved in the computation of the relative risk, the procedure might differ. Indeed, the relative risk concept requires as input two dichotomous variables, such as gender.

However, most of the variables in the PISA databases are not dichotomous; they are categorical or continuous variables.

The recoding of a categorical into a dichotomous variable does not raise any specific issues. From a theoretical point of view, the purpose of the comparison needs to be decided; the recoding will follow. For instance, in PISA 2003, the education levels of the parents are reported by using the ISCED classification (OECD, 1999b). If the comparison is based on the distinction between tertiary versus non-tertiary education, then the categorical variable can be recoded into a dichotomous variable.

Numerical variables also have to be recoded into dichotomous variables. As stated earlier, the OECD has decided to divide numerical variables based on the 25th percentile.

If plausible values are involved as outcome measures, after the recoding of the five estimates of the student performance into dichotomous variables, five relative risks will be computed and then combined.

In the PISA databases however, most numerical variables are discontinuous variables. To ensure that the 25th percentile will divide the variables into two categories that will include 25% and 75% respectively, a random component has to be added to the initial variable, as described in the previous section on analyses per quarter. Five relative risk estimates are computed and then combined.

Box 14.2 presents the SPSS® syntax for computing the increased likelihood of the students in the bottom quarter of HISEI (international socio-economic index of occupational status) scoring the bottom quarter of the science performance distribution, with the PISA 2006 data in France. As HISEI is a discontinuous variable with a limited number of values, it is necessary to add a random component. This example therefore has five antecedent variables and five outcome variables.

The first macro devoted to the computation of the relative risk requires five dummy variables as antecedents and five dummy variables as outcomes. Value 1 will be assigned if the risk is present; otherwise the assigned value will be 0. Value 1 will also be assigned if the outcome is present; otherwise it will be 0. It is of prime importance to respect these application conditions. Inverting the values will not stop the macro running, but it will change the meaning of the results.

Table 14.7 shows that a student in the bottom quarter of the international socio-economic index of occupational status has 2.38 times more chance of appearing in the bottom quarter of the science performance distribution.



Box 14.2 SPSS® syntax for computing the relative risk with five antecedent variables and five outcome variables (e.g. PISA 2006)

```

GET FILE="C:\PISA\2006\DATA\INT_STU06_DEC07.SAV".
SELECT IF CNT="FRA".
RECODE ST04Q01 (1=0) (2=1) (ELSE=SYSMIS) INTO GENDER.
IF (INTSCIE<=0) INT=1.
IF (INTSCIE>0) INT=0.

SET SEED=1.
COMPUTE SES1=HISEI+RV.NORMAL(0,0.01).
SET SEED=2.
COMPUTE SES2=HISEI+RV.NORMAL(0,0.01).
SET SEED=3.
COMPUTE SES3=HISEI+RV.NORMAL(0,0.01).
SET SEED=4.
COMPUTE SES4=HISEI+RV.NORMAL(0,0.01).
SET SEED=5.
COMPUTE SES5=HISEI+RV.NORMAL(0,0.01).
EXE.

WEIGHT BY W_FSTUWT.
FRE PV1SCIE PV2SCIE PV3SCIE PV4SCIE PV5SCIE SES1 SES2 SES3 SES4 SES5 /
FORMAT=NOTABLE /PERCENTILES=25.

COMPUTE PV25_1=(PV1SCIE<=421.6089).
COMPUTE PV25_2=(PV2SCIE<=425.1523).
COMPUTE PV25_3=(PV3SCIE<=424.2198).
COMPUTE PV25_4=(PV4SCIE<=422.5414).
COMPUTE PV25_5=(PV5SCIE<=426.0848).
COMPUTE SES25_1=(SES1<=34.0114).
COMPUTE SES25_2=(SES2<=34.0104).
COMPUTE SES25_3=(SES3<=34.0095).
COMPUTE SES25_4=(SES4<=34.0127).
COMPUTE SES25_5=(SES5<=34.0089).
EXE.

FRE PV25_1 PV25_2 PV25_3 PV25_4 PV25_5 SES25_1 SES25_2 SES25_3 SES25_4 SES25_5.
FORMATS GENDER INT PV25_1 PV25_2 PV25_3 PV25_4 PV25_5 SES25_1 SES25_2 SES25_3
SES25_4 SES25_5 (F1.0).
WEIGHT OFF.
SAVE OUTFILE="C:\TEMP\FRA.SAV".

*** RELATIVE RISK BETWEEN TWO SETS OF 5 DICHOTOMOUS VARIABLES (AT LEAST ONE SET
IS BASED ON PLAUSIBLE VALUES) ***.

INSERT FILE="C:\PISA\MACRO\MCR_SE_RR_PV.SPS".

SET MPRINT=YES.
RelRisk_PV OUTCOME=PV25_1 PV25_2 PV25_3 PV25_4 PV25_5/
ANTECEDENT=SES25_1 SES25_2 SES25_3 SES25_4 SES25_5/
INFILE=C:\TEMP\FRA.sav/
GRP=NOGRP/ /* default is NOGRP */
WGT=W_FSTUWT/ /* is default */
RWGT=W_FSTR/ /* is default */
CONS=0.05/ /* is default */
NREP=80/. /* is default */

```

Table 14.7
Output data file from Box 14.2

CNT	STAT	SE
FRA	2.42	0.18

A second macro presented in Box 14.3 has been developed for analyses that involve only one antecedent variable and one outcome variable.



Box 14.3 SPSS® syntax for computing the relative risk with one antecedent variable and one outcome variable (e.g. PISA 2006)

```

INSERT FILE="C:\PISA\MACRO\MCR_SE_RR.SPS" .
SET MPRINT=YES.
RelRisk OUTCOME=INT/
      ANTECEDENT=GENDER/
      INFILE=C:\TEMP\FRA.SAV/
      GRP=CNT/                          /* default is NOGRP */
      WGT=W FSTUWT/                      /* is default */
      RWGT=W FSTR/                       /* is default */
      CONS=0.05/                          /* is default */
      NREP=80/.                          /* is default */

```

It should be noted that the variables listed in the antecedent and outcome arguments should be assigned a F1.0 format.

No macro has been developed for analyses that involve one antecedent variable and five outcome variables. However, Box 14.4 presents the SPSS® syntax for computing the relative risk in that case. It consists of running the macro for relative risk without plausible values five times and then, as usual, combining the results.

Box 14.4 SPSS® syntax for computing the relative risk with one antecedent variable and five outcome variables (e.g. PISA 2006)

```

GET FILE="C:\TEMP\FRA.SAV" .
DO REPEAT A=PV1SCIE PV2SCIE PV3SCIE PV4SCIE PV5SCIE /
      B=LOW1 TO LOW5.
IF (A<=409.5) B=1.
IF (A>409.5) B=0.
END REPEAT.
FORMATS GENDER LOW1 LOW2 LOW3 LOW4 LOW5 (F1.0) .
SAVE OUTFILE="C:\TEMP\FRA.SAV" .

RelRisk OUTCOME=LOW1/ ANTECEDENT=GENDER/ INFILE=C:\TEMP\FRA.SAV/.
COMPUTE PV=1.
SAVE OUTFILE="C:\TEMP\RESULTS1.SAV" .
RelRisk OUTCOME=LOW2/ ANTECEDENT=GENDER/ INFILE=C:\TEMP\FRA.SAV/.
COMPUTE PV=2.
SAVE OUTFILE="C:\TEMP\RESULTS2.SAV" .
RelRisk OUTCOME=LOW3/ ANTECEDENT=GENDER/ INFILE=C:\TEMP\FRA.SAV/.
COMPUTE PV=3.
SAVE OUTFILE="C:\TEMP\RESULTS3.SAV" .
RelRisk OUTCOME=LOW4/ ANTECEDENT=GENDER/ INFILE=C:\TEMP\FRA.SAV/.
COMPUTE PV=4.
SAVE OUTFILE="C:\TEMP\RESULTS4.SAV" .
RelRisk OUTCOME=LOW5/ ANTECEDENT=GENDER/ INFILE=C:\TEMP\FRA.SAV/.
COMPUTE PV=5.
SAVE OUTFILE="C:\TEMP\RESULTS5.SAV" .

GET FILE="C:\TEMP\RESULTS1.SAV" .
ADD FILES /FILE=*
      /FILE="C:\TEMP\RESULTS2.SAV"
      /FILE="C:\TEMP\RESULTS3.SAV"
      /FILE="C:\TEMP\RESULTS4.SAV"
      /FILE="C:\TEMP\RESULTS5.SAV" .
EXTENSION.
COMPUTE VAR=SE*SE.
AGGREGATE
  /OUTFILE=* MODE=ADDVARIABLES
  /BREAK=
  /STAT_MEAN=MEAN(STAT) .
COMPUTE MESVAR=((STAT-STAT_MEAN)**2)/4.
AGGREGATE
  /OUTFILE="C:\TEMP\RESULTS.SAV"
  /BREAK=
  /STAT=MEAN(STAT)
  /SAMPVAR=MEAN(VAR)
  /MESVAR=SUM(MESVAR) .
GET FILE="C:\TEMP\RESULTS.SAV" .
COMPUTE SE=(SAMPVAR+(1.2*MESVAR))**.5.
EXECUTE.

```




EFFECT SIZE

An effect size is a measure of the strength of the relationship between two variables. PISA requires sampling a substantial number of students from each participating country. As standard errors are inversely proportional to the number of observations, small differences will be statistically different from 0. It is therefore recommended to analyse the strength of a relationship that is statistically significant. In other words, the effect size helps researchers decide whether a statistically significant difference is of practical concern.

The term effect size is commonly used to refer to standardised differences. Standardising a difference is useful when the metric has no intrinsic meaning. With variables that have an intrinsic meaning, it is preferable to use non-standardised differences. For instance, the differences between male and female averages of height or weight are more meaningful if they are expressed in metres or kilos than if they are expressed in standardised differences.

Mathematically, the effect size is equal to:

$$\frac{\hat{\mu}_1 - \hat{\mu}_2}{\sqrt{\frac{\sigma_1^2 + \sigma_2^2}{2}}}$$

$\hat{\mu}_1$ and $\hat{\mu}_2$ respectively represent the mean estimates for groups 1 and 2 and σ_1^2 , σ_2^2 their variance.

Effect sizes are particularly interesting in PISA. Firstly, as differences are also compared across countries, using the difference or using the effect size might change the interpretation of the results. Differences will mainly affect countries with large or small standard deviations. For instance, the mean and the standard deviation of the student performance in reading for males and for females are presented in Table 14.8.

Table 14.8

Mean and standard deviation for the student performance in reading by gender, gender difference and effect size (PISA 2006)

	Performance in reading							
	Mean		Standard deviation		Difference in means between females and males	Effect size	Rank of the difference	Rank of the effect size
	Females	Males	Females	Males				
AUS	531.8	494.9	86.9	96.5	36.9	0.40	14	16
AUT	512.9	468.3	102.8	108.8	44.6	0.42	24	22
BEL	521.7	482.0	102.1	113.5	39.7	0.37	17	12
CAN	543.0	511.1	90.7	98.9	31.9	0.34	7	6
CHE	515.2	484.4	91.2	94.3	30.8	0.33	5	5
CZE	508.6	462.8	107.5	110.0	45.8	0.42	25	21
DEU	516.6	474.6	105.6	113.9	42.0	0.38	22	14
DNK	509.3	479.5	86.0	90.0	29.8	0.34	3	8
ESP	478.7	443.3	82.4	91.4	35.4	0.41	13	19
FIN	572.0	521.4	73.4	80.9	50.6	0.66	28	29
FRA	504.6	469.8	97.1	107.9	34.9	0.34	11	9
GBR	509.5	480.4	95.3	106.3	29.2	0.29	2	2
GRC	488.1	431.6	87.9	108.3	56.6	0.57	29	28
HUN	503.0	463.4	87.1	96.8	39.6	0.43	16	24
IRL	534.0	500.2	87.0	94.7	33.8	0.37	10	13
ISL	508.9	460.4	87.6	99.8	48.5	0.52	27	27
ITA	489.0	447.7	100.6	112.8	41.3	0.39	20	15
JPN	513.3	482.7	95.0	107.1	30.6	0.30	4	3
KOR	573.8	538.8	82.4	90.4	35.0	0.41	12	18
LUX	495.4	463.7	95.2	102.5	31.7	0.32	6	4
MEX	426.7	393.1	91.7	96.8	33.6	0.36	9	10
NLD	519.0	494.9	93.1	98.4	24.2	0.25	1	1
NOR	508.0	462.1	95.5	108.8	45.9	0.45	26	25
NZL	539.1	501.7	99.2	107.9	37.4	0.36	15	11
POL	527.6	487.4	92.6	103.5	40.1	0.41	18	20
PRT	488.2	455.3	94.1	100.8	32.8	0.34	8	7
SVK	487.8	446.1	98.5	107.0	41.7	0.40	21	17
SWE	528.1	487.6	92.6	99.3	40.5	0.42	19	23
TUR	471.0	427.3	84.6	94.8	43.7	0.49	23	26



The differences and the effect sizes have been sorted and ranked. In most cases, the two rankings presented in the two last columns of Table 14.8 are equal or close. However, for some countries, they differ substantially. For instance, Germany is ranked 22 in the difference, and 14 in the effect size; Hungary ranks 16 and 24, respectively. The German and Hungarian researchers and policy makers would certainly interpret the results differently if their analyses were based on the absolute difference or on the effect size.

An effect size also allows a comparison of differences across measures that differ in their metric. For example, it is possible to compare effect sizes between the PISA indices and the PISA test scores, as for example, gender differences in performance in science compared to gender differences in several indices.

Two SPSS® macros have been developed for the computation of the effect size, depending on whether the independent variables consist of plausible values or not. Box 14.5 presents the SPSS® syntax for running effect size analyses.

Box 14.5 SPSS® syntax for computing effect size (e.g. PISA 2006)

```
GET FILE="C:\PISA\2006\DATA\INT_STU06_DEC07.SAV" .
SELECT IF CNT="AUT" .
SAVE OUTFILE="C:\TEMP\AUT.SAV" .
INSERT FILE="C:\PISA\MACRO\MCR_SE_EFFECT_PV.SPS" .
OMSEND.
SET MPRINT=YES.
EFF_PV DEP=PV1SCIE PV2SCIE PV3SCIE PV4SCIE PV5SCIE /
  COMPARE=ST04Q01 /
  CATEG= 1 2/
  INFILE=C:\TEMP\AUT.SAV/
  WITHIN=CNT/ /*default=NOWITHIN*/
  WGT=W_FSTUWT/ /*default=W_FSTUWT*/
  RWGT=W_FSTR/ /*default=W_FSTR */
  CONS=0.05/ /*default=0.05 */
  NREP=80/. /*default=80 */

INSERT FILE="C:\PISA\MACRO\MCR_SE_EFFECT.SPS" .
SET MPRINT=YES.
EFFnoPV DEP=INSTSCIE/
  COMPARE=ST04Q01/
  CATEG= 1 2/
  INFILE=C:\TEMP\AUT.SAV/
  WITHIN=CNT/ /*default=NOWITHIN*/
  WGT=W_FSTUWT/ /*default=W_FSTUWT*/
  RWGT=W_FSTR/ /*default=W_FSTR */
  CONS=0.05/ /*default=0.05 */
  NREP=80/. /*default=80 */
```

The dependent variable is listed in the **DEP** argument and the independent variable is listed in the **COMPARE** argument. The two values that define the two subgroups are listed in the **CATEG** argument. The order of two values is important: with **1 2**, the effect size will be equal to:

$$\frac{\hat{\mu}_1 - \hat{\mu}_2}{\sqrt{\frac{\sigma_1^2 + \sigma_2^2}{2}}}$$

With **2 1**, the effect size will be equal to:

$$\frac{\hat{\mu}_2 - \hat{\mu}_1}{\sqrt{\frac{\sigma_1^2 + \sigma_2^2}{2}}}$$



Table 14.9 and Table 14.10 present the structure of the output data files.

Table 14.9
Output data file from the first model in Box 14.5

CNT	ST04Q01	STAT	SESTAT
AUT	1-2	-0.08	0.05
TOTAL	1-2	-0.08	0.05

Table 14.10
Output data file from the second model in Box 14.5

CNT	EFF	STAT	SESTAT
AUT	ST04Q01.1-2	-0.21	0.04
TOTAL	ST04Q01.1-2	-0.21	0.04

LINEAR REGRESSION AND RESIDUAL ANALYSIS

This section is devoted to linear regression analyses. As it will be demonstrated, such models can be used, even if there is some dependency between the errors within schools. Further, analyses of regression residuals might also provide researchers an opportunity to investigate composition effects outside complex multilevel modelling, for instance.

Independence of errors

As expressed by P. Bressoux (2008), “*qui se ressemble s’assemble mais aussi qui s’assemble tend à se ressembler*” (all that appear similar, group together; but also, all that group together tend to appear similar).

Selected students that attend the same school cannot be considered as independent observations as they are usually more similar to each other than students that attend different schools. This assumption of independence of errors can also be translated as a requirement of the absence of intraclass correlation. The PISA initial and thematic reports have reported the school variances and the intraclass correlation coefficients. In some countries such as Austria, Belgium or Germany, more than 50% of the variance is accounted for by the schools. The assumption of independence between errors cannot be maintained.

The violation of this assumption does not bias the regression coefficient estimates, but underestimates its standard error, which leads to an increase of type I error (Bressoux, 2008, p. 108). Fortunately, standard errors in PISA are estimated by replication techniques and are therefore unbiased.

The PISA initial and thematic reports have extensively used the linear regression, mainly for reporting the change in the student performance score per unit of indices derived from contextual questionnaires. The detailed description of conducting linear regression analysis with using replicates and plausible values is presented in Chapters 7 and 8.

This section will describe an alternative use of the linear regression. It consists of analysing the residual according to a particular criterion. An illustration from Monseur and Crahay (forthcoming) is also provided to show the potential of such analyses.

The underlying hypothesis of the illustration relates to the impact of social segregation in schools on student performance. Typically, this hypothesis should be tested with a multilevel regression analysis. However, such models are quite complex and their results are not always easy to explain.



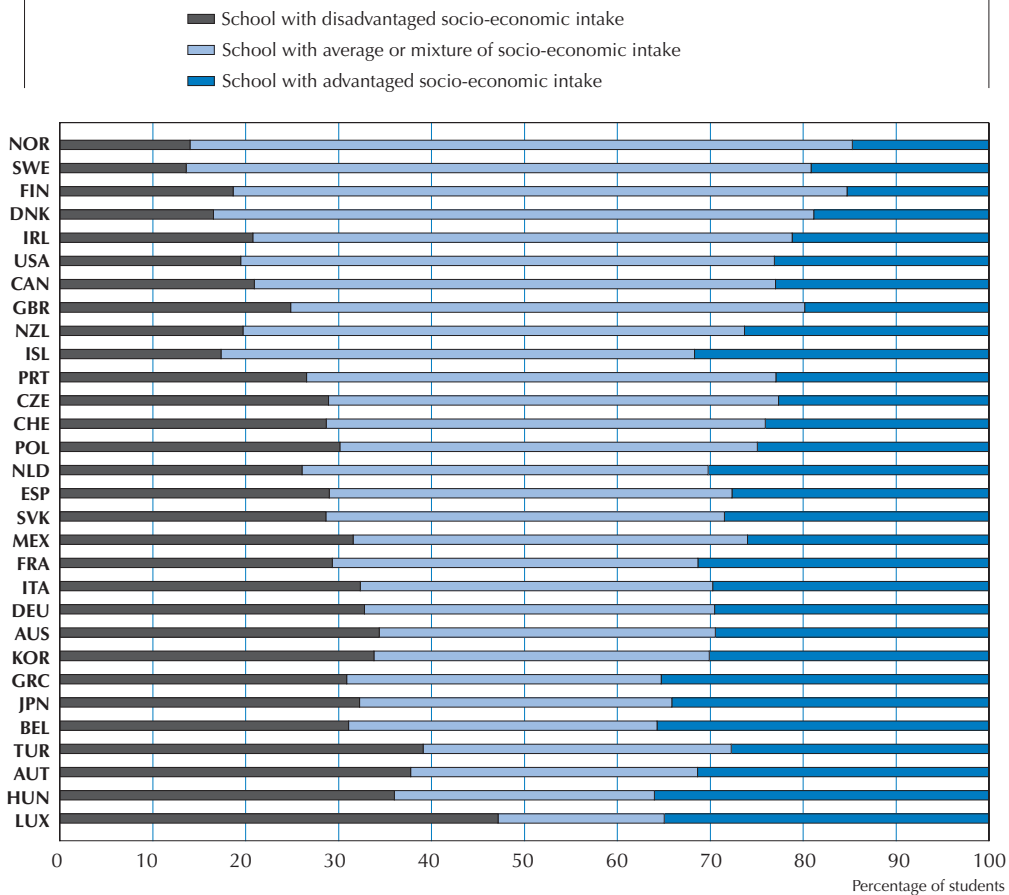
The first step implemented by Monseur and Crahay (forthcoming) was to allocate schools to one of the three following groups: (i) schools mainly attended by students with disadvantaged socio-economic backgrounds; (ii) schools mainly attended by students with average socio-economic backgrounds, or students with a mixture of disadvantaged and advantaged socio-economic backgrounds; and (iii) schools mainly attended by students with advantaged socio-economic backgrounds.

As the within-school samples are simple random samples, the standard error of the school average of the student socio-economic background can be estimated only with the student final weights without replicates. Then, the difference of a school average and the country average is statistically tested. If the null hypothesis is accepted, the school is allocated to the second group, *i.e.* the average or mixed schools; otherwise, the school is allocated to the first or to the third group, depending on whether the difference is negative or positive.

Figure 14.1 presents the percentage of students by the three school groups in PISA 2003.

Northern European countries, and to a lesser extent English-speaking countries, tend to have a large percentage of mixed schools. Highly tracked educational systems, such as Belgium or Germany, present small percentages of the average or mixed schools.

Figure 14.1
Percentage of schools by three school groups (PISA 2003)





The mathematics performance of the students was predicted by the PISA index of economic, social and cultural status (ESCS) of the student. The residuals were then saved and their means were computed by school categories and by the national quartiles of ESCS. Table 14.11 presents the average residual for the students in the bottom quarter of ESCS and for the students in the top quarter of ESCS, by the three school groups.

Table 14.11
Mean of the residuals in mathematics performance for the bottom and top quarters of the PISA index of economic, social and cultural status, by school group (PISA 2003)

	<i>(i) School with disadvantaged socio-economic intake</i>				<i>(ii) School with average or mixture of socio-economic intake</i>				<i>(iii) School with advantaged socio-economic intake</i>			
	Students in the bottom quarter of ESCS		Students in the top quarter of ESCS		Students in the bottom quarter of ESCS		Students in the top quarter of ESCS		Students in the bottom quarter of ESCS		Students in the top quarter of ESCS	
	Mean	S.E.	Mean	S.E.	Mean	S.E.	Mean	S.E.	Mean	S.E.	Mean	S.E.
AUS	-12.7	(5.4)	-32.5	(7.9)	10.3	(6.4)	-13.3	(4.4)	47.7	(11.6)	18.6	(4.1)
AUT	-20.1	(5.3)	-79.3	(8.6)	25.5	(7.7)	-25.5	(8.0)	87.3	(11.9)	20.4	(4.9)
BEL	-30.3	(4.4)	-71.9	(8.6)	22.2	(4.7)	-18.9	(5.0)	97.3	(5.7)	15.3	(3.5)
CAN	-8.5	(3.6)	-19.2	(5.9)	4.0	(2.7)	-9.5	(3.0)	23.9	(7.7)	17.1	(4.2)
CHE	-13.9	(6.3)	-43.3	(11.6)	-0.3	(4.7)	-24.6	(5.1)	72.1	(17.1)	21.9	(5.2)
CZE	-23.2	(5.5)	-71.8	(11.8)	6.6	(4.0)	-32.0	(4.4)	78.6	(15.1)	29.4	(5.4)
DEU	-30.2	(5.2)	-85.1	(9.2)	25.8	(5.8)	-32.0	(7.3)	109.3	(9.8)	24.9	(3.4)
DNK	-15.6	(7.6)	-18.5	(14.7)	2.5	(4.0)	-1.6	(3.8)	19.8	(13.7)	9.2	(5.7)
ESP	-7.2	(5.4)	-31.9	(9.3)	9.8	(4.5)	-9.8	(4.2)	43.0	(8.9)	13.8	(3.7)
FIN	1.6	(4.2)	-7.5	(7.5)	-0.3	(3.2)	-0.1	(3.8)	-15.8	(12.7)	1.6	(5.9)
FRA	-29.7	(5.2)	-70.8	(15.9)	27.4	(6.6)	-8.5	(7.0)	69.6	(7.7)	9.7	(5.2)
GBR	-12.4	(4.0)	-30.0	(8.0)	11.0	(4.3)	-10.9	(3.8)	47.5	(11.3)	25.9	(5.3)
GRC	-16.3	(5.5)	-63.8	(11.4)	21.3	(6.9)	-10.4	(5.6)	57.6	(9.4)	14.0	(4.8)
HUN	-21.2	(4.8)	-84.6	(10.2)	36.7	(5.8)	-37.0	(8.2)	76.8	(11.8)	14.4	(4.3)
IRL	-22.3	(5.7)	-38.8	(12.1)	10.0	(3.8)	-9.2	(4.0)	25.9	(11.3)	9.2	(5.7)
ISL	2.5	(4.9)	-0.2	(15.9)	-0.4	(3.8)	4.9	(4.5)	-1.1	(9.6)	0.9	(3.7)
ITA	-27.5	(5.6)	-71.9	(7.4)	24.2	(7.3)	-22.7	(6.9)	71.1	(12.4)	12.5	(5.9)
JPN	-27.7	(7.6)	-75.4	(11.0)	19.1	(7.9)	-41.1	(6.8)	78.3	(11.0)	26.6	(8.7)
KOR	-25.0	(5.5)	-79.8	(10.5)	41.8	(6.2)	-9.6	(5.1)	59.3	(9.5)	20.1	(8.4)
LUX	-9.4	(2.7)	-51.2	(6.3)	14.7	(6.6)	0.0	(8.2)	73.9	(8.1)	25.3	(3.0)
MEX	-13.9	(4.7)	-56.8	(9.3)	20.8	(5.2)	-30.5	(5.2)	71.7	(7.1)	22.4	(4.5)
NLD	-30.3	(6.1)	-87.7	(8.4)	15.1	(7.8)	-28.3	(6.2)	105.3	(7.5)	34.0	(5.0)
NOR	-0.5	(7.2)	-14.2	(15.3)	-1.4	(3.3)	-2.1	(4.3)	-11.4	(15.8)	5.6	(5.1)
NZL	-19.1	(6.3)	-45.0	(12.8)	7.7	(4.0)	-2.3	(4.1)	30.9	(10.6)	19.2	(3.9)
POL	-8.3	(5.0)	-19.5	(7.9)	0.7	(4.8)	-4.2	(4.7)	23.4	(14.2)	4.8	(3.8)
PRT	-22.3	(5.5)	-20.1	(11.3)	20.0	(6.4)	-5.4	(5.8)	58.6	(12.8)	15.1	(3.8)
SVK	-28.6	(4.9)	-75.2	(12.2)	10.3	(5.3)	-27.8	(6.3)	59.0	(9.1)	18.2	(5.2)
SWE	-12.4	(8.6)	-24.4	(12.5)	6.4	(3.5)	-2.4	(4.0)	15.7	(10.3)	9.5	(7.0)
TUR	0.6	(5.0)	-79.3	(10.2)	30.0	(7.0)	-38.7	(5.4)	94.4	(13.6)	38.1	(13.2)
USA	-22.3	(6.0)	-53.5	(12.5)	13.2	(3.6)	-1.6	(3.6)	27.3	(10.5)	15.3	(4.2)

In all countries with significant differences:

- Students attending schools with disadvantaged socio-economic intake perform on average lower than what would be predicted based on their socio-economic background.
- Students attending schools with advantaged socio-economic intake perform on average higher than what would be predicted based on their socio-economic background.
- Students attending schools with average or a mixture of socio-economic intake and from the bottom quarter of ESCS on average perform higher than what would be predicted based on their socio-economic background; the reverse is observed for students from the top quarter.

These results confirm a school socio-economic composition effect on student performance in mathematics. These composition effects appear small in northern European countries but large in countries such as Belgium, Czech Republic and Germany and the Netherlands.

As stated previously, composition effect can be estimated by multilevel regression modelling. Such modelling does not require allocating schools to some groups. On the other hand, residual analyses can easily be explained to policy makers.



STATISTICAL PROCEDURE

Box 14.6 presents the SPSS® syntax for running a residual analysis that involves plausible values. First, five regression models need to be computed, each of them on one of the five plausible values. The regression residuals are saved. These five regression residuals are then analysed according to a similar procedure as any plausible values.

Box 14.6 [1/2] SPSS® syntax for residual analyses (e.g. PISA 2003)

```

GET FILE="C:\PISA\2003\DATA\INT_STUI_2003.SAV" .
SELECT IF (CNT="AUS" | CNT="AUT" | CNT="BEL" | CNT="CAN" | CNT="CHE" |
CNT="CZE" | CNT="DEU" | CNT="DNK" | CNT="ESP" | CNT="FIN" | CNT="FRA" |
CNT="GBR" | CNT="GRC" | CNT="HUN" | CNT="ISL" | CNT="IRL" | CNT="ITA" | CNT="JPN" |
CNT="KOR" | CNT="LUX" | CNT="MEX" | CNT="NLD" | CNT="NZL" | CNT="NOR" | CNT="POL" |
CNT="PRT" | CNT="SVK" | CNT="SWE" | CNT="TUR" | CNT="USA" ).
SAVE OUTFILE="C:\TEMP\STUDENT.SAV"
/KEEP=CNT SCHOOLID STIDSTD W_FSTUWT W_FSTR1 to W_FSTR80 ESCS
PV1MATH PV2MATH PV3MATH PV4MATH PV5MATH.
GET FILE="C:\TEMP\STUDENT.SAV" .
SORT CASES BY CNT(A) SCHOOLID(A) STIDSTD (A) .
WEIGHT BY W_FSTUWT.
AGGREGATE
  /OUTFILE=* MODE=ADDVARIABLES
  /BREAK=CNT
  /CNT_ESCS=MEAN(ESCS) .
AGGREGATE
  /OUTFILE=* MODE=ADDVARIABLES
  /BREAK=CNT SCHOOLID
  /SCH_ESCS=MEAN(ESCS) .
WEIGHT OFF.
AGGREGATE
  /OUTFILE=* MODE=ADDVARIABLES
  /BREAK=CNT SCHOOLID
  /SCH_STD=SD(ESCS)
  /SCH_NU=NU(ESCS) .
DO IF (SCH_NU >0) .
COMPUTE SE=((SCH_STD**2)/(SCH_NU))**0.5 .
END IF.
DO IF (SE>0) .
COMPUTE Ttest=(SCH_ESCS - CNT_ESCS)/SE.
IF (Ttest<=-1.96) TYPE=1.
IF (Ttest>-1.96 and Ttest<1.96) TYPE=2.
IF (Ttest>=1.96) TYPE=3.
END IF.
EXECUTE .
SAVE OUTFILE="C:\TEMP\STUDENT2.SAV" .

OMS /SELECT TABLES /IF SUBTYPES ="Statistics" /DESTINATION FORMAT=SAV
OUTFILE="C:\TEMP\PCT.SAV" /TAG="PCT" .
SPLIT FILE LAYERED BY CNT.
WEIGHT BY W_FSTUWT.
FRE ESCS /FORMAT=NOTABLE /PERCENTILES 25 75.
SPLIT FILE OFF.
OMSEND TAG="PCT" .
EXECUTE.

GET FILE="C:\TEMP\PCT.SAV" .
SELECT IF (SUBSTR(VAR2,1,11)="Percentiles") .
RENAME VARIABLES (VAR1=CNT) (VAR3=PCT) (VAR5=PCT25) .
SELECT IF (SUBSTR(PCT,1,2)="25") .
SAVE OUTFILE="C:\TEMP\PCT25.SAV" /KEEP=CNT PCT25.
GET FILE="C:\TEMP\PCT.SAV" .
SELECT IF (SUBSTR(VAR2,1,11)="Percentiles") .
RENAME VARIABLES (VAR1=CNT) (VAR3=PCT) (VAR5=PCT75) .
SELECT IF (SUBSTR(PCT,1,2)="75") .
SAVE OUTFILE="C:\TEMP\PCT75.SAV" /KEEP=CNT PCT75.

```



Box 14.6 [2/2] SPSS® syntax for residual analyses (e.g. PISA 2003)

```

GET FILE="C:\TEMP\STUDENT2.SAV" .
MATCH FILES /FILE=*
  /TABLE="C:\TEMP\pct25.sav"
  /TABLE="C:\TEMP\pct75.sav"
  /BY CNT.
SAVE OUTFILE="C:\TEMP\STUDENT3.SAV" .
IF (ESCS<=PCT25) STATUT=1.
IF (ESCS>PCT25 AND ESCS<=PCT75) STATUT=2.
IF (ESCS>PCT75) STATUT=3.
EXECUTE.

SPLIT FILE LAYERED BY CNT.
WEIGHT BY W_FSTUWT.
REGRESSION _
  /DEPENDENT PV1MATH
  /METHOD=ENTER ESCS
  /SAVE RESID.
REGRESSION
  /DEPENDENT PV2MATH
  /METHOD=ENTER ESCS
  /SAVE RESID.
REGRESSION
  /DEPENDENT PV3MATH
  /METHOD=ENTER ESCS
  /SAVE RESID.
REGRESSION
  /DEPENDENT PV4MATH
  /METHOD=ENTER ESCS
  /SAVE RESID.
REGRESSION
  /DEPENDENT PV5MATH
  /METHOD=ENTER ESCS
  /SAVE RESID.
SAVE OUTFILE="C:\TEMP\STUDENT4.SAV" .

WEIGHT OFF.

INSERT FILE="C:\PISA\MACRO\MCR_SE_GRPPT.SPS" .

GRPPT WITHIN=CNT/
  GRP=TYPE/
  INFILE="C:\TEMP\STUDENT4.SAV" / .
EXEC.

GRPPT WITHIN=CNT/
  GRP=TYPE/ .
EXEC.

INSERT FILE="C:\PISA\MACRO\MCR_SE_PV.SPS" .

SET MPRINT=YES.

PV STAT=MEAN/
  DEP=RES_1 RES_2 RES_3 RES_4 RES_5/
  INFILE="C:\TEMP\STUDENT4.SAV"/
  GRP=CNT TYPE STATUT/ .

```

CONCLUSION

This chapter was devoted to some statistical issues related to the way the OECD reported the relationship between questionnaire indices and student performance in the initial reports.

The PISA initial and thematic reports extensively use the linear regression, in particular for estimating the impact of contextual indices on student performance. An alternative use of the linear regression that consists of analysing the residuals was also presented.



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 Boeck, 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?* *aSoc. Forces*, No. 78, pp. 1349-79.
- Cochran, W.G.** (1977), *Sampling Techniques*, J. Wiley and Sons, Inc., New York.
- Dunn, O.J.** (1961), "Multiple 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

FOREWORD	3
USER'S GUIDE	17
CHAPTER 1 THE USEFULNESS OF PISA DATA FOR POLICY MAKERS, RESEARCHERS AND EXPERTS ON METHODOLOGY	19
PISA – an overview	20
▪ The PISA surveys.....	20
How can PISA contribute to educational policy, practice and research?	22
▪ Key results from PISA 2000, PISA 2003 and PISA 2006.....	23
Further analyses of PISA datasets	25
▪ Contextual framework of PISA 2006.....	28
▪ Influence of the methodology on outcomes.....	31
CHAPTER 2 EXPLORATORY ANALYSIS PROCEDURES	35
Introduction	36
Weights	36
Replicates for computing the standard error	39
Plausible values	43
Conclusion	45
CHAPTER 3 SAMPLE WEIGHTS	47
Introduction	48
Weights for simple random samples	49
Sampling designs for education surveys	51
Why do the PISA weights vary?	55
Conclusion	56
CHAPTER 4 REPLICATE WEIGHTS	57
Introduction	58
Sampling variance for simple random sampling	58
Sampling variance for two-stage sampling	63
Replication methods for simple random samples	68
Replication methods for two-stage samples	70
▪ The Jackknife for unstratified two-stage sample designs.....	70
▪ The Jackknife for stratified two-stage sample designs.....	71
▪ The Balanced Repeated Replication method.....	72
Other procedures for accounting for clustered samples	74
Conclusion	74

CHAPTER 5 THE RASCH MODEL	77
Introduction	78
How can the information be summarised?	78
The Rasch Model for dichotomous items	79
▪ Introduction to the Rasch Model.....	79
▪ Item calibration.....	83
▪ Computation of a student's score.....	85
▪ Computation of a student's score for incomplete designs.....	89
▪ Optimal conditions for linking items.....	90
▪ Extension of the Rasch Model.....	91
Other item response theory models	92
Conclusion	92
 CHAPTER 6 PLAUSIBLE VALUES	 93
Individual estimates versus population estimates	94
The meaning of plausible values (PVs)	94
Comparison of the efficiency of WLEs, EAP estimates and PVs for the estimation of some population statistics	97
How to perform analyses with plausible values	100
Conclusion	101
 CHAPTER 7 COMPUTATION OF STANDARD ERRORS	 103
Introduction	104
The standard error on univariate statistics for numerical variables	104
The SPSS® macro for computing the standard error on a mean	107
The standard error on percentages	110
The standard error on regression coefficients	112
The standard error on correlation coefficients	114
Conclusion	115
 CHAPTER 8 ANALYSES WITH PLAUSIBLE VALUES	 117
Introduction	118
Univariate statistics on plausible values	118
The standard error on percentages with PVs	121
The standard error on regression coefficients with PVs	121
The standard error on correlation coefficients with PVs	124
Correlation between two sets of plausible values	124
A fatal error shortcut	128
An unbiased shortcut	129
Conclusion	130
 CHAPTER 9 USE OF PROFICIENCY LEVELS	 133
Introduction	134
Generation of the proficiency levels	134
Other analyses with proficiency levels	139
Conclusion	141



CHAPTER 10 ANALYSES WITH SCHOOL-LEVEL VARIABLES	143
Introduction	144
Limits of the PISA school samples	145
Merging the school and student data files	146
Analyses of the school variables	146
Conclusion	148
CHAPTER 11 STANDARD ERROR ON A DIFFERENCE	149
Introduction	150
Statistical issues and computing standard errors on differences	150
The standard error on a difference without plausible values	152
The standard error on a difference with plausible values	157
Multiple comparisons	161
Conclusion	162
CHAPTER 12 OECD TOTAL AND OECD AVERAGE	163
Introduction	164
Recoding of the database to estimate the pooled OECD total and the pooled OECD average	166
Duplication of the data to avoid running the procedure three times	168
Comparisons between the pooled OECD total or pooled OECD average estimates and a country estimate	169
Comparisons between the arithmetic OECD total or arithmetic OECD average estimates and a country estimate	171
Conclusion	171
CHAPTER 13 TRENDS	173
Introduction	174
The computation of the standard error for trend indicators on variables other than performance	175
The computation of the standard error for trend indicators on performance variables	177
Conclusion	181
CHAPTER 14 STUDYING THE RELATIONSHIP BETWEEN STUDENT PERFORMANCE AND INDICES DERIVED FROM CONTEXTUAL QUESTIONNAIRES	183
Introduction	184
Analyses by quarters	184
The concept of relative risk	186
▪ Instability of the relative risk	187
▪ Computation of the relative risk	188
Effect size	191
Linear regression and residual analysis	193
▪ Independence of errors	193
Statistical procedure	196
Conclusion	197



CHAPTER 15 MULTILEVEL ANALYSES	199
Introduction	200
Two-level modelling with SPSS®	202
▪ Decomposition of the variance in the empty model.....	202
▪ Models with only random intercepts.....	205
▪ Shrinkage factor.....	207
▪ Models with random intercepts and fixed slopes.....	207
▪ Models with random intercepts and random slopes.....	209
▪ Models with Level 2 independent variables.....	214
▪ Computation of final estimates and their respective standard errors.....	217
Three-level modelling	219
Limitations of the multilevel model in the PISA context	221
Conclusion	222
 CHAPTER 16 PISA AND POLICY RELEVANCE – THREE EXAMPLES OF ANALYSES	 223
Introduction	224
Example 1: Gender differences in performance	224
Example 2: Promoting socio-economic diversity within school?	228
Example 3: The influence of an educational system on the expected occupational status of students at age 30	234
Conclusion	237
 CHAPTER 17 SPSS® MACRO	 239
Introduction	240
Structure of the SPSS® Macro	240
 REFERENCES	 321
 APPENDICES	 323
Appendix 1 Three-level regression analysis.....	324
Appendix 2 PISA 2006 International database.....	332
Appendix 3 PISA 2006 Student questionnaire.....	341
Appendix 4 PISA 2006 Information communication technology (ICT) Questionnaire.....	350
Appendix 5 PISA 2006 School questionnaire.....	352
Appendix 6 PISA 2006 Parent questionnaire.....	359
Appendix 7 Codebook for PISA 2006 student questionnaire data file.....	363
Appendix 8 Codebook for PISA 2006 non-scored cognitive and embedded attitude items.....	407
Appendix 9 Codebook for PISA 2006 scored cognitive and embedded attitude items.....	427
Appendix 10 Codebook for PISA 2006 school questionnaire data file.....	439
Appendix 11 Codebook for PISA 2006 parents questionnaire data file.....	450
Appendix 12 PISA 2006 questionnaire indices.....	456



LIST OF BOXES

Box 2.1	WEIGHT statement in SPSS®.....	37
<hr/>		
Box 7.1	SPSS® syntax for computing 81 means (e.g. PISA 2003).....	104
Box 7.2	SPSS® syntax for computing the mean of HISEI and its standard error (e.g. PISA 2003).....	107
Box 7.3	SPSS® syntax for computing the standard deviation of HISEI and its standard error by gender (e.g. PISA 2003).....	109
Box 7.4	SPSS® syntax for computing the percentages and their standard errors for gender (e.g. PISA 2003).....	110
Box 7.5	SPSS® syntax for computing the percentages and its standard errors for grades by gender (e.g. PISA 2003).....	112
Box 7.6	SPSS® syntax for computing regression coefficients, R^2 and its respective standard errors: Model 1 (e.g. PISA 2003).....	113
Box 7.7	SPSS® syntax for computing regression coefficients, R^2 and its respective standard errors: Model 2 (e.g. PISA 2003).....	114
Box 7.8	SPSS® syntax for computing correlation coefficients and its standard errors (e.g. PISA 2003).....	114
<hr/>		
Box 8.1	SPSS® syntax for computing the mean on the science scale by using the MCR_SE_UNIV macro (e.g. PISA 2006).....	119
Box 8.2	SPSS® syntax for computing the mean and its standard error on PVs (e.g. PISA 2006).....	120
Box 8.3	SPSS® syntax for computing the standard deviation and its standard error on PVs by gender (e.g. PISA 2006).....	131
Box 8.4	SPSS® syntax for computing regression coefficients and their standard errors on PVs by using the MCR_SE_REG macro (e.g. PISA 2006).....	122
Box 8.5	SPSS® syntax for running the simple linear regression macro with PVs (e.g. PISA 2006).....	123
Box 8.6	SPSS® syntax for running the correlation macro with PVs (e.g. PISA 2006).....	124
Box 8.7	SPSS® syntax for the computation of the correlation between mathematics/quantity and mathematics/space and shape by using the MCR_SE_COR_2PV macro (e.g. PISA 2003).....	126
<hr/>		
Box 9.1	SPSS® syntax for generating the proficiency levels in science (e.g. PISA 2006).....	135
Box 9.2	SPSS® syntax for computing the percentages of students by proficiency level in science and its standard errors (e.g. PISA 2006).....	136
Box 9.3	SPSS® syntax for computing the percentage of students by proficiency level in science and its standard errors (e.g. PISA 2006).....	138
Box 9.4	SPSS® syntax for computing the percentage of students by proficiency level and its standard errors by gender (e.g. PISA 2006).....	138
Box 9.5	SPSS® syntax for generating the proficiency levels in mathematics (e.g. PISA 2003).....	139
Box 9.6	SPSS® syntax for computing the mean of self-efficacy in mathematics and its standard errors by proficiency level (e.g. PISA 2003).....	140
<hr/>		
Box 10.1	SPSS® syntax for merging the student and school data files (e.g. PISA 2006).....	146
Box 10.2	Question on school location in PISA 2006.....	147
Box 10.3	SPSS® syntax for computing the percentage of students and the average performance in science, by school location (e.g. PISA 2006).....	147
<hr/>		
Box 11.1	SPSS® syntax for computing the mean of job expectations by gender (e.g. PISA 2003).....	152
Box 11.2	SPSS® macro for computing standard errors on differences (e.g. PISA 2003).....	155



Box 11.3	Alternative SPSS® macro for computing the standard error on a difference for a dichotomous variable (e.g. PISA 2003).....	156
Box 11.4	SPSS® syntax for computing standard errors on differences which involve PVs (e.g. PISA 2003).....	158
Box 11.5	SPSS® syntax for computing standard errors on differences that involve PVs (e.g. PISA 2006).....	160
<hr/>		
Box 12.1	SPSS® syntax for computing the pooled OECD total for the mathematics performance by gender (e.g. PISA 2003).....	166
Box 12.2	SPSS® syntax for the pooled OECD average for the mathematics performance by gender (e.g. PISA 2003).....	167
Box 12.3	SPSS® syntax for the creation of a larger dataset that will allow the computation of the pooled OECD total and the pooled OECD average in one run (e.g. PISA 2003).....	168
<hr/>		
Box 14.1	SPSS® syntax for the quarter analysis (e.g. PISA 2006).....	185
Box 14.2	SPSS® syntax for computing the relative risk with five antecedent variables and five outcome variables (e.g. PISA 2006).....	189
Box 14.3	SPSS® syntax for computing the relative risk with one antecedent variable and one outcome variable (e.g. PISA 2006).....	190
Box 14.4	SPSS® syntax for computing the relative risk with one antecedent variable and five outcome variables (e.g. PISA 2006).....	190
Box 14.5	SPSS® syntax for computing effect size (e.g. PISA 2006).....	192
Box 14.6	SPSS® syntax for residual analyses (e.g. PISA 2003).....	196
<hr/>		
Box 15.1	Normalisation of the final student weights (e.g. PISA 2006).....	203
Box 15.2	SPSS® syntax for the decomposition of the variance in student performance in science (e.g. PISA 2006).....	203
Box 15.3	SPSS® syntax for normalising PISA 2006 final student weights with deletion of cases with missing values and syntax for variance decomposition (e.g. PISA 2006).....	206
Box 15.4	SPSS® syntax for a multilevel regression model with random intercepts and fixed slopes (e.g. PISA 2006).....	208
Box 15.5	Results for the multilevel model in Box 15.4.....	208
Box 15.6	SPSS® syntax for a multilevel regression model (e.g. PISA 2006).....	210
Box 15.7	Results for the multilevel model in Box 15.6.....	211
Box 15.8	Results for the multilevel model with covariance between random parameters.....	212
Box 15.9	Interpretation of the within-school regression coefficient.....	214
Box 15.10	SPSS® syntax for a multilevel regression model with a school-level variable (e.g. PISA 2006).....	214
Box 15.11	SPSS® syntax for a multilevel regression model with interaction (e.g. PISA 2006).....	215
Box 15.12	Results for the multilevel model in Box 15.11.....	216
Box 15.13	SPSS® syntax for using the multilevel regression macro (e.g. PISA 2006).....	217
Box 15.14	SPSS® syntax for normalising the weights for a three-level model (e.g. PISA 2006).....	219
<hr/>		
Box 16.1	SPSS® syntax for testing the gender difference in standard deviations of reading performance (e.g. PISA 2000).....	225
Box 16.2	SPSS® syntax for computing the 5th percentile of the reading performance by gender (e.g. PISA 2000).....	227
Box 16.3	SPSS® syntax for preparing a data file for the multilevel analysis.....	230



Box 16.4	SPSS® syntax for running a preliminary multilevel analysis with one PV	231
Box 16.5	Estimates of fixed parameters in the multilevel model.....	231
Box 16.6	SPSS® syntax for running preliminary analysis with the MCR_ML_PV macro.....	233
Box 17.1	SPSS® macro of MCR_SE_UNI.sps.....	243
Box 17.2	SPSS® macro of MCR_SE_PV.sps.....	247
Box 17.3	SPSS® macro of MCR_SE_PERCENTILES_PV.sps	251
Box 17.4	SPSS® macro of MCR_SE_GrpPct.sps.....	254
Box 17.5	SPSS® macro of MCR_SE_PctLev.sps.....	257
Box 17.6	SPSS® macro of MCR_SE_REG.sps	261
Box 17.7	SPSS® macro of MCR_SE_REG_PV.sps.....	265
Box 17.8	SPSS® macro of MCR_SE_COR.sps.....	270
Box 17.9	SPSS® macro of MCR_SE_COR_1PV.sps.....	273
Box 17.10	SPSS® macro of MCR_SE_COR_2PV.sps.....	277
Box 17.11	SPSS® macro of MCR_SE_DIFF.sps.....	281
Box 17.12	SPSS® macro of MCR_SE_DIFF_PV.sps.....	285
Box 17.13	SPSS® macro of MCR_SE_PV_WLEQRT.sps.....	290
Box 17.14	SPSS® macro of MCR_SE_RR.sps.....	295
Box 17.15	SPSS® macro of MCR_SE_RR_PV.sps.....	298
Box 17.16	SPSS® macro of MCR_SE_EFFECT.sps.....	302
Box 17.17	SPSS® macro of MCR_SE_EFFECT_PV.sps	306
Box 17.18	SPSS® macro of MCR_ML.sps.....	311
Box 17.19	SPSS® macro of MCR_ML_PV.sps	315
Box A1.1	Descriptive statistics of background and explanatory variables.....	326
Box A1.2	Background model for student performance.....	327
Box A1.3	Final net combined model for student performance.....	328
Box A1.4	Background model for the impact of socio-economic background.....	329
Box A1.5	Model of the impact of socio-economic background: “school resources” module.....	330
Box A1.6	Model of the impact of socio-economic background: “accountability practices” module	331
Box A1.7	Final combined model for the impact of socio-economic background.....	331

LIST OF FIGURES

Figure 1.1	Relationship between social and academic segregations.....	27
Figure 1.2	Relationship between social segregation and the correlation between science performance and student HISEI	27
Figure 1.3	Conceptual grid of variable types.....	29
Figure 1.4	Two-dimensional matrix with examples of variables collected or available from other sources	30
Figure 2.1	Science mean performance in OECD countries (PISA 2006).....	37
Figure 2.2	Gender differences in reading in OECD countries (PISA 2000).....	38
Figure 2.3	Regression coefficient of ESCS on mathematic performance in OECD countries (PISA 2003).....	38
Figure 2.4	Design effect on the country mean estimates for science performance and for ESCS in OECD countries (PISA 2006)	41
Figure 2.5	Simple random sample and unbiased standard errors of ESCS on science performance in OECD countries (PISA 2006)	42



Figure 4.1	Distribution of the results of 36 students.....	58
Figure 4.2	Sampling variance distribution of the mean.....	60
Figure 5.1	Probability of success for two high jumpers by height (dichotomous).....	80
Figure 5.2	Probability of success for two high jumpers by height (continuous).....	81
Figure 5.3	Probability of success to an item of difficulty zero as a function of student ability.....	81
Figure 5.4	Student score and item difficulty distributions on a Rasch continuum.....	84
Figure 5.5	Response pattern probabilities for the response pattern (1, 1, 0, 0).....	86
Figure 5.6	Response pattern probabilities for a raw score of 1.....	87
Figure 5.7	Response pattern probabilities for a raw score of 2.....	88
Figure 5.8	Response pattern probabilities for a raw score of 3.....	88
Figure 5.9	Response pattern likelihood for an easy test and a difficult test.....	89
Figure 5.10	Rasch item anchoring.....	90
Figure 6.1	Living room length expressed in integers.....	94
Figure 6.2	Real length per reported length.....	95
Figure 6.3	A posterior distribution on a test of six items.....	96
Figure 6.4	EAP estimators.....	97
Figure 8.1	A two-dimensional distribution.....	125
Figure 8.2	Axes for two-dimensional normal distributions.....	125
Figure 13.1	Trend indicators in PISA 2000, PISA 2003 and PISA 2006.....	175
Figure 14.1	Percentage of schools by three school groups (PISA 2003).....	194
Figure 15.1	Simple linear regression analysis versus multilevel regression analysis.....	201
Figure 15.2	Graphical representation of the between-school variance reduction.....	209
Figure 15.3	A random multilevel model.....	210
Figure 15.4	Change in the between-school residual variance for a fixed and a random model.....	212
Figure 16.1	Relationship between the segregation index of students' expected occupational status and the segregation index of student performance in reading (PISA 2000).....	236
Figure 16.2	Relationship between the segregation index of students' expected occupational status and the correlation between HISEI and students' expected occupational status.....	236

LIST OF TABLES

Table 1.1	Participating countries/economies in PISA 2000, PISA 2003, PISA 2006 and PISA 2009.....	21
Table 1.2	Assessment domains covered by PISA 2000, PISA 2003 and PISA 2006.....	22
Table 1.3	Correlation between social inequities and segregations at schools for OECD countries.....	28
Table 1.4	Distribution of students per grade and per ISCED level in OECD countries (PISA 2006).....	31
Table 2.1	Design effect and type I errors.....	40
Table 2.2	Mean estimates and standard errors.....	44



Table 2.3	Standard deviation estimates and standard errors.....	44
Table 2.4	Correlation estimates and standard errors.....	45
Table 2.5	ESCS regression coefficient estimates and standard errors.....	45
<hr/>		
Table 3.1	Height and weight of ten persons	50
Table 3.2	Weighted and unweighted standard deviation estimate	50
Table 3.3	School, within-school, and final probability of selection and corresponding weights for a two-stage, simple random sample with the first-stage units being schools of equal size.....	52
Table 3.4	School, within-school, and final probability of selection and corresponding weights for a two-stage, simple random sample with the first-stage units being schools of unequal size	52
Table 3.5	School, within-school, and final probability of selection and corresponding weights for a simple and random sample of schools of unequal size (smaller schools)	53
Table 3.6	School, within-school, and final probability of selection and corresponding weights for a simple and random sample of schools of unequal size (larger schools)	53
Table 3.7	School, within-school, and final probability of selection and corresponding weights for PPS sample of schools of unequal size	54
Table 3.8	Selection of schools according to a PPS and systematic procedure.....	55
<hr/>		
Table 4.1	Description of the 630 possible samples of 2 students selected from 36 students, according to their mean.....	59
Table 4.2	Distribution of all possible samples with a mean between 8.32 and 11.68.....	61
Table 4.3	Distribution of the mean of all possible samples of 4 students out of a population of 36 students.....	62
Table 4.4	Between-school and within-school variances on the mathematics scale in PISA 2003.....	65
Table 4.5	Current status of sampling errors.....	65
Table 4.6	Between-school and within-school variances, number of participating schools and students in Denmark and Germany in PISA 2003	66
Table 4.7	The Jackknives replicates and sample means.....	68
Table 4.8	Values on variables X and Y for a sample of ten students.....	69
Table 4.9	Regression coefficients for each replicate sample.....	69
Table 4.10	The Jackknife replicates for unstratified two-stage sample designs.....	70
Table 4.11	The Jackknife replicates for stratified two-stage sample designs.....	71
Table 4.12	Replicates with the Balanced Repeated Replication method.....	72
Table 4.13	The Fay replicates	73
<hr/>		
Table 5.1	Probability of success when student ability equals item difficulty.....	82
Table 5.2	Probability of success when student ability is less than the item difficulty by 1 unit.....	82
Table 5.3	Probability of success when student ability is greater than the item difficulty by 1 unit	82
Table 5.4	Probability of success when student ability is less than the item difficulty by 2 units	83
Table 5.5	Probability of success when student ability is greater than the item difficulty by 2 units.....	83
Table 5.6	Possible response pattern for a test of four items.....	85
Table 5.7	Probability for the response pattern (1, 1, 0, 0) for three student abilities.....	85
Table 5.8	Probability for the response pattern (1, 0) for two students of different ability in an incomplete test design.....	89
Table 5.9	PISA 2003 test design	91

Table 6.1	Structure of the simulated data.....	98
Table 6.2	Means and variances for the latent variables and the different student ability estimators.....	98
Table 6.3	Percentiles for the latent variables and the different student ability estimators.....	99
Table 6.4	Correlation between HISEI, gender and the latent variable, the different student ability estimators.....	99
Table 6.5	Between- and within-school variances.....	100
Table 7.1	HISEI mean estimates	105
Table 7.2	Squared differences between replicate estimates and the final estimate.....	106
Table 7.3	Output data file from Box 7.2.....	108
Table 7.4	Available statistics with the UNIVAR macro	109
Table 7.5	Output data file from Box 7.3.....	109
Table 7.6	Output data file from Box 7.4.....	110
Table 7.7	Percentage of girls for the final and replicate weights and squared differences.....	111
Table 7.8	Output data file from Box 7.5.....	112
Table 7.9	Output data file from Box 7.6.....	113
Table 7.10	Output data file from Box 7.7.....	114
Table 7.11	Output data file from Box 7.8.....	114
Table 8.1	The 405 mean estimates.....	118
Table 8.2	Mean estimates and their respective sampling variances on the science scale for Belgium (PISA 2006).....	119
Table 8.3	Output data file from Box 8.2.....	121
Table 8.4	Output data file from Box 8.3.....	121
Table 8.5	The 450 regression coefficient estimates.....	123
Table 8.6	HISEI regression coefficient estimates and their respective sampling variance on the science scale in Belgium after accounting for gender (PISA 2006).....	123
Table 8.7	Output data file from Box 8.5.....	123
Table 8.8	Output data file from Box 8.6.....	124
Table 8.9	Correlation between the five plausible values for each domain, mathematics/quantity and mathematics/space and shape.....	126
Table 8.10	The five correlation estimates between mathematics/quantity and mathematics/space and shape and their respective sampling variance.....	127
Table 8.11	Standard deviations for mathematics scale using the correct method (plausible values) and by averaging the plausible values at the student level (pseudo-EAP) (PISA 2003).....	128
Table 8.12	Unbiased shortcut for a population estimate and its standard error	129
Table 8.13	Standard errors from the full and shortcut computation (PISA 2006).....	130
Table 9.1	The 405 percentage estimates for a particular proficiency level	136
Table 9.2	Estimates and sampling variances per proficiency level in science for Germany (PISA 2006)	137
Table 9.3	Final estimates of the percentage of students, per proficiency level, in science and its standard errors for Germany (PISA 2006).....	137
Table 9.4	Output data file from Box 9.3.....	138
Table 9.5	Output data file from Box 9.4.....	138
Table 9.6	Mean estimates and standard errors for self-efficacy in mathematics per proficiency level (PISA 2003).....	141
Table 9.7	Output data file from Box 9.6.....	141



Table 10.1	Percentage of students per grade and ISCED level, by country (PISA 2006)	144
Table 10.2	Output data file from the first model in Box 10.3	148
Table 10.3	Output data file from the second model in Box 10.3	148
<hr/>		
Table 11.1	Output data file from Box 11.1	153
Table 11.2	Mean estimates for the final and 80 replicate weights by gender (PISA 2003)	153
Table 11.3	Difference in estimates for the final weight and 80 replicate weights between females and males (PISA 2003)	155
Table 11.4	Output data file from Box 11.2	156
Table 11.5	Output data file from Box 11.3	157
Table 11.6	Gender difference estimates and their respective sampling variances on the mathematics scale (PISA 2003)	157
Table 11.7	Output data file from Box 11.4	158
Table 11.8	Gender differences on the mathematics scale, unbiased standard errors and biased standard errors (PISA 2003)	159
Table 11.9	Gender differences in mean science performance and in standard deviation for science performance (PISA 2006)	159
Table 11.10	Regression coefficient of HISEI on the science performance for different models (PISA 2006)	160
Table 11.11	Cross tabulation of the different probabilities	161
<hr/>		
Table 12.1	Regression coefficients of the index of instrumental motivation in mathematics on mathematic performance in OECD countries (PISA 2003)	165
Table 12.2	Output data file from Box 12.1	166
Table 12.3	Output data file from Box 12.2	167
Table 12.4	Difference between the country mean scores in mathematics and the OECD total and average (PISA 2003)	170
<hr/>		
Table 13.1	Trend indicators between PISA 2000 and PISA 2003 for HISEI, by country	176
Table 13.2	Linking error estimates	178
Table 13.3	Mean performance in reading by gender in Germany	180
<hr/>		
Table 14.1	Distribution of the questionnaire index of cultural possession at home in Luxembourg (PISA 2006)	184
Table 14.2	Output data file from Box 14.1	186
Table 14.3	Labels used in a two-way table	186
Table 14.4	Distribution of 100 students by parents' marital status and grade repetition	187
Table 14.5	Probabilities by parents' marital status and grade repetition	187
Table 14.6	Relative risk for different cutpoints	187
Table 14.7	Output data file from Box 14.2	189
Table 14.8	Mean and standard deviation for the student performance in reading by gender, gender difference and effect size (PISA 2006)	191
Table 14.9	Output data file from the first model in Box 14.5	197
Table 14.10	Output data file from the second model in Box 14.5	197
Table 14.11	Mean of the residuals in mathematics performance for the bottom and top quarters of the PISA index of economic, social and cultural status, by school group (PISA 2003)	195

Table 15.1	Between- and within-school variance estimates and intraclass correlation (PISA 2006).....	204
Table 15.2	Fixed parameter estimates	211
Table 15.3	Variance/covariance estimates before and after centering.....	213
Table 15.4	Output data file of the fixed parameters file.....	215
Table 15.5	Average performance and percentage of students by student immigrant status and by type of school.....	216
Table 15.6	Variables for the four groups of students	216
Table 15.7	Comparison of the regression coefficient estimates and their standard errors in Belgium (PISA 2006).....	218
Table 15.8	Comparison of the variance estimates and their respective standard errors in Belgium (PISA 2006)	218
Table 15.9	Three-level regression analyses.....	220
Table 16.1	Differences between males and females in the standard deviation of student performance (PISA 2000).....	226
Table 16.2	Distribution of the gender differences (males – females) in the standard deviation of the student performance	226
Table 16.3	Gender difference on the PISA combined reading scale for the 5 th , 10 th , 90 th and 95 th percentiles (PISA 2000)	227
Table 16.4	Gender difference in the standard deviation for the two different item format scales in reading (PISA 2000)	228
Table 16.5	Random and fixed parameters in the multilevel model with student and school socio-economic background.....	229
Table 16.6	Random and fixed parameters in the multilevel model with socio-economic background and grade retention at the student and school levels	233
Table 16.7	Segregation indices and correlation coefficients by country (PISA 2000).....	234
Table 16.8	Segregation indices and correlation coefficients by country (PISA 2006).....	235
Table 16.9	Country correlations (PISA 2000).....	237
Table 16.10	Country correlations (PISA 2006).....	237
Table 17.1	Synthesis of the 19 SPSS® macros.....	241
Table A2.1	Cluster rotation design used to form test booklets for PISA 2006	332
Table A12.1	Mapping of ISCED to accumulated years of education	457
Table A12.2	ISCO major group white-collar/blue-collar classification	459
Table A12.3	ISCO occupation categories classified as science-related occupations	459
Table A12.4	Household possessions and home background indices.....	463
Table A12.5	Factor loadings and internal consistency of ESCS 2006 in OECD countries.....	473
Table A12.6	Factor loadings and internal consistency of ESCS 2006 in partner countries/economies.....	474



User's Guide

Preparation of data files

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

SPSS® users

By running the SPSS® control files, the PISA data files are created in the SPSS® format. Before starting analysis in the following chapters, save the PISA 2000 data files in the folder of “c:\pisa2000\data\”, the PISA 2003 data files in “c:\pisa2003\data\”, and the PISA 2006 data files in “c:\pisa2006\data\”.

SPSS® syntax and macros

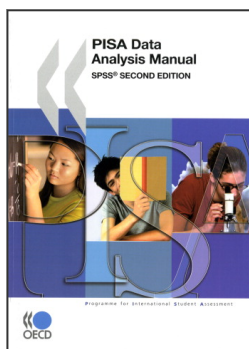
All syntaxes and macros in this manual can be copied from the PISA website (www.pisa.oecd.org). These macros were developed for SPSS 17.0. The 19 SPSS® macros presented in Chapter 17 need to be saved under “c:\pisa\macro\”, before starting analysis. Each chapter of the manual contains a complete set 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: SPSS, Second Edition

Access the complete publication at:
<https://doi.org/10.1787/9789264056275-en>

Please cite this chapter as:

OECD (2009), “Studying the Relationship between Student Performance and Indices Derived from Contextual Questionnaires”, in *PISA Data Analysis Manual: SPSS, Second Edition*, OECD Publishing, Paris.

DOI: <https://doi.org/10.1787/9789264056275-15-en>

This work is published under the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of OECD member countries.

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgment of OECD as source and copyright owner is given. All requests for public or commercial use and translation rights should be submitted to rights@oecd.org. Requests for permission to photocopy portions of this material for public or commercial use shall be addressed directly to the Copyright Clearance Center (CCC) at info@copyright.com or the Centre français d'exploitation du droit de copie (CFC) at contact@cfcopies.com.