



3

# Sample Weights

Introduction .....	50
Weights for simple random samples.....	51
Sampling designs for education surveys .....	53
Why do the PISA weights vary? .....	57
Conclusion.....	58



## INTRODUCTION

National and international surveys usually collect data from a sample. Dealing with a sample rather than the whole population is preferable for several reasons.

First, for a census, all members of the population need to be identified. This identification process presents no major difficulty for human populations in some countries, where national databases with the name and address of all, or nearly all, citizens may be available. However, in other countries, it is not possible for the researcher to identify all members or sampling units of the target population, mainly because it would be too time-consuming or because of the nature of the target population.

Second, even if all members of a population are easily identifiable, researchers may still draw from a sample, because dealing with the whole population:

- might require unreasonable budgets;
- is time-consuming and thus incompatible with publication deadlines;
- does not necessarily help with obtaining additional and/or required information.

Drawing a sample can be done in several ways depending on the population characteristics and the survey research questions. All sample designs aim to avoid bias in the selection procedure and achieve the maximum precision in view of the available resources. Nevertheless, biases in the selection can arise:

- If the sampling is done by a non-random method, which generally means that the selection is consciously or unconsciously influenced by human choices. The importance of randomness in the selection procedure should not be underestimated;
- If the sampling frame (list, index, or other population record) that serves as the basis for selection does not cover the population adequately, completely or accurately.

Biases can also arise if some sections of the population are impossible to find or refuse to co-operate. In educational surveys, schools might refuse to participate and within participating schools, some students might refuse to participate or simply be absent on the day of the assessment. The size of the bias introduced by the school or student non-response is proportional to the correlation between the school, or the student, propensity to participate and the variables measured with cognitive tests or contextual questionnaires. For instance, it may be that low achievers are more likely to be absent on the day of the assessment than high achievers. On the other hand, it would be less likely to observe a correlation between the height of a student and his/her propensity to participate. The non-response would therefore not introduce a bias in the height mean estimate.

To limit the size of the bias due to non-response, international education surveys require a minimal student participation rate. For PISA, this minimum is 80%.

Finally, if the sampling units do not have the same chances to be selected and if the population parameters are estimated without taking into account these varying probabilities, then results might also be biased. To compensate for these varying probabilities, data need to be weighted. Weighting consists of acknowledging that some units in the sample are more important than others and have to contribute more than others for any population estimates. A sampling unit with a very small probability of selection will be considered as more important than a sampling unit with a high probability of selection. Weights are therefore inversely proportional to the probability of selection.

Nevertheless, a sample is only useful to the extent that it can estimate some characteristics of the whole population. This means that statistical estimates computed on the sample, including a mean, a standard



deviation, a correlation, a regression coefficient, and so on, can be generalised to the population. This generalisation is more reliable if the sampling requirements have been met.

Depending on the sampling design, selection probabilities and procedures to compute the weights will vary. These variations are discussed in the following sections.

## WEIGHTS FOR SIMPLE RANDOM SAMPLES

Selecting members of a population by simple random sampling is the most straightforward procedure. There are several ways to draw such a sample, *e.g.*:

- The  $N$  members<sup>1</sup> of a population are numbered and  $n$  of them are selected by random numbers without replacement;
- $N$  numbered discs are placed in a container, mixed well, and  $n$  of them are selected at random;
- The  $N$  population members are arranged in a random order, and every  $\frac{N}{n}$ <sup>th</sup> member is then selected; or
- The  $N$  population members are each assigned a random number. The random numbers are sorted from lowest to highest or highest to lowest. The first  $n$  members make up one random sample.

The simple random sample gives an equal probability of selection to each member of the population. If  $n$  members are selected from a population of  $N$  members according to a simple random procedure, then the probability of each member  $i$  to be part of the sample is equal to:

$$p_i = \frac{n}{N}$$

For example, if 40 students are randomly selected from a population of 400 students, the probability of each student  $i$  to be part of the sample is equal to:

$$p_i = \frac{n}{N} = \frac{40}{400} = 0.1$$

In other words, each student has one chance out of ten of being selected.

As mentioned previously, weights are usually defined as the inverse of the probability of selection. In the case of a simple random sample, the weight will be equal to:

$$w_i = \frac{1}{p_i} = \frac{N}{n}$$

The weight of each of the 40 students selected from a population of 400 students will therefore be equal to:

$$w_i = \frac{1}{p_i} = \frac{N}{n} = \frac{400}{40} = 10$$

This means that each student in the sample represents himself or herself, as well as nine other students. Since each unit has the same selection probability in a simple random sample, the weight attached to each selected unit will also be identical. Therefore, the sum of the weights of the selected units will be equal to the population size, *i.e.*  $N$ .

$$\sum_{i=1}^n w_i = \sum_{i=1}^n \frac{N}{n} = N$$



In the example,

$$\sum_{i=1}^{40} 10 = 400$$

Furthermore, since all sampled units have the same weight, the estimation of any population parameter should not be affected by the weights. For instance, consider the mean of some characteristic,  $X$ . The weighted mean is equivalent to the sum of the product of the weight and  $X$  divided by the sum of the weights.

$$\hat{\mu}_{(X)} = \frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i}$$

Since  $w_i$  is a constant, the weighted mean and the unweighted mean will be equal.

$$\hat{\mu}_{(X)} = \frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i} = \frac{w_i \sum_{i=1}^n x_i}{w_i \sum_{i=1}^n 1} = \frac{\sum_{i=1}^n x_i}{n}$$

However, even with an equi-probabilistic sample, statistical software packages might return different results for weighted and unweighted data. As mentioned in Chapter 2, SAS® proposes for instance four options for dividing the weighted sum of square, *i.e.* (i) the number of valid observations; (ii) the number of valid observations minus 1; (iii) the sum of the weights for the valid observations; and (iv) the sum of the weights for the valid observations minus 1. By default, SAS® divides the weighted sum of square by (n-1) while SPSS® divides it by the sum of the weight minus 1.

**Table 3.1**  
Height and weight of ten persons

Individual	Weight	Height
1	10	160
2	10	162
3	10	164
4	10	166
5	10	168
6	10	170
7	10	172
8	10	174
9	10	176
10	10	178

**Table 3.2**  
Weighted and unweighted standard deviation estimate

	Standard deviation estimate
SAS® unweighted estimate	6.0553
SAS® weighted estimate default option	19.14854
SAS® weighted estimate option = N	18.1659
SAS® weighted estimate option = DF	19.14854
SAS® weighted estimate option = WGT	5.74456
SAS® weighted estimate option = WDF	5.7735
SPSS® unweighted estimate	6.0553
SPSS® weighted estimate	5.7735



Table 3.1 presents the height of ten individuals and Table 3.2, the different standard deviation estimates returned by SPSS® and SAS®.

Table 3.2 clearly indicates how a population estimate can be affected by the weighting process offered in the statistical software, even with an equi-probabilistic sample. Data analysts are strongly recommended to carefully read the software documentation related to the weights.

## SAMPLING DESIGNS FOR EDUCATION SURVEYS

Simple random sampling is very rarely used in education surveys because:

- It is too expensive. Indeed, depending on the school population size, it is quite possible that selected students would attend many different schools. This would require the training of a large number of test administrators, the reimbursement of a large amount of travel expenses and so on;
- It is not practical. One would have to contact too many schools; and
- It would be impossible to link, from a statistical point of view, student variables to school, class, or teacher variables. Educational surveys usually try to understand the statistical variability of the student's outcome measure by school or class level variables. With just one or only a few students per school, this statistical relationship would have no stability.

Therefore, surveys in education usually draw up a student sample in two steps. First, a sample of schools is selected from a complete list of schools containing the student population of interest. Then, a simple random sample of students or classes is drawn from within the selected schools. In PISA, usually 35 students from the population of 15-year-olds are randomly selected within the selected schools. If less than 35 15-year-olds attend a selected school, then all of the students will be invited to participate.

This two-stage sampling procedure will have an impact on the calculation of the weights and, similarly, the school selection procedure will affect the characteristics and properties of the student sample.

Suppose that the population of 400 students is distributed in 10 schools, each school containing 40 students. Four schools are selected randomly and within schools, ten students are selected according to a similar procedure. Each school, denoted  $i$ , has a selection probability equal to:

$$p_{1-i} = \frac{n_{sc}}{N_{sc}} = \frac{4}{10} = 0.4 \text{ with } N_{sc} \text{ being the number of schools and } n_{sc} \text{ the number of schools sampled.}$$

Within the four selected schools, each student, denoted  $j$ , has a selection probability equal to:

$$p_{2-ij} = \frac{n_i}{N_i} = \frac{10}{40} = 0.25$$

with  $N_i$  being the number of students in school  $i$  and  $n_i$  the number of students sampled in school  $i$ . This means that within each selected school, each student has a chance of one in four of being sampled.

The final selection probability for student  $j$  attending school  $i$  is equal to the product of the school selection probability by the student selection probability within the school, *i.e.*:

$$p_{ij} = p_{1-i} p_{2-ij} = \frac{n_{sc} n_i}{N_{sc} N_i}$$

In the example, the final student probability is equal to:

$$p_{ij} = p_{1-i} p_{2-ij} = \frac{n_{sc} n_i}{N_{sc} N_i} = \frac{4 * 10}{10 * 40} = 0.4 * 0.25 = 0.10$$



The school weight, denoted  $w_{1_i}$ , the within-school weight, denoted  $w_{2_{ij}}$ , and the final student weight, denoted  $w_{ij}$ , are respectively equal to:

$$w_{1_i} = \frac{1}{p_{1_i}} = \frac{1}{0.4} = 2.5$$

$$w_{2_{ij}} = \frac{1}{p_{2_{ij}}} = \frac{1}{0.25} = 4$$

$$w_{ij} = \frac{1}{p_{ij}} = \frac{1}{0.1} = 10$$

Table 3.3 presents the selection probability at the school level, at the within-school level, and the final probability of selection for the selected students, as well as the weight for these different levels where schools 2, 5, 7 and 10 have been selected.

**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**

School label	School size $N_i$	School probability $p_{1_i}$	School weight $w_{1_i}$	Within-school probability $p_{2_{ij}}$	Within-school weight $w_{2_{ij}}$	Final student probability $p_{ij}$	Final student weight $w_{ij}$	Sum of final weights $n_i w_{ij}$
1	40							
2	40	0.4	2.5	0.25	4	0.1	10	100
3	40							
4	40							
5	40	0.4	2.5	0.25	4	0.1	10	100
6	40							
7	40	0.4	2.5	0.25	4	0.1	10	100
8	40							
9	40							
10	40	0.4	2.5	0.25	4	0.1	10	100
<b>Total</b>			<b>10.0</b>					<b>400</b>

As shown in Table 3.3, the sum of the school weights corresponds to the number of schools in the population, *i.e.* 10, and the sum of the final student weights corresponds to the number of students in the population, *i.e.* 400.

In practice, schools differ in size. Often, school enrolment numbers tend to be larger in urban areas than rural areas. If schools are selected by simple, random sampling, the school selection probability will not change, but within the selected schools, the student selection probability will vary according to the school size. In a small school, the student selection probability will be large, while in a very large school, this probability will be small. Table 3.4 shows an example of the results obtained from schools of different enrolment sizes.

**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**

School label	School size	School probability	School weight	Within-school probability	Within-school weight	Final student probability	Final student weight	Sum of final weights
1	10							
2	15	0.4	2.5	0.66	1.5	0.27	3.75	37.5
3	20							
4	25							
5	30	0.4	2.5	0.33	3.0	0.13	7.50	75.0
6	35							
7	40	0.4	2.5	0.25	4.0	0.10	10.00	100.0
8	45							
9	80							
10	100	0.4	2.5	0.10	10.0	0.04	25.00	250.0
<b>Total</b>	<b>400</b>		<b>10.0</b>					<b>462.5</b>



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)

School label	School size	School probability	School weight	Within-school probability	Within-school weight	Final student probability	Final student weight	Sum of final weight
1	10	0.4	2.5	1.00	1.0	0.40	4.00	40.0
2	15	0.4	2.5	0.66	1.5	0.27	3.75	37.5
3	20	0.4	2.5	0.50	2.0	0.20	5.00	50.0
4	25	0.4	2.5	0.40	2.5	0.16	6.25	62.5
<b>Total</b>			<b>10.0</b>					<b>190.0</b>

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)

School label	School size	School probability	School weight	Within-school probability	Within-school weight	Final student probability	Final student weight	Sum of final weight
7	40	0.4	2.5	0.250	4.0	0.10	10.00	100.0
8	45	0.4	2.5	0.222	4.5	0.88	11.25	112.5
9	80	0.4	2.5	0.125	8.0	0.05	20.00	200.0
10	100	0.4	2.5	0.100	10.0	0.04	25.00	250.0
<b>Total</b>			<b>10.0</b>					<b>662.5</b>

With a simple, random sample of schools of unequal size, all schools have the same selection probability and the sum of school weights is equal to the number of schools in the population. However, the sum of the final student weights are not necessarily equal to the number of students in the population. Further, the final student weights differ among schools depending on the size of each school. This variability reduces the reliability of all population parameter estimates.

Table 3.5 and Table 3.6 present the different probabilities and weights if the four smallest schools or the four largest schools are selected. As shown in these two tables, the sums of final student weights vary substantially from the expected value of 400. The sum of school weights, however, is always equal to the number of schools in the population.

The focus of international education surveys such as PISA is more on the student sample than on the school sample. Many authors even consider that such studies do not draw a school sample *per se*. They just consider the school sample as an operational stage to draw the student sample. Therefore, a sampling design that consists of a simple random sample of schools is inappropriate as it would underestimate or overestimate the student population size. It would also result in an important variability of final student weights and consequently increase the sampling variance.

In order to avoid these disadvantages, schools are selected with probabilities proportional to their size (PPS). Larger schools will therefore have a higher probability of selection than smaller schools, but students in larger schools have a smaller within-school probability of being selected than students in small schools. With such procedures, the probability of a school to be selected is equal to the ratio of the school size multiplied by the number of schools to be sampled and divided by the total number of students in the population:

$$p_{1-i} = \frac{N_i * n_{sc}}{N}$$



The formulae for computing the within-school probabilities and weights remain unchanged. The final probability and weight are still the product of the school and within-school probabilities or weights. For instance, the school probability for school 9 is equal to:

$$p_{1\_9} = \frac{N_9 * n_{sc}}{N} = \frac{80 * 4}{400} = \frac{4}{5} = 0.8$$

The student within-school probability for school 9 is equal to:

$$p_{2\_9j} = \frac{n_9}{N_9} = \frac{10}{80} = 0.125$$

The final probability is equal to:

$$p_{9j} = 0.8 * 0.125 = 0.1$$

As shown in Table 3.7, the school and within-school weights differ among schools, but final student weights do not vary. The weights therefore do not increase sampling variability. Further, the sum of final student weights corresponds to the total number of students in the population. However, the sum of school weight differs from the expected value of ten, but this does not present a major problem as such educational surveys are primarily and mainly interested in the student sample.

**Table 3.7**  
School, within-school, and final probability of selection and corresponding weights  
for PPS sample of schools of unequal size

School label	School size	School probability	School weight	Within-school probability	Within-school weight	Final student probability	Final student weight	Sum of final weight
1	10							
2	15							
3	20	0.2	5.00	0.500	2.0	0.1	10	100
4	25							
5	30							
6	35							
7	40	0.4	2.50	0.250	4.0	0.1	10	100
8	45							
9	80	0.8	1.25	0.125	8.0	0.1	10	100
10	100	1.0	1.00	0.100	10.0	0.1	10	100
<b>Total</b>	<b>400</b>		<b>9.75</b>					<b>400</b>

With a PPS sample of schools, and an equal number of students selected in each selected school, the sum of the final student weights is always equal to the total number of students in the population (non-response being ignored at this stage). This will be the case even if the smallest or the largest schools get selected. The sum of the school weights, however, is not equal to the number of schools in the population. If the four smallest schools get selected, the sum of school weights is equal to 25.666. If the four largest schools get selected, the sum of school weights is equal to 6.97.

In order to keep the difference between the number of schools in the population and the sum of the school weights in the sample minimal, schools are selected according to a systematic procedure. The procedure consists of first sorting the schools according to their size. A sampling interval is computed as the ratio between the total number of students in the population and the number of schools in the sample, *i.e.*:

$$Int = \frac{N}{n_{sc}} = \frac{400}{4} = 100$$





A random number from a uniform distribution  $[0;1]$  is drawn. Let us say 0.752. This random number is then multiplied by the sampling interval, *i.e.* 0.752 by  $100 = 75.2$ . The school which contains the student number 76 is selected. Then the sampling interval is added to the value 75.2. The school which contains the student having the student number 176 will be selected. This systematic procedure is applied until the number of schools needed in the sample has been reached. In the example, the four selection numbers will be the following: 75.2, 175.2, 275.2 and 375.2. See Table 3.8.

**Table 3.8**  
**Selection of schools according to a PPS and systematic procedure**

School label	School size	From student number	To student number	Part of the sample
1	10	1	10	No
2	15	11	25	No
3	20	26	45	No
4	25	46	70	No
5	30	71	100	Yes
6	35	101	135	No
7	40	136	175	No
8	45	176	220	Yes
9	80	221	300	Yes
10	100	301	400	Yes

Sorting the school sampling frame by the measure of size and then using a systematic selection procedure prevents obtaining a sample of only small schools or (more likely) a sample with only large schools. This therefore reduces the sampling variance on the sum of the school weights, which is an estimate of the school population size.

### WHY DO THE PISA WEIGHTS VARY?

As demonstrated in the previous section, a two-stage sample design with a PPS sample of schools should guarantee that all students have the same probability of selection and therefore the same weight. However, the PISA data still needs to be weighted.

Different factors contribute to the variability of weights:

- Oversampling or undersampling of some strata of the population.** Usually, the school population is divided into different subgroups, called strata. For instance, a country might decide for convenience to separate the urban schools from the rural schools in the list of schools. In most cases, the number of students selected in the rural stratum and in the urban stratum will be proportional to what these two strata represent in the whole population. This stratification process guarantees for instance that a predefined number of schools within each stratum will be selected. Without the stratification, this number might vary. Nevertheless, for national reporting purposes, a country might decide to sample more students than what would have been sampled based on a proportional allocation in some part of the student population. Suppose that 90% of the student population in a country pursue academic tracks and 10% of the students pursue vocational tracks. If the national centre staff wants to compare the performance of the students by track, then it would be necessary to sample more vocational students than what would be sampled based on a proportional allocation. Further, since PISA 2003, the OECD offers countries the opportunity to adjudicate the data at a subnational level. This process however requires countries to sample at least 50 schools and 1 500 students per subnational entities. This requirement of course leads to some oversampling. Some subnational entities were separately adjudicated for Italy, Spain and the United-Kingdom in PISA 2003 and PISA 2006, and Belgium in PISA 2006.



- Lack of accuracy or no updated size measure for schools on the school sampling frame.** When schools are selected with a probability proportional to their size, a measure of size needs to be included in the school list. In PISA, this measure of size is the number of 15-year-olds in each school in the population, but national statistics per school and per date of birth are not always available. Therefore, the measure of size can be the number of students in the modal grade for 15-year-olds, or the total number of students in the school divided by the number of grades. Further, even if national statistics per school and per date of birth are available, these data might be one or two years old. Therefore, inconsistencies between the number of 15-year-olds at the time of testing and the measure of size used in the school sample frame generate some variability in the final student weights. Let us suppose that school 9 in Table 3.7 has 100 15-year-old students at the time of testing. When schools were selected from the list of schools, the measure of size was set at 80. The school weight was set at 1.25. The within-school weight will be equal to 100 divided by 80, *i.e.* 1.25 rather than 1. Therefore, the final student weight will be equal to 12.5 instead of the expected 10.
- School and within-school weight adjustment for school and student non-response.** Some schools, and within the selected and participating schools, some students might refuse to participate. To compensate for this non-response, a weight adjustment is applied at each level where non-response occurs. For instance, if only 25 students out of the 35 selected students from a participating school are present on the day of the assessment, then the weight of the participating students will be multiplied by a ratio of 35 by 25. The student participation rates vary from one school to another, and therefore the final student weights vary. A similar procedure is also applied to compensate for the school non-response. It should be noted that student non-response adjustment has been modified for counterbalancing different participation rates. More information about these adjustment factors is available in the PISA Technical Reports (Adams and Wu, 2000; OECD, 2005, forthcoming).

## CONCLUSION

This chapter briefly described: (i) what a weight is and how to compute it; (ii) what the PISA sampling design is and why such a design is considered the most appropriate; (iii) why the PISA final student weights show some variability.

All statistical analyses or procedures concerning the PISA data should be weighted. Unweighted analyses will provide biased population parameter estimates.

## Notes

1.  $N$  usually represents the size of the population and  $n$  the size of the sample.



# 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

<b>FOREWORD</b> .....	<b>3</b>
<b>USER'S GUIDE</b> .....	<b>17</b>
<b>CHAPTER 1 THE USEFULNESS OF PISA DATA FOR POLICY MAKERS, RESEARCHERS AND EXPERTS ON METHODOLOGY</b> .....	<b>19</b>
<b>PISA – an overview</b> .....	<b>20</b>
▪ The PISA surveys.....	20
<b>How can PISA contribute to educational policy, practice and research?</b> .....	<b>22</b>
▪ Key results from PISA 2000, PISA 2003 and PISA 2006.....	23
<b>Further analyses of PISA datasets</b> .....	<b>25</b>
▪ Contextual framework of PISA 2006.....	28
▪ Influence of the methodology on outcomes.....	31
<b>CHAPTER 2 EXPLORATORY ANALYSIS PROCEDURES</b> .....	<b>35</b>
<b>Introduction</b> .....	<b>36</b>
<b>Weights</b> .....	<b>36</b>
<b>Replicates for computing the standard error</b> .....	<b>39</b>
<b>Plausible values</b> .....	<b>43</b>
<b>Conclusion</b> .....	<b>46</b>
<b>CHAPTER 3 SAMPLE WEIGHTS</b> .....	<b>49</b>
<b>Introduction</b> .....	<b>50</b>
<b>Weights for simple random samples</b> .....	<b>51</b>
<b>Sampling designs for education surveys</b> .....	<b>53</b>
<b>Why do the PISA weights vary?</b> .....	<b>57</b>
<b>Conclusion</b> .....	<b>58</b>
<b>CHAPTER 4 REPLICATE WEIGHTS</b> .....	<b>59</b>
<b>Introduction</b> .....	<b>60</b>
<b>Sampling variance for simple random sampling</b> .....	<b>60</b>
<b>Sampling variance for two-stage sampling</b> .....	<b>65</b>
<b>Replication methods for simple random samples</b> .....	<b>70</b>
<b>Replication methods for two-stage samples</b> .....	<b>72</b>
▪ The Jackknife for unstratified two-stage sample designs.....	72
▪ The Jackknife for stratified two-stage sample designs.....	73
▪ The Balanced Repeated Replication method.....	74
<b>Other procedures for accounting for clustered samples</b> .....	<b>76</b>
<b>Conclusion</b> .....	<b>76</b>



<b>CHAPTER 5 THE RASCH MODEL</b> .....	<b>79</b>
<b>Introduction</b> .....	80
<b>How can the information be summarised?</b> .....	80
<b>The Rasch Model for dichotomous items</b> .....	81
▪ Introduction to the Rasch Model.....	81
▪ Item calibration.....	85
▪ Computation of a student's score.....	87
▪ Computation of a student's score for incomplete designs.....	91
▪ Optimal conditions for linking items.....	92
▪ Extension of the Rasch Model.....	93
<b>Other item response theory models</b> .....	94
<b>Conclusion</b> .....	94
 <b>CHAPTER 6 PLAUSIBLE VALUES</b> .....	 <b>95</b>
<b>Individual estimates versus population estimates</b> .....	96
<b>The meaning of plausible values (PVs)</b> .....	96
<b>Comparison of the efficiency of WLEs, EAP estimates and PVs for the estimation of some population statistics</b> .....	99
<b>How to perform analyses with plausible values</b> .....	102
<b>Conclusion</b> .....	103
 <b>CHAPTER 7 COMPUTATION OF STANDARD ERRORS</b> .....	 <b>105</b>
<b>Introduction</b> .....	106
<b>The standard error on univariate statistics for numerical variables</b> .....	106
<b>The SAS® macro for computing the standard error on a mean</b> .....	109
<b>The standard error on percentages</b> .....	112
<b>The standard error on regression coefficients</b> .....	115
<b>The standard error on correlation coefficients</b> .....	117
<b>Conclusion</b> .....	117
 <b>CHAPTER 8 ANALYSES WITH PLAUSIBLE VALUES</b> .....	 <b>119</b>
<b>Introduction</b> .....	120
<b>Univariate statistics on plausible values</b> .....	120
<b>The standard error on percentages with PVs</b> .....	123
<b>The standard error on regression coefficients with PVs</b> .....	123
<b>The standard error on correlation coefficients with PVs</b> .....	126
<b>Correlation between two sets of plausible values</b> .....	126
<b>A fatal error shortcut</b> .....	130
<b>An unbiased shortcut</b> .....	131
<b>Conclusion</b> .....	133
 <b>CHAPTER 9 USE OF PROFICIENCY LEVELS</b> .....	 <b>135</b>
<b>Introduction</b> .....	136
<b>Generation of the proficiency levels</b> .....	136
<b>Other analyses with proficiency levels</b> .....	141
<b>Conclusion</b> .....	143



<b>CHAPTER 10 ANALYSES WITH SCHOOL-LEVEL VARIABLES</b> .....	<b>145</b>
<b>Introduction</b> .....	146
<b>Limits of the PISA school samples</b> .....	147
<b>Merging the school and student data files</b> .....	148
<b>Analyses of the school variables</b> .....	148
<b>Conclusion</b> .....	150
<b>CHAPTER 11 STANDARD ERROR ON A DIFFERENCE</b> .....	<b>151</b>
<b>Introduction</b> .....	152
<b>Statistical issues and computing standard errors on differences</b> .....	152
<b>The standard error on a difference without plausible values</b> .....	154
<b>The standard error on a difference with plausible values</b> .....	159
<b>Multiple comparisons</b> .....	163
<b>Conclusion</b> .....	164
<b>CHAPTER 12 OECD TOTAL AND OECD AVERAGE</b> .....	<b>167</b>
<b>Introduction</b> .....	168
<b>Recoding of the database to estimate the pooled OECD total and the pooled OECD average</b> .....	170
<b>Duplication of the data to avoid running the procedure three times</b> .....	172
<b>Comparisons between the pooled OECD total or pooled OECD average estimates and a country estimate</b> .....	173
<b>Comparisons between the arithmetic OECD total or arithmetic OECD average estimates and a country estimate</b> .....	175
<b>Conclusion</b> .....	175
<b>CHAPTER 13 TRENDS</b> .....	<b>177</b>
<b>Introduction</b> .....	178
<b>The computation of the standard error for trend indicators on variables other than performance</b> .....	179
<b>The computation of the standard error for trend indicators on performance variables</b> .....	181
<b>Conclusion</b> .....	185
<b>CHAPTER 14 STUDYING THE RELATIONSHIP BETWEEN STUDENT PERFORMANCE AND INDICES DERIVED FROM CONTEXTUAL QUESTIONNAIRES</b> .....	<b>187</b>
<b>Introduction</b> .....	188
<b>Analyses by quarters</b> .....	188
<b>The concept of relative risk</b> .....	190
▪ <b>Instability of the relative risk</b> .....	191
▪ <b>Computation of the relative risk</b> .....	192
<b>Effect size</b> .....	195
<b>Linear regression and residual analysis</b> .....	197
▪ <b>Independence of errors</b> .....	197
<b>Statistical procedure</b> .....	200
<b>Conclusion</b> .....	201



<b>CHAPTER 15 MULTILEVEL ANALYSES</b> .....	<b>203</b>
<b>Introduction</b> .....	204
<b>Two-level modelling with SAS®</b> .....	206
▪ Decomposition of the variance in the empty model.....	206
▪ Models with only random intercepts.....	209
▪ Shrinkage factor.....	213
▪ Models with random intercepts and fixed slopes.....	213
▪ Models with random intercepts and random slopes.....	215
▪ Models with Level 2 independent variables.....	220
▪ Computation of final estimates and their respective standard errors.....	223
<b>Three-level modelling</b> .....	225
<b>Limitations of the multilevel model in the PISA context</b> .....	227
<b>Conclusion</b> .....	228
<b>CHAPTER 16 PISA AND POLICY RELEVANCE – THREE EXAMPLES OF ANALYSES</b> .....	<b>231</b>
<b>Introduction</b> .....	232
<b>Example 1: Gender differences in performance</b> .....	232
<b>Example 2: Promoting socio-economic diversity within school?</b> .....	236
<b>Example 3: The influence of an educational system on the expected occupational status of students at age 30</b> .....	242
<b>Conclusion</b> .....	246
<b>CHAPTER 17 SAS® MACRO</b> .....	<b>247</b>
<b>Introduction</b> .....	248
<b>Structure of the SAS® Macro</b> .....	248
<b>REFERENCES</b> .....	<b>313</b>
<b>APPENDICES</b> .....	<b>315</b>
<b>Appendix 1</b> Three-level regression analysis.....	316
<b>Appendix 2</b> PISA 2006 International database.....	324
<b>Appendix 3</b> PISA 2006 Student questionnaire.....	333
<b>Appendix 4</b> PISA 2006 Information communication technology (ICT) Questionnaire.....	342
<b>Appendix 5</b> PISA 2006 School questionnaire.....	344
<b>Appendix 6</b> PISA 2006 Parent questionnaire.....	351
<b>Appendix 7</b> Codebook for PISA 2006 student questionnaire data file.....	355
<b>Appendix 8</b> Codebook for PISA 2006 non-scored cognitive and embedded attitude items.....	399
<b>Appendix 9</b> Codebook for PISA 2006 scored cognitive and embedded attitude items.....	419
<b>Appendix 10</b> Codebook for PISA 2006 school questionnaire data file.....	431
<b>Appendix 11</b> Codebook for PISA 2006 parents questionnaire data file.....	442
<b>Appendix 12</b> PISA 2006 questionnaire indices.....	448





## LIST OF BOXES

Box 2.1	WEIGHT statement in the proc means procedure .....	37
<hr/>		
Box 7.1	SAS® syntax for computing 81 means (e.g. PISA 2003).....	106
Box 7.2	SAS® syntax for computing the mean of HISEI and its standard error (e.g. PISA 2003).....	109
Box 7.3	SAS® syntax for computing the standard deviation of HISEI and its standard error by gender (e.g. PISA 2003).....	112
Box 7.4	SAS® syntax for computing the percentages and their standard errors for gender (e.g. PISA 2003).....	112
Box 7.5	SAS® syntax for computing the percentages and its standard errors for grades by gender (e.g. PISA 2003).....	114
Box 7.6	SAS® syntax for computing regression coefficients, R <sup>2</sup> and its respective standard errors: Model 1 (e.g. PISA 2003).....	115
Box 7.7	SAS® syntax for computing regression coefficients, R <sup>2</sup> and its respective standard errors: Model 2 (e.g. PISA 2003).....	116
Box 7.8	SAS® syntax for computing correlation coefficients and its standard errors (e.g. PISA 2003).....	117
<hr/>		
Box 8.1	SAS® syntax for computing the mean on the science scale by using the PROC_MEANS_NO_PV macro (e.g. PISA 2006).....	121
Box 8.2	SAS® syntax for computing the mean and its standard error on PVs (e.g. PISA 2006).....	122
Box 8.3	SAS® syntax for computing the standard deviation and its standard error on PVs by gender (e.g. PISA 2006).....	123
Box 8.4	SAS® syntax for computing regression coefficients and their standard errors on PVs by using the PROC_REG_NO_PV macro (e.g. PISA 2006).....	124
Box 8.5	SAS® syntax for running the simple linear regression macro with PVs (e.g. PISA 2006).....	125
Box 8.6	SAS® syntax for running the correlation macro with PVs (e.g. PISA 2006).....	126
Box 8.7	SAS® syntax for the computation of the correlation between mathematics/quantity and mathematics/space and shape by using the PROC_CORR_NO_PV macro (e.g. PISA 2003).....	129
<hr/>		
Box 9.1	SAS® syntax for generating the proficiency levels in science (e.g. PISA 2006).....	137
Box 9.2	SAS® syntax for computing the percentages of students by proficiency level in science and its standard errors by using the PROC_FREQ_NO_PV macro (e.g. PISA 2006).....	138
Box 9.3	SAS® syntax for computing the percentage of students by proficiency level in science and its standard errors by using the PROC_FREQ_PV macro (e.g. PISA 2006).....	140
Box 9.4	SAS® syntax for computing the percentage of students by proficiency level and its standard errors by gender (e.g. PISA 2006).....	140
Box 9.5	SAS® syntax for generating the proficiency levels in mathematics (e.g. PISA 2003).....	141
Box 9.6	SAS® syntax for computing the mean of self-efficacy in mathematics and its standard errors by proficiency level (e.g. PISA 2003).....	142
<hr/>		
Box 10.1	SAS® syntax for merging the student and school data files (e.g. PISA 2006).....	148
Box 10.2	Question on school location in PISA 2006.....	149
Box 10.3	SAS® syntax for computing the percentage of students and the average performance in science, by school location (e.g. PISA 2006).....	149
<hr/>		
Box 11.1	SAS® syntax for computing the mean of job expectations by gender (e.g. PISA 2003).....	154
Box 11.2	SAS® macro for computing standard errors on differences (e.g. PISA 2003).....	157



Box 11.3	Alternative SAS® macro for computing the standard error on a difference for a dichotomous variable (e.g. PISA 2003).....	158
Box 11.4	SAS® syntax for computing standard errors on differences which involve PVs (e.g. PISA 2003).....	160
Box 11.5	SAS® syntax for computing standard errors on differences that involve PVs (e.g. PISA 2006).....	162
<hr/>		
Box 12.1	SAS® syntax for computing the pooled OECD total for the mathematics performance by gender (e.g. PISA 2003).....	170
Box 12.2	SAS® syntax for the pooled OECD average for the mathematics performance by gender (e.g. PISA 2003).....	171
Box 12.3	SAS® 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).....	172
<hr/>		
Box 14.1	SAS® syntax for the quarter analysis (e.g. PISA 2006).....	189
Box 14.2	SAS® syntax for computing the relative risk with five antecedent variables and five outcome variables (e.g. PISA 2006).....	193
Box 14.3	SAS® syntax for computing the relative risk with one antecedent variable and one outcome variable (e.g. PISA 2006).....	194
Box 14.4	SAS® syntax for computing the relative risk with one antecedent variable and five outcome variables (e.g. PISA 2006).....	194
Box 14.5	SAS® syntax for computing effect size (e.g. PISA 2006).....	196
Box 14.6	SAS® syntax for residual analyses (e.g. PISA 2003).....	200
<hr/>		
Box 15.1	Normalisation of the final student weights (e.g. PISA 2006).....	207
Box 15.2	SAS® syntax for the decomposition of the variance in student performance in science (e.g. PISA 2006).....	208
Box 15.3	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).....	211
Box 15.4	SAS® syntax for a multilevel regression model with random intercepts and fixed slopes (e.g. PISA 2006).....	214
Box 15.5	SAS® output for the multilevel model in Box 15.4.....	214
Box 15.6	SAS® syntax for a multilevel regression model (e.g. PISA 2006).....	216
Box 15.7	SAS® output for the multilevel model in Box 15.6.....	217
Box 15.8	SAS® output for the multilevel model with covariance between random parameters.....	218
Box 15.9	Interpretation of the within-school regression coefficient.....	220
Box 15.10	SAS® syntax for a multilevel regression model with a school-level variable (e.g. PISA 2006).....	221
Box 15.11	SAS® syntax for a multilevel regression model with interaction (e.g. PISA 2006).....	222
Box 15.12	SAS® output for the multilevel model in Box 15.11.....	222
Box 15.13	SAS® syntax for using the multilevel regression macro (e.g. PISA 2006).....	224
Box 15.14	SAS® syntax for normalising the weights for a three-level model (e.g. PISA 2006).....	226
<hr/>		
Box 16.1	SAS® syntax for testing the gender difference in standard deviations of reading performance (e.g. PISA 2000).....	233
Box 16.2	SAS® syntax for testing the gender difference in the 5th percentile of the reading performance (e.g. PISA 2006).....	235
Box 16.3	SAS® syntax for preparing a data file for the multilevel analysis.....	238



Box 16.4	SAS® syntax for running a preliminary multilevel analysis with one PV .....	239
Box 16.5	SAS® output for fixed parameters in the multilevel model.....	239
Box 16.6	SAS® syntax for running multilevel models with the PROC_MIXED_PV macro .....	242
<hr/>		
Box 17.1	SAS® macro of PROC_MEANS_NO_PV.sas.....	250
Box 17.2	SAS® macro of PROC_MEANS_PV.sas.....	253
Box 17.3	SAS® macro of PROC_FREQ_NO_PV.sas.....	256
Box 17.4	SAS® macro of PROC_FREQ_PV.sas.....	259
Box 17.5	SAS® macro of PROC_REG_NO_PV.sas.....	263
Box 17.6	SAS® macro of PROC_REG_PV.sas.....	266
Box 17.7	SAS® macro of PROC_CORR_NO_PV.sas.....	270
Box 17.8	SAS® macro of PROC_CORR_PV.sas.....	273
Box 17.9	SAS® macro of PROC_DIF_NO_PV.sas .....	276
Box 17.10	SAS® macro of PROC_DIF_PV.sas .....	279
Box 17.11	SAS® macro of QUARTILE_PV.sas .....	282
Box 17.12	SAS® macro of RELATIVE_RISK_NO_PV.sas.....	288
Box 17.13	SAS® macro of RELATIVE_RISK_PV.sas.....	291
Box 17.14	SAS® macro of EFFECT_SIZE_NO_PV.sas .....	296
Box 17.15	SAS® macro of EFFECT_SIZE_PV.sas.....	298
Box 17.16	SAS® macro of PROC_MIXED_NO_PV.sas.....	301
Box 17.17	SAS® macro of PROC_MIXED_PV.sas .....	306
<hr/>		
Box A1.1	Descriptive statistics of background and explanatory variables.....	318
Box A1.2	Background model for student performance.....	319
Box A1.3	Final net combined model for student performance.....	320
Box A1.4	Background model for the impact of socio-economic background.....	321
Box A1.5	Model of the impact of socio-economic background: “school resources” module .....	322
Box A1.6	Model of the impact of socio-economic background: “accountability practices” module .....	323
Box A1.7	Final combined model for the impact of socio-economic background.....	323

## 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
<hr/>		
Figure 2.1	Science mean performance in OECD countries (PISA 2006).....	38
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).....	39
Figure 2.4	Design effect on the country mean estimates for science performance and for ESCS in OECD countries (PISA 2006) .....	42
Figure 2.5	Simple random sample and unbiased standard errors of ESCS on science performance in OECD countries (PISA 2006) .....	43



Figure 4.1	Distribution of the results of 36 students.....	60
Figure 4.2	Sampling variance distribution of the mean.....	62
Figure 5.1	Probability of success for two high jumpers by height (dichotomous).....	82
Figure 5.2	Probability of success for two high jumpers by height (continuous).....	83
Figure 5.3	Probability of success to an item of difficulty zero as a function of student ability.....	83
Figure 5.4	Student score and item difficulty distributions on a Rasch continuum.....	86
Figure 5.5	Response pattern probabilities for the response pattern (1, 1, 0, 0).....	88
Figure 5.6	Response pattern probabilities for a raw score of 1.....	89
Figure 5.7	Response pattern probabilities for a raw score of 2.....	90
Figure 5.8	Response pattern probabilities for a raw score of 3.....	90
Figure 5.9	Response pattern likelihood for an easy test and a difficult test.....	91
Figure 5.10	Rasch item anchoring.....	92
Figure 6.1	Living room length expressed in integers.....	96
Figure 6.2	Real length per reported length.....	97
Figure 6.3	A posterior distribution on a test of six items.....	98
Figure 6.4	EAP estimators.....	99
Figure 8.1	A two-dimensional distribution.....	127
Figure 8.2	Axes for two-dimensional normal distributions.....	127
Figure 13.1	Trend indicators in PISA 2000, PISA 2003 and PISA 2006.....	179
Figure 14.1	Percentage of schools by three school groups (PISA 2003).....	198
Figure 15.1	Simple linear regression analysis versus multilevel regression analysis.....	205
Figure 15.2	Graphical representation of the between-school variance reduction.....	215
Figure 15.3	A random multilevel model.....	216
Figure 15.4	Change in the between-school residual variance for a fixed and a random model.....	218
Figure 16.1	Relationship between the segregation index of students' expected occupational status and the segregation index of student performance in reading (PISA 2000).....	244
Figure 16.2	Relationship between the segregation index of students' expected occupational status and the correlation between HISEI and students' expected occupational status.....	245

**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.....	41
Table 2.2	Mean estimates and standard errors.....	45



Table 2.3	Standard deviation estimates and standard errors.....	45
Table 2.4	Correlation estimates and standard errors.....	45
Table 2.5	ESCS regression coefficient estimates and standard errors.....	46
<hr/>		
Table 3.1	Height and weight of ten persons .....	52
Table 3.2	Weighted and unweighted standard deviation estimate .....	52
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.....	54
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 .....	54
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) .....	55
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) .....	55
Table 3.7	School, within-school, and final probability of selection and corresponding weights for PPS sample of schools of unequal size .....	56
Table 3.8	Selection of schools according to a PPS and systematic procedure.....	57
<hr/>		
Table 4.1	Description of the 630 possible samples of 2 students selected from 36 students, according to their mean.....	61
Table 4.2	Distribution of all possible samples with a mean between 8.32 and 11.68.....	63
Table 4.3	Distribution of the mean of all possible samples of 4 students out of a population of 36 students.....	64
Table 4.4	Between-school and within-school variances on the mathematics scale in PISA 2003.....	67
Table 4.5	Current status of sampling errors.....	67
Table 4.6	Between-school and within-school variances, number of participating schools and students in Denmark and Germany in PISA 2003 .....	68
Table 4.7	The Jackknives replicates and sample means.....	70
Table 4.8	Values on variables X and Y for a sample of ten students.....	71
Table 4.9	Regression coefficients for each replicate sample.....	71
Table 4.10	The Jackknife replicates for unstratified two-stage sample designs.....	72
Table 4.11	The Jackknife replicates for stratified two-stage sample designs.....	73
Table 4.12	Replicates with the Balanced Repeated Replication method.....	74
Table 4.13	The Fay replicates .....	75
<hr/>		
Table 5.1	Probability of success when student ability equals item difficulty.....	84
Table 5.2	Probability of success when student ability is less than the item difficulty by 1 unit.....	84
Table 5.3	Probability of success when student ability is greater than the item difficulty by 1 unit .....	84
Table 5.4	Probability of success when student ability is less than the item difficulty by 2 units .....	85
Table 5.5	Probability of success when student ability is greater than the item difficulty by 2 units.....	85
Table 5.6	Possible response pattern for a test of four items.....	87
Table 5.7	Probability for the response pattern (1, 1, 0, 0) for three student abilities.....	87
Table 5.8	Probability for the response pattern (1, 0) for two students of different ability in an incomplete test design.....	91
Table 5.9	PISA 2003 test design .....	93

Table 6.1	Structure of the simulated data.....	100
Table 6.2	Means and variances for the latent variables and the different student ability estimators.....	100
Table 6.3	Percentiles for the latent variables and the different student ability estimators.....	101
Table 6.4	Correlation between HISEI, gender and the latent variable, the different student ability estimators.....	101
Table 6.5	Between- and within-school variances.....	102
<hr/>		
Table 7.1	HISEI mean estimates.....	107
Table 7.2	Squared differences between replicate estimates and the final estimate.....	108
Table 7.3	Output data file exercise1 from Box 7.2.....	111
Table 7.4	Available statistics with the PROC_MEANS_NO_PV macro.....	111
Table 7.5	Output data file exercise2 from Box 7.3.....	112
Table 7.6	Output data file exercise3 from Box 7.4.....	112
Table 7.7	Percentage of girls for the final and replicate weights and squared differences.....	113
Table 7.8	Output data file exercise4 from Box 7.5.....	114
Table 7.9	Output data file exercise5 from Box 7.6.....	115
Table 7.10	Output data file exercise6 from Box 7.7.....	116
Table 7.11	Output data file exercise6_criteria from Box 7.7.....	117
Table 7.12	Output data file exercise7 from Box 7.8.....	117
<hr/>		
Table 8.1	The 405 mean estimates.....	120
Table 8.2	Mean estimates and their respective sampling variances on the science scale for Belgium (PISA 2006).....	121
Table 8.3	Output data file exercise6 from Box 8.2.....	123
Table 8.4	Output data file exercise7 from Box 8.3.....	123
Table 8.5	The 450 regression coefficient estimates.....	125
Table 8.6	HISEI regression coefficient estimates and their respective sampling variance on the science scale in Belgium after accounting for gender (PISA 2006).....	125
Table 8.7	Output data file exercise8 from Box 8.5.....	125
Table 8.8	Output data file exercise9 from Box 8.6.....	126
Table 8.9	Correlation between the five plausible values for each domain, mathematics/quantity and mathematics/space and shape.....	128
Table 8.10	The five correlation estimates between mathematics/quantity and mathematics/space and shape and their respective sampling variance.....	129
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).....	131
Table 8.12	Unbiased shortcut for a population estimate and its standard error.....	132
Table 8.13	Standard errors from the full and shortcut computation (PISA 2006).....	132
<hr/>		
Table 9.1	The 405 percentage estimates for a particular proficiency level.....	138
Table 9.2	Estimates and sampling variances per proficiency level in science for Germany (PISA 2006).....	139
Table 9.3	Final estimates of the percentage of students, per proficiency level, in science and its standard errors for Germany (PISA 2006).....	139
Table 9.4	Output data file exercise6 from Box 9.3.....	140
Table 9.5	Output data file exercise7 from Box 9.4.....	140
Table 9.6	Mean estimates and standard errors for self-efficacy in mathematics per proficiency level (PISA 2003).....	143
Table 9.7	Output data file exercise8 from Box 9.6.....	143



Table 10.1	Percentage of students per grade and ISCED level, by country (PISA 2006) .....	146
Table 10.2	Output data file exercise1 from Box 10.3 .....	150
Table 10.3	Output data file exercise2 from Box 10.3 .....	150
Table 11.1	Output data file exercise1 from Box 11.1 .....	155
Table 11.2	Mean estimates for the final and 80 replicate weights by gender (PISA 2003) .....	155
Table 11.3	Difference in estimates for the final weight and 80 replicate weights between females and males (PISA 2003) .....	157
Table 11.4	Output data file exercise2 from Box 11.2 .....	158
Table 11.5	Output data file exercise3 from Box 11.3 .....	159
Table 11.6	Gender difference estimates and their respective sampling variances on the mathematics scale (PISA 2003) .....	159
Table 11.7	Output data file exercise4 from Box 11.4 .....	160
Table 11.8	Gender differences on the mathematics scale, unbiased standard errors and biased standard errors (PISA 2003) .....	161
Table 11.9	Gender differences in mean science performance and in standard deviation for science performance (PISA 2006) .....	161
Table 11.10	Regression coefficient of HISEI on the science performance for different models (PISA 2006) .....	163
Table 11.11	Cross tabulation of the different probabilities .....	163
Table 12.1	Regression coefficients of the index of instrumental motivation in mathematics on mathematic performance in OECD countries (PISA 2003) .....	169
Table 12.2	Output data file exercise1 from Box 12.1 .....	170
Table 12.3	Output data file exercise2 from Box 12.2 .....	171
Table 12.4	Difference between the country mean scores in mathematics and the OECD total and average (PISA 2003) .....	174
Table 13.1	Trend indicators between PISA 2000 and PISA 2003 for HISEI, by country .....	180
Table 13.2	Linking error estimates .....	182
Table 13.3	Mean performance in reading by gender in Germany .....	184
Table 14.1	Distribution of the questionnaire index of cultural possession at home in Luxembourg (PISA 2006) .....	188
Table 14.2	Output data file exercise1 from Box 14.1 .....	190
Table 14.3	Labels used in a two-way table .....	190
Table 14.4	Distribution of 100 students by parents' marital status and grade repetition .....	191
Table 14.5	Probabilities by parents' marital status and grade repetition .....	191
Table 14.6	Relative risk for different cutpoints .....	191
Table 14.7	Output data file exercise2 from Box 14.2 .....	193
Table 14.8	Mean and standard deviation for the student performance in reading by gender, gender difference and effect size (PISA 2006) .....	195
Table 14.9	Output data file exercise4 from Box 14.5 .....	197
Table 14.10	Output data file exercise5 from 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) .....	199

Table 15.1	Between- and within-school variance estimates and intraclass correlation (PISA 2006).....	209
Table 15.2	Output data file “ranparm1” from Box 15.3.....	212
Table 15.3	Output data file “fixparm3” from Box 15.6.....	217
Table 15.4	Output data file “ranparm3” from Box 15.6.....	217
Table 15.5	Variance/covariance estimates before and after centering.....	219
Table 15.6	Output data file of the fixed parameters file.....	221
Table 15.7	Average performance and percentage of students by student immigrant status and by type of school.....	223
Table 15.8	Variables for the four groups of students.....	223
Table 15.9	Comparison of the regression coefficient estimates and their standard errors in Belgium (PISA 2006).....	224
Table 15.10	Comparison of the variance estimates and their respective standard errors in Belgium (PISA 2006).....	225
Table 15.11	Three-level regression analyses.....	226
Table 16.1	Differences between males and females in the standard deviation of student performance (PISA 2000).....	234
Table 16.2	Distribution of the gender differences (males – females) in the standard deviation of the student performance.....	234
Table 16.3	Gender difference on the PISA combined reading scale for the 5 <sup>th</sup> , 10 <sup>th</sup> , 90 <sup>th</sup> and 95 <sup>th</sup> percentiles (PISA 2000).....	235
Table 16.4	Gender difference in the standard deviation for the two different item format scales in reading (PISA 2000).....	236
Table 16.5	Random and fixed parameters in the multilevel model with student and school socio-economic background.....	237
Table 16.6	Random and fixed parameters in the multilevel model with socio-economic background and grade retention at the student and school levels.....	241
Table 16.7	Segregation indices and correlation coefficients by country (PISA 2000).....	243
Table 16.8	Segregation indices and correlation coefficients by country (PISA 2006).....	244
Table 16.9	Country correlations (PISA 2000).....	245
Table 16.10	Country correlations (PISA 2006).....	246
Table 17.1	Synthesis of the 17 SAS® macros.....	249
Table A2.1	Cluster rotation design used to form test booklets for PISA 2006.....	324
Table A12.1	Mapping of ISCED to accumulated years of education.....	449
Table A12.2	ISCO major group white-collar/blue-collar classification.....	451
Table A12.3	ISCO occupation categories classified as science-related occupations.....	451
Table A12.4	Household possessions and home background indices.....	455
Table A12.5	Factor loadings and internal consistency of ESCS 2006 in OECD countries.....	465
Table A12.6	Factor loadings and internal consistency of ESCS 2006 in partner countries/economies.....	466





# 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](http://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

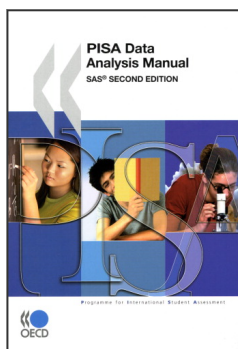
All syntaxes and macros in this manual can be copied from the PISA website ([www.pisa.oecd.org](http://www.pisa.oecd.org)). The 17 SAS® macros presented in Chapter 17 need to be saved under "c:\pisa\macro\", before starting analysis. Each chapter of the manual contains a complete set of syntaxes, which 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: SAS, Second Edition**

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

**Please cite this chapter as:**

OECD (2009), "Sample Weights", in *PISA Data Analysis Manual: SAS, Second Edition*, OECD Publishing, Paris.

DOI: <https://doi.org/10.1787/9789264056251-4-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](mailto: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](mailto:info@copyright.com) or the Centre français d'exploitation du droit de copie (CFC) at [contact@cfcopies.com](mailto:contact@cfcopies.com).