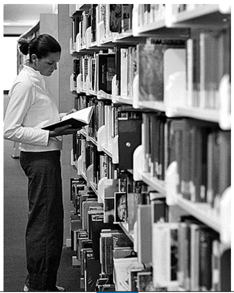
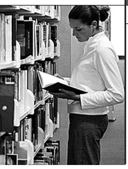


8

Survey Weighting and the Calculation of Sampling Variance



Survey weights were required to analyse PISA data, to calculate appropriate estimates of sampling error and to make valid estimates and inferences. The consortium calculated survey weights for all assessed, ineligible and excluded students, and provided variables in the data that permit users to make approximately unbiased estimates of standard errors, to conduct significance tests and to create confidence intervals appropriately, given the sample design for PISA in each individual country.

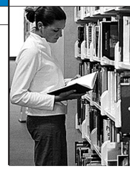
SURVEY WEIGHTING

While the students included in the final PISA sample for a given country were chosen randomly, the selection probabilities of the students vary. Survey weights must therefore be incorporated into the analysis to ensure that each sampled student represents the correct number of students in the full PISA population.

There are several reasons why the survey weights are not the same for all students in a given country:

- A school sample design may intentionally over- or under-sample certain sectors of the school population: in the former case, so that they could be effectively analysed separately for national purposes, such as a relatively small but politically important province or region, or a sub-population using a particular language of instruction; and in the latter case, for reasons of cost, or other practical considerations, such as very small or geographically remote schools;[1]
- Information about school size available at the time of sampling may not have been completely accurate. If a school was expected to be very large, the selection probability was based on the assumption that only a sample of its students would be selected for PISA. But if the school turned out to be quite small, all students would have to be included and would have, overall, a higher probability of selection in the sample than planned, making these inclusion probabilities higher than those of most other students in the sample. Conversely, if a school thought to be small turned out to be large, the students included in the sample would have had smaller selection probabilities than others;
- School non-response, where no replacement school participated, may have occurred, leading to the under-representation of students from that kind of school, unless weighting adjustments were made. It is also possible that only part of the eligible population in a school (such as those 15-year-olds in a particular grade) were represented by its student sample, which also requires weighting to compensate for the missing data from the omitted grades;
- Student non-response, within participating schools, occurred to varying extents. Sampled students who were eligible and not excluded, but did not participate in the assessment will be under-represented in the data unless weighting adjustments are made;
- Trimming weights to prevent undue influence of a relatively small subset of the school or student sample might have been necessary if a small group of students would otherwise have much larger weights than the remaining students in the country. Such large sampling weights can lead to unstable estimates – large sampling errors – but cannot be well estimated. Trimming weights introduces a small bias into estimates but greatly reduces standard errors (Kish, 1992).

The procedures used to derive the survey weights for PISA reflect the standards of best practice for analysing complex survey data, and the procedures used by the world's major statistical agencies. The same procedures were used in other international studies of educational achievement: the Trends in International Mathematics and Science Study (TIMSS) and the Progress in International Reading Literacy Studies (PIRLS), which were all implemented by the International Association for the Evaluation of Educational Achievement (IEA). See Cochran, (1977), Lohr (1999) and Särndal, Swensson and Wretman (1992) for the underlying statistical theory for the analysis of survey data.



The weight, W_{ij}, for student j in school i consists of two base weights – the school and the within-school – and five adjustment factors, and can be expressed as:

8.1

$$W_{ij} = t_{2ij} f_{1i} f_{2ij} f_{1ij}^{A} t_{1i} w_{2ij} w_{1i}$$

Where:

w_{1i} is the school base weight, is given as the reciprocal of the probability of inclusion of school i into the sample;

w_{2ij} is the within-school base weight, is given as the reciprocal of the probability of selection of student j from within the selected school i;

f_{1i} is an adjustment factor to compensate for non-participation by other schools that are somewhat similar in nature to school i (not already compensated for by the participation of replacement schools);

f_{1ij}^{A} is an adjustment factor to compensate for the fact that, in some countries, in some schools only 15-year-old students who were enrolled in the modal grade for 15-year-olds were included in the assessment;

f_{2ij} is an adjustment factor to compensate for non-participation by students within the same school non-response cell and explicit stratum, and, where permitted by the sample size, within the same high/low grade and gender categories;

t_{1i} is a school trimming factor, used to reduce unexpectedly large values of w_{1i}; and

t_{2ij}, is a student trimming factor, used to reduce the weights of students with exceptionally large values for the product of all the preceding weight components.

THE SCHOOL BASE WEIGHT

The term w_{1i} is referred to as the school base weight. For the systematic probability proportional- to-size school sampling method used in PISA, this is given as:

8.2

$$w_{1i} = \begin{cases} \dfrac{int(g/i)}{mos(i)} & \text{if } mos(i) < int(g/i) \\ 1 & \text{otherwise} \end{cases}$$

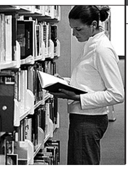
The term *mos* (*i*) denotes the measure of size given to each school on the sampling frame.

Despite country variations, *mos* (*i*) was usually equal to the estimated number of 15-year-olds in the school, if it was greater than the predetermined Target Cluster Size (*TCS*) (35 in most countries).

If the enrolment of 15-year-olds was less than the *TCS*, then *mos* (*i*) = *TCS*.

The term *int*(*g/i*) denotes the sampling interval used within the explicit sampling stratum *g* that contains school *i* and is calculated as the total of the *mos*(*i*) values for all schools in stratum *g*, divided by the school sample size for that stratum.

Thus, if school *i* was estimated to have 100 15-year-olds at the time of sample selection, *mos* (*i*) = 100. If the country had a single explicit stratum (*g*=1) and the total of the *mos*(*i*) values over all schools was 150 000, with a school sample size of 150, then int(1/*i*) = 150000/150 = 1000, for school *i* (and others in the sample),



giving w_{1i} = 1000/100 = 10.0. Roughly speaking, the school can be thought of as representing about 10 schools from the population. In this example, any school with 1 000 or more 15-year-old students would be included in the sample with certainty, with a base weight of $w_{1i} = 1$.

The school weight trimming factor

Once school base weights were established for each sampled school in the country, verifications were made separately within each explicit sampling stratum to see if the school weights required trimming. The school trimming factor t_{1i}, is the ratio of the trimmed to the untrimmed school base weight, and is equal to 1.0000 for most schools and therefore most students, and never exceeds this value.

The school-level trimming adjustment was applied to schools that turned out to be much larger than was believed at the time of sampling – where 15-year-old enrolment exceeded $3 \times \max(TCS, mos(i))$. For example, if *TCS* = 35, then a school flagged for trimming had more than 105 PISA-eligible students, and more than three times as many students as was indicated on the school sampling frame. Because the student sample size was set at *TCS* regardless of the actual enrolment, the student sampling rate was much lower than anticipated during the school sampling. This meant that the weights for the sampled students in these schools would have been more than three times greater than anticipated when the school sample was selected. These schools had their school base weights trimmed by having *mos*(*i*) replaced by $3 \times \max(TCS, mos(i))$ in the school base weight formula.

The student base weight

The term w_{2ij} is referred to as the student base weight. With the PISA procedure for sampling students, w_{2ij} did not vary across students (*j*) within a particular school *i*. This is given as:

8.3

$$w_{2ij} = \frac{enr(i)}{sam(i)}$$

where *enr*(*i*) is the actual enrolment of 15-year-olds in the school (and so, in general, is somewhat different from the estimated *mos*(*i*)), and *sam*(*i*) is the sample size within school *i*. It follows that if all students from the school were selected, then $w_{2ij} = 1$ for all eligible students in the school. For all other cases $w_{2ij} > 1$.

In the case of the international grade sampling option, for direct sampled grade students, the sampling interval for the extra grade students was the same as that for the PISA students. Therefore, countries with extra direct sampled grade students (the Czech Republic, Korea, Mexico, Norway, Sweden, certain explicit strata in Switzerland and Uruguay) have the same within school student weights for the extra grade students as those for PISA students from the same school.

Additional weight components were needed for the grade students in Chile, Germany, Liechtenstein, Mexico and certain strata in Switzerland. For these first four countries, the extra weight component consisted of the class weight for the selected class(es) (All students were selected into the grade sample in the selected class(es).) For Mexico, the extra weight component at the school level accounted for the sub-sampling of schools in which the grade sample would be implemented. In these five countries, the extra weight component resulted in the necessity of a second weighting stream for the extra grade students.

School non-response adjustment

In order to adjust for the fact that those schools that declined to participate, and were not replaced by a replacement school, were not in general typical of the schools in the sample as a whole, school-level non-response adjustments were made. Several groups of somewhat similar schools were formed within a country, and within each group the weights of the responding schools were adjusted to compensate for the missing schools and their students.

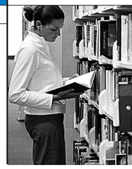
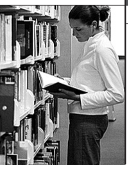


Table 8.1
Non-response classes

		Implicit stratification variables used to create school non-response cells (within explicit stratum) and number of original and final cells	Number of original cells	Number of final cells
OECD	Australia	Geographic Zone (8); School Level for TAS and ACT Government Schools (3)	88	78
	Austria	Province-District (121)	114	27
	Belgium	Flanders -- Index of Overaged students; French Community -- Public/Private School Types for Special Education Schools (4); Index of Overaged students for Regular Schools; German Community – None	154	38
	Canada	Public/Private(2); Urban/Rural(2)	107	51
	Czech Republic	Region for Programmes 3. 4. 5. 6 (14)	151	38
	Denmark	School Type (5); Geo Area (5)	37	16
	Finland	None	33	13
	France	None	20	11
	Germany	School Type for Normal Schools (5); State for Other Schools (16)	63	27
	Greece	School Type (3); Public/Private (2) when both in an explicit stratum	40	18
	Hungary	Region (7); National Grade 10 math Score Categories (5) for Non-Primary Schools	105	24
	Iceland	Urban/Rural (2); School Size (4)	33	24
	Ireland	School Type (3); School Gender Composition Categories (5)	31	13
	Italy	Public/Private (2)	107	54
	Japan	Levels of proportion of students taking University/College Entrance Exams(4)	16	12
	Korea	School Level (2)	12	8
	Luxembourg	None	16	4
	Mexico	School Size (3); Public/Private (2); Urban/Rural (2); School Level (2); School Program (4 For Each School Level)	343	107
	Netherlands	School Type (6 for School Track A and 3 for School Track B)	9	5
	New Zealand	Public/Private (2); Socio-Economic Status Category (3) and Urban/Rural (2) for Public Schools	7	6
	Norway	None	12	8
	Poland	Urbanicity (4)	11	8
	Portugal	Public/Private (2); Socio-Economic Status Category (4)	50	15
	Slovak Republic	Programme (9); Language (2) in 4 of the Regions	60	16
	Spain	2 or 3 digits of Postal Code for adjudicated regions	323	84
	Sweden	Urbanicity (5) for Private Lower Secondary schools; Public/private (2) for Upper Secondary schools; Administrative Province (25) for Upper Secondary schools; Income Quartiles (4) for all except Private Lower Secondary schools	55	23
	Switzerland	School Type (28); Canton (26)	186	52
	Turkey	School Level (3); Public/Private (2); Urban/Rural (2)	39	12
	United Kingdom	England: School Type (6). GCSE Performance (6). Region (4). Local Authority Northern Ireland: School Type (3). Education and Library Board Region (5) Scotland: None Wales: School Type (2). Region (3). Local Authority	252	65
	United States	Public/Private (2); Region (4); Urbanicity (3); Minority Status (2); Grade Span (4); Postal Code	79	15
Partners	Argentina	Sector (2); School Type (5); School Level (5)	96	25
	Azerbaijan	Urbanicity (4); Education Department (5)	108	9
	Brazil	School Type (3); HDI Category (2); School Size (3); Urban/Rural (2); Capital/Non-Capital (2)	355	124
	Bulgaria	School Type (3); Settlement Size (5); State/Municipal (2); Public/Private (2)	94	79
	Chile	Urban/Rural (2); Region (13)	114	29
	Columbia	Urban/Rural (2); Public/Private(2)	4	3
	Croatia	County (21)	110	21
	Estonia	Urbanicity (4); School Type (4); County (15)	67	18
	Hong Kong-China	Student Academic Intake (4)	10	7
	Indonesia	School Type (5); Public/Private (2); National Achievement Score Categories (3)	225	62
	Israel	Location (9) for Public Schools. Except For Schools in Druz Migzar Sector; Group Size (5) for Regular Public Schools; Gender Composition (3) for Religious Public Schools; Migzar Sector (3) for Regular Public Arabic Schools; Cultivation Categories (4) for Public Jewish Schools; Cultivation (Continuous Measure) in All	61	31
	Jordan	Urbanicity (2); School Gender Composition (3); School form (2)	28	16
	Kyrgyzstan	School Type (5)	60	18
	Latvia	Urbanicity (4); School Type (4)	15	8
	Liechtenstein	Urbanicity (3); Public/Private(2)	2	2
	Lithuania	None	12	8
	Macao-China	Secondary Levels (3)	14	3
	Montenegro	Region (3) for Primary Schools; Urban/Rural (2); School Type (3)	14	10
	Qatar	Qatari/Non-Qatari (2)	26	18
	Romania	Language (3); Urbanicity (2)	13	6
	Russian Federation	Urbanicity (9); School Type (4); School Sub-Type (16)	194	94
	Serbia	Urban/Rural (2); School Type (7)	77	19
	Slovenia	Urbanicity (4)	24	18
	Chinese Taipei	Public/Private (2)	60	30
	Thailand	Local Area (9)	57	22
	Tunisia	Category of Grade Repeating (3) for General Public Schools; East/West (2) for Private Schools and Vocational Schools; North/South (2) for all	39	13
	Uruguay	Area (4); Shift (4) for Public Secondary Schools; Shift (4) for Public Technical Schools	65	40



The compositions of the non-response groups varied from country to country, but were based on cross-classifying the explicit and implicit stratification variables used at the time of school sample selection. Usually, about 10 to 15 such groups were formed within a given country depending upon school distribution with respect to stratification variables. If a country provided no implicit stratification variables, schools were divided into three roughly equal groups, within each stratum, based on their enrolment size. It was desirable to ensure that each group had at least six participating schools, as small groups can lead to unstable weight adjustments, which in turn would inflate the sampling variances. However, it was not necessary to collapse cells where all schools participated, as the school non-response adjustment factor was 1.0 regardless of whether cells were collapsed or not. Adjustments greater than 2.0 were flagged for review, as they can cause increased variability in the weights and lead to an increase in sampling variances. In either of these situations, cells were generally collapsed over the last implicit stratification variable(s) until the violations no longer existed. In countries with very high overall levels of school non-response after school replacement, the requirement for school non-response adjustment factors all to be below 2.0 was waived.

Within the school non-response adjustment group containing school i, the non-response adjustment factor was calculated as:

8.4

$$f_{1i} = \frac{\sum\limits_{k \in \Omega(i)} w_{1k} enr(k)}{\sum\limits_{k \in \Gamma(i)} w_{1k} enr(k)}$$

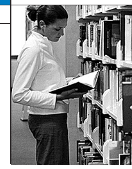
where the sum in the denominator is over $\Gamma(i)$ the schools within the group (originals and replacements) that participated, while the sum in the numerator is over $\Omega(i)$, those same schools, plus the original sample schools that refused and were not replaced. The numerator estimates the population of 15-year-olds in the group, while the denominator gives the size of the population of 15-year-olds directly represented by participating schools. The school non-response adjustment factor ensures that participating schools are weighted to represent all students in the group. If a school did not participate because it had no eligible students enrolled, no adjustment was necessary since this was neither non-response nor under-coverage.

Table 8.1 shows the number of school non-response classes that were formed for each country, and the variables that were used to create the cells.

Grade non-response adjustment

Because of perceived administrative inconvenience, individual schools may occasionally agree to participate in PISA but require that participation be restricted to 15-year-olds in the modal grade for 15-year-olds, rather than all 15-year-olds. Since the modal grade generally includes the majority of the population to be covered, such schools may be accepted as participants. For the part of the 15-year-old population in the modal grade, these schools are respondents, while for the rest of the grades in the school with 15-year-olds, such a school is a refusal. To account for this, a special non-response adjustment can be calculated at the school level for students not in the modal grade (and is automatically 1.0 for all students in the modal grade). No countries had this type of non-response for PISA 2006, so the weight adjustment for grade non-response was automatically 1.0 for all students in both the modal and non-modal grades, and therefore did not affect the final weights.

If the weight adjustment for grade non-response had been needed (as it was in earlier cycles of PISA in a few countries), it would have been calculated as follows:



Within the same non-response adjustment groups used for creating school non-response adjustment factors, the grade non-response adjustment factor for all students in school i, f_{1i}^{A}, is given as:

8.5

$$f_{1i}^{A} = \begin{cases} \sum_{k \in C(i)} w_{1k}\, enra(k) \Big/ \sum_{k \in B(i)} w_{1k}\, enra(k) \\ 1 \end{cases}$$

The variable $enra(k)$ is the approximate number of 15-year-old students in school k but not in the modal grade. The set B(i) is all schools that participated for all eligible grades (from within the non-response adjustment group with school (i)), while the set C(i) includes these schools and those that only participated for the modal responding grade.

This procedure gives, for each school, a single grade non-response adjustment factor that depends upon its non-response adjustment class. Each individual student has this factor applied to the weight if he/she did not belong to the modal grade, and 1.0000 if belonging to the modal grade. In general, this factor is not the same for all students within the same school when a country has some grade non-response.

Student non-response adjustment

Within each final school non-response adjustment cell, explicit stratum and high/low grade, gender, and school combination, the student non-response adjustment f_{2i} was calculated as:

8.6

$$f_{2i} = \frac{\sum_{k \in X(i)} f_{1i} w_{1i} w_{2ik}}{\sum_{k \in \Delta(i)} f_{1i} w_{1i} w_{2ik}}$$

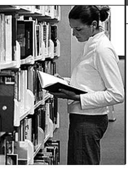
Where

$\Delta(i)$ is all assessed students in the final school non-response adjustment cell and explicit stratum-grade-gender-school combination; and,

$X(i)$ is all assessed students in the final school non-response adjustment cell and explicit stratum-grade-gender-school combination plus all others who should have been assessed (*i.e.* who were absent, but not excluded or ineligible).

The high and low grade categories in each country were defined so that each contain a substantial proportion of the PISA population in each explicit stratum of larger schools.

The definition was then applied to all schools of the same explicit stratum characteristics but regardless of school size. In most cases, this student non-response factor reduces to the ratio of the number of students who should have been assessed to the number who were assessed. In some cases of small cells (*i.e.* final school non-response adjustment cell and explicit stratum/grade/gender/school category combinations) sizes (fewer than 15 respondents), it was necessary to collapse cells together, then apply the more complex formula shown above. Additionally, an adjustment factor greater than 2.0 was not allowed for the same reasons noted under school non-response adjustments. If this occurred, the cell with the large adjustment was collapsed with the closest cell within grade and gender combinations in the same school non-response cell and explicit stratum. Note that the calculation of the high / low grades, the use of gender, and the order of cell collapsing represent differences from the student non-response adjustment strategy used for PISA 2003.



Some schools in some countries had very low student response levels. In these cases it was determined that the small sample of assessed students was potentially too biased as a representation of the school to be included in the PISA data. For any school where the student response rate was below 25%, the school was therefore treated as a non-respondent, and its student data were removed. In schools with between 25 and 50% student response, the student non-response adjustment described above would have resulted in an adjustment factor of between 2.0000 and 4.0000, and so the grade / gender cells of these schools were collapsed with others to create student non-response adjustments[2].

For countries with extra direct grade sampled students (the Czech Republic, Korea, Mexico, Norway, Sweden, certain explicit strata in Switzerland and Uruguay), care was taken to ensure that student non-response cells were formed separately for PISA students and the extra non-PISA grade students. No procedural changes were needed for Chile, Germany, Liechtenstein, Mexico and certain strata in Switzerland since a separate weighting stream was needed for the grade students.

As noted above, a few beneficial changes were introduced to the 2006 strategy for student non-response adjustments: namely the calculation of the high/low grade categories within explicit strata rather than over the whole set of schools, the use of gender in forming the student non-response cells, and the removal of the school as the basis for the final cell formation. As a result of this latter change, the final weights for students within schools could vary.

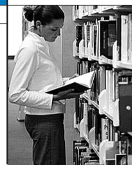
Trimming student weights

This final trimming check was used to detect student records that were unusually large compared to those of other students within the same explicit stratum. The sample design was intended to give all students from within the same explicit stratum an equal probability of selection and therefore equal weight, in the absence of school and student non-response. As already noted, poor prior information about the number of eligible students in each school could lead to substantial violations of this principle. Moreover, school, grade, and student non-response adjustments, and, occasionally, inappropriate student sampling could, in a few cases, accumulate to give a few students in the data relatively very large weights, which adds considerably to sampling variance. The weights of individual students were therefore reviewed, and where the weight was more than four times the median weight of students from the same explicit sampling stratum, it was trimmed to be equal to four times the median weight for that explicit stratum.

The student trimming factor, t_{2ij} , is equal to the ratio of the final student weight to the student weight adjusted for student non-response, and therefore equal to 1.0000 for the great majority of students. The final weight variable on the data file was called *w_fstuwt*, which is the final student weight that incorporates any student-level trimming. As in PISA 2000 and PISA 2003, little trimming was required at either the school or the student levels.

Comparing the PISA 2006 student non-response adjustment strategy with the strategy used for PISA 2003

The student non-response adjustment section of this chapter noted that changes had been made to the 2006 student non-response adjustment strategy. While the changes were thought to be beneficial because they used more student information in the adjustments, an assessment of the impact of the change was nevertheless conducted. This section is devoted to that investigation, which compares the 2006 student non-response adjustment strategy to the 2003 non-response adjustment strategy for countries that also participated in PISA 2003.



Recall that the final student weight consists of:

- The school base weight;
- The school weight trimming factor;
- The school non response adjustment;
- The student base weight;
- The student non response adjustment;,
- The student weight trimming factor;

as well as potentially the grade non-response adjustment factor if needed (but this was not needed for 2006).

The student non-response adjustment is designed to reduce the potential bias introduced by the non-participating students. As described in the *PISA 2000 Technical Report* (Adams & Wu, 2002), the student non- response adjustment factor was computed in PISA 2000 as follows:

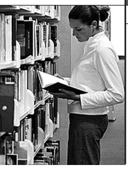
- Within each participating school, the student non-response adjustment was equal to the sum of all sampled student initial weights divided by the sum of the participating student initial weights;
- If the school sample had fewer than 15 participating students, the school was collapsed with the next school within the same school non-response cell;
- If the adjustment factor was greater than 2, the school was collapsed with the next school within the same school non-response cell.

Secondary analyses of the PISA 2000 student tracking forms have revealed substantial differential non-response rates in some countries (Monseur, 2005) countries. For instance, the response rates of Portuguese students attending grade 7 to grade 10 were respectively equal to 0.76, 0.80, 0.87 and 0.88. As grade highly correlates with performance, these differential response rates introduced a bias.

In 2003, it was therefore decided to include the grade attended by the student in the computation of the student non-response adjustment. Concretely:

- Grades were grouped into two categories: lower grades and higher grades so that each had a substantial proportion of students;
- Within each participating school and high/low grade combination, the student non-response adjustment was calculated;
- If the combination of school and high/low grade cells had fewer than 10 participating students or the corresponding adjustment factor was greater than 2, the two initial student non-response cells were collapsed within that school;
- If this collapsing within a particular school did not allow satisfying the two criteria (*i.e.* a minimum of 10 students and an adjustment factor lower than 2), further collapsing was done as in PISA 2000.

This procedure, however, had a limited impact as in most countries, the within school sample size was equal to 35 students. Therefore, the requirement of 10 participating students per student non-response cell was not reached in a large number of schools, so that the two non-response cells (lower versus higher grade cells) were collapsed.



Previous analyses (Monseur, 2005) had also shown a differential participation rate for boys and girls. For instance, in Portugal, the student response rate for boys was 82.6% and for girls 87.8%. As gender also correlated with performance, particularly in reading literacy, the student non-response adjustment developed for PISA 2006 aimed to better compensate for differential grade and gender response rates.

As described above the student non-response adjustment was computed in PISA 2006 as follows:

For each school, four student non response cells were created:

- Higher grades/girls;
- Higher grades/boys;
- Lower grades/girls;,
- Lower grade/boys,

where the high/low grades were derived within each explicit stratum.

In single sex schools or in schools with students enrolled in only one grade, only two student non-response cells were created.

The major change between the previous procedures and the PISA 2006 procedure is not the addition of the gender variable for creating the non- response cell but the ordering of the collapsing.

In PISA 2003, non-response cells were first collapsed within school, then, if required, schools were collapsed.

In 2006, a non-response cell from a school was first collapsed with a non-response cell sharing the same gender and grade but from another school. However, these two schools had to be in the same school non-response cell and explicit stratum. If further collapsing was required, usually non-response cells were collapsed across gender and then across grade.

As this modification in the computation of the student non-response adjustment might have an impact on population estimates, in particular on performance estimates, it was decided to compute the 2006 data the student non response adjustment according to (i) the new algorithm and (ii) the PISA 2003 algorithm for only the countries that participated in the 2003 and 2006 surveys. Comparing population estimates for the two sets of weights allows measuring the impact of the weighting modification.

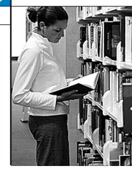
The comparison

Three sets of weights have been used in the subsequent analyses:

- The initial student weight that consists of:
  a. The school initial base weight;
  b. The school trimming factor;
  c. The school non-response adjustment factor;
  d. The student initial within school weight;
- The final student weight based on the 2003 non response adjustment method; and,
- The final student weight based on the 2006 non response adjustment method.

For the second and third sets of weights, only participating students were included in the analyses while absent and participating students were included for the first set of weights.

Each set of weights can be used with the gender and grade data collected through the student tracking form.



In several countries, the difference between the males' response rate and the females' response rate was greater than 2%. Even with these response rate differences, the 2006 method weighted estimates were equal or very close to the estimates computed using the initial student weights and data from the student tracking form, while the 2003 method weighted estimates differed to a greater extent. The 2006 adjustment method appears more efficient in reducing a potential bias due to the differential participation rates between males and females (as expected since gender was not used in the PISA 2003 strategy).

Regarding grade, there were also several countries where the difference between the initial weighted estimate and the 2003 method adjusted estimate was at least 1%. In almost all cases, the 2006 method adjusted weights reduced the differences when compared to the initial weighted estimates.

Finally, looking at the three literacy scales, in all countries, the differences in mean performance between the two sets of weight results was always less than two PISA scale points and for most countries the difference was less than one scale point.

Country comparisons are not provided since all differences were small.

In summary, the new method of student non-response adjustment used in 2006 does not appear to have generated any spurious changes in achievement means of any consequence.

CALCULATING SAMPLING VARIANCE

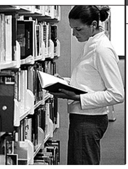
To estimate the sampling variances of PISA parameter estimates, a replication methodology was employed. This reflected the variance in estimates due to the sampling of schools and students. Additional variance due to the use of plausible values from the posterior distributions of scaled scores was captured separately, although computationally the calculation of the two components can be carried out in a single program, such as *WesVar 5.1* (Westat, 2007).

The balanced repeated replication variance estimator

The approach used for calculating sampling variances for PISA is known as balanced repeated replication (BRR), or balanced half- samples; the particular variant known as Fay's method was used. This method is very similar in nature to the jackknife method used in other international studies of educational achievement, such as TIMSS, and it is well documented in the survey sampling literature (*see* Rust, 1985; Rust and Rao, 1996; Shao, 1996; Wolter, 2007). The major advantage of BRR over the jackknife is that the jackknife method is not fully appropriate for use with non-differentiable functions of the survey data, most noticeably quantiles, for which it does not provide a statistically consistent estimator of variance. This means that, depending upon the sample design, the variance estimator can be very unstable, and despite empirical evidence that it can behave well in a PISA-like design, theory is lacking. In contrast BRR does not have this theoretical flaw. The standard BRR procedure can become unstable when used to analyse sparse population subgroups, but Fay's modification overcomes this difficulty, and is well justified in the literature (Judkins, 1990).

The BRR approach was implemented as follows, for a country where the student sample was selected from a sample of schools, rather than all schools:

- Schools were paired on the basis of the explicit and implicit stratification and frame ordering used in sampling. The pairs were originally sampled schools, with each participating replacement taking the place of the original school that it replaced. For an odd number of schools within a stratum, a triple was formed consisting of the last three schools on the sorted list;
- Pairs were numbered sequentially, 1 to H, with pair number denoted by the subscript h. Other studies and the literature refer to such pairs as variance strata or zones, or pseudo-strata;



- Within each variance stratum, one school (the Primary Sampling Unit, PSU) was randomly numbered as 1, the other as 2 (and the third as 3, in a triple), which defined the variance unit of the school. Subscript *j* refers to this numbering;
- These variance strata and variance units (1, 2, 3) assigned at school level are attached to the data for the sampled students within the corresponding school;
- Let the estimate of a given statistic from the full student sample be denoted as X^*. This is calculated using the full sample weights;
- A set of 80 replicate estimates, X_t^* (where *t* runs from 1 to 80), was created. Each of these replicate estimates was formed by multiplying the sampling weights from one of the two PSUs in each stratum by 1.5, and the weights from the remaining PSUs by 0.5. The determination as to which PSUs received inflated weights, and which received deflated weights, was carried out in a systematic fashion, based on the entries in a Hadamard matrix of order 80. A Hadamard matrix contains entries that are +1 and –1 in value, and has the property that the matrix, multiplied by its transpose, gives the identity matrix of order 80, multiplied by a factor of 80. Details concerning Hadamard matrices are given in Wolter (2007);
- In cases where there were three units in a triple, either one of the schools (designated at random) received a factor of 1.7071 for a given replicate, with the other two schools receiving factors of 0.6464, or else the one school received a factor of 0.2929 and the other two schools received factors of 1.3536. The explanation of how these particular factors came to be used is explained in Appendix 12 of the PISA 2000 Technical Report (Adams & Wu, 2002);
- To use a Hadamard matrix of order 80 requires that there be no more than 80 variance strata within a country, or else that some combining of variance strata be carried out prior to assigning the replication factors via the Hadamard matrix. The combining of variance strata does not cause any bias in variance estimation, provided that it is carried out in such a way that the assignment of variance units is independent from one stratum to another within strata that are combined. That is, the assignment of variance units must be completed before the combining of variance strata takes place, and this approach was used for PISA;
- The reliability of variance estimates for important population subgroups is enhanced if any combining of variance strata that is required is conducted by combining variance strata from different subgroups. Thus in PISA, variance strata that were combined were selected from different explicit sampling strata and, to the extent possible, from different implicit sampling strata also;
- In some countries, it was not the case that the entire sample was a two-stage design, of first sampling schools and then sampling students. In some countries for part of the sample (and for the entire samples for Iceland, Liechtenstein, Luxembourg, Macao - China, and Qatar), schools were included with certainty into the sampling, so that only a single stage of student sampling was carried out for this part of the sample. In these cases instead of pairing schools, pairs of individual students were formed from within the same school (and if the school had an odd number of sampled students, a triple of students was formed). The procedure of assigning variance units and replicate weight factors was then conducted at the student level, rather than at the school level;
- In contrast, in one country, the Russian Federation, there was a stage of sampling that preceded the selection of schools. Then the procedure for assigning variance strata, variance units and replicate factors was applied at this higher level of sampling. The schools and students then inherited the assignment from the higher-level unit in which they were located;

- Procedural changes were in general not needed in the formation of variance strata for countries with extra direct grade sampled students (the Czech Republic, Korea, Mexico, Norway, Sweden, certain explicit strata in Switzerland and Uruguay) since the extra grade sample came from the same schools as the PISA students. However, if there were certainty schools in these countries, students within the certainty schools were paired so that PISA non-grade students were together, PISA grade students were together and non-PISA grade students were together. No procedural changes were required for the grade students for Chile, Germany, Liechtenstein, certain strata in Switzerland, and Mexico, since a separate weighting stream was needed in these cases;
- The variance estimator is then:

8.7

$$V_{BRR}(X^*) = 0.05 \sum_{t=1}^{80} \left\{ (X_t^* - X^*)^2 \right\}$$

The properties of BRR have been established by demonstrating that it is unbiased and consistent for simple linear estimators (*i.e.* means from straightforward sample designs), and that it has desirable asymptotic consistency for a wide variety of estimators under complex designs, and through empirical simulation studies.

Reflecting weighting adjustments

This description glosses over one aspect of the implementation of the BRR method. Weights for a given replicate are obtained by applying the adjustment to the weight components that reflect selection probabilities (the school base weight in most cases), and then re-computing the non-response adjustment replicate by replicate.

Implementing this approach required that the consortium produce a set of replicate weights in addition to the full sample weight. Eighty such replicate weights were needed for each student in the data file. The school and student non-response adjustments had to be repeated for each set of replicate weights.

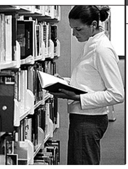
To estimate sampling errors correctly, the analyst must use the variance estimation formula above, by deriving estimates using the *t*-th set of replicate weights. Because of the weight adjustments (and the presence of occasional triples), this does not mean merely increasing the final full sample weights for half the schools by a factor of 1.5 and decreasing the weights from the remaining schools by a factor of 0.5. Many replicate weights will also be slightly disturbed, beyond these adjustments, as a result of repeating the non-response adjustments separately by replicate.

Formation of variance strata

With the approach described above, all original sampled schools were sorted in stratum order (including refusals, excluded and ineligible schools) and paired. An alternative would have been to pair participating schools only. However, the approach used permits the variance estimator to reflect the impact of non-response adjustments on sampling variance, which the alternative does not. This is unlikely to be a big component of variance in any PISA country, but the procedure gives a more accurate estimate of sampling variance.

Countries where all students were selected for PISA

In Iceland, Liechtenstein, Luxembourg, and Qatar, all eligible students were selected for PISA. It might be considered surprising that the PISA data should reflect any sampling variance in these countries, but students have been assigned to variance strata and variance units, and the BRR formula does give a positive



estimate of sampling variance for two reasons. First, in each country there was some student non-response, and, in the case of Iceland and Qatar, some school non-response. Not all eligible students were assessed, giving sampling variance. Second, the intent is to make inference about educational systems and not particular groups of individual students, so it is appropriate that a part of the sampling variance reflect random variation between student populations, even if they were to be subjected to identical educational experiences. This is consistent with the approach that is generally used whenever survey data are used to try to make direct or indirect inference about some underlying system.

Notes

1. Note that this is not the same as excluding certain portions of the school population. This also happened in some cases, but cannot be addressed adequately through the use of survey weights.

2. Chapter 12 describes these schools as being treated as non-respondents for the purpose of response rate calculation, even though their student data were used in the analyses.

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

Reader's Guide

Country codes – the following country codes are used in this report:

OECD countries

AUS	Australia
AUT	Austria
BEL	Belgium
BEF	Belgium (French Community)
BEN	Belgium (Flemish Community)
CAN	Canada
CAE	Canada (English Community)
CAF	Canada (French Community)
CZE	Czech Republic
DNK	Denmark
FIN	Finland
FRA	France
DEU	Germany
GRC	Greece
HUN	Hungary
ISL	Iceland
IRL	Ireland
ITA	Italy
JPN	Japan
KOR	Korea
LUX	Luxembourg
LXF	Luxembourg (French Community)
LXG	Luxembourg (German Community)
MEX	Mexico
NLD	Netherlands
NZL	New Zealand
NOR	Norway
POL	Poland
PRT	Portugal
SVK	Slovak Republic
ESP	Spain
ESB	Spain (Basque Community)
ESC	Spain (Catalonian Community)
ESS	Spain (Castillian Community)
SWE	Sweden
CHE	Switzerland
CHF	Switzerland (French Community)
CHG	Switzerland (German Community)
CHI	Switzerland (Italian Community)
TUR	Turkey
GBR	United Kingdom
IRL	Ireland
SCO	Scotland
USA	United States

Partner countries and economies

ARG	Argentina
AZE	Azerbaijan
BGR	Bulgaria
BRA	Brazil
CHL	Chile
COL	Colombia
EST	Estonia
HKG	Hong Kong-China
HRV	Croatia
IDN	Indonesia
JOR	Jordan
KGZ	Kyrgyztan
LIE	Liechtenstein
LTU	Lithuania
LVA	Latvia
LVL	Latvia (Latvian Community)
LVR	Latvia (Russian Community)
MAC	Macao-China
MNE	Montenegro
QAT	Qatar
ROU	Romania
RUS	Russian Federation
SRB	Serbia
SVN	Slovenia
TAP	Chinese Taipei
THA	Thailand
TUN	Tunisia
URY	Uruguay

References

Adams, R.J., Wilson, M. & **Wang, W.C.** (1997), The multidimensional random coefficients multinomial logit model. *Applied Psychological Measurement*, No. 21, pp. 1-23.

Adams, R.J., Wilson, M. R. & **Wu, M.L.** (1997), Multilevel item response models: An approach to errors in variables regression, *Journal of Educational and Behavioural Statistics,* No. 22 (1), pp. 46-75.

Adams, R.J. & **Wu, M.L.** (2002), *PISA 2000 Technical Report,* OECD, Paris.

Bollen, K.A. & **Long, S.J.** (1993) (eds.), *Testing Structural Equation Models*, Newbury Park: London.

Beaton, A.E. (1987), Implementing the new design: The NAEP 1983-84 technical report (Rep. No. 15-TR-20), Princeton, NJ: Educational Testing Service.

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

Buchmann, C. (2002), Measuring Family Background in International Studies of Education: Conceptual Issues and Methodological Challenges, in Porter, A.C. and Gamoran, A. (eds.). *Methodological Advances in Cross-National Surveys of Educational Achievement* (pp. 150-97), Washington, DC: National Academy Press.

Creemers, B.P.M. (1994), *The Effective Classroom*, London: Cassell.

Cochran, W.G. (1977), *Sampling techniques,* third edition, New York, NY: John Wiley and Sons.

Ganzeboom, H.B.G., de Graaf, P.M. & **Treiman, D.J.** (1992), A standard international socio-economic index of occupational status, *Social Science Research*, No. 21, pp. 1-56.

Ganzeboom H.B. & **Treiman, D.J.** (1996), Internationally comparable measures of occupational status for the 1988 international standard classification of occupations, *Social Science Research*, No. 25, pp. 201-239.

Grisay, A. (2003), Translation procedures in OECD/PISA 2000 international assessment, *Language Testing,* No. 20 (2), pp. 225-240.

Hambleton, R.K., Swaminathan, H. & **Rogers, H.J.** (1991), *Fundamentals of item response theory*, Newbury Park, London, New Delhi: SAGE Publications.

Hambleton, R.K., Merenda, P.F. & **Spielberger, C.D.** (2005), *Adapting Educational and Psychological Tests for Cross-Cultural Assessment,* IEA Lawrence Erlbaum Associates, Publishers, Mahwah, New Jersey.

Harkness, J.A., Van de Vijver, F.J.R. & **Mohler, P.Ph** (2003), *Cross-Cultural Survey Methods,* Wiley-Interscience, John Wiley & Sons, Inc., Hoboken, New Jersey.

Harvey-Beavis, A. (2002), Student and School Questionnaire Development, in R.J. Adams and M.L. Wu (eds.), *PISA 2000 Technical Report,* (pp. 33-38), OECD, Paris.

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

Jöreskog, K.G. & **Sörbom, Dag** (1993), *LISREL 8 User's Reference Guide*, Chicago: SSI.

Judkins, D.R. (1990), Fay's Method of Variance Estimation, *Journal of Official Statistics*, No. 6 (3), pp. 223-239.

Kaplan, D. (2000), *Structural equation modeling: Foundation and extensions*, Thousand Oaks: SAGE Publications.

Keyfitz, N. (1951), Sampling with probabilities proportionate to science: Adjustment for changes in probabilities, *Journal of the American Statistical Association,* No. 46, American Statistical Association, Alexandria, pp. 105-109.

Kish, L. (1992), Weighting for Unequal, *Pi. Journal of Official Statistics*, No. 8 (2), pp. 183-200.

LISREL (1993), K.G. Jöreskog & D. Sörbom, [computer software], Lincolnwood, IL: Scientific Software International, Inc.

Lohr, S.L. (1999), *Sampling: Design and Analysis*, Duxberry: Pacific Grove.

Macaskill, G., Adams, R.J. & **Wu, M.L.** (1998), Scaling methodology and procedures for the mathematics and science literacy, advanced mathematics and physics scale, in M. Martin and D.L. Kelly, Editors, *Third International Mathematics and Science Study, technical report Volume 3: Implementation and analysis*, Boston College, Chestnut Hill, MA.

Masters, G.N. & **Wright, B.D.** (1997), The Partial Credit Model, in W.J. van der Linden, & R.K. Hambleton (eds.), *Handbook of Modern Item Response Theory* (pp. 101-122), New York/Berlin/Heidelberg: Springer.

Mislevy, R.J. (1991), Randomization-based inference about latent variables from complex samples, *Psychometrika,* No. 56, pp. 177-196.

Mislevy, R.J., Beaton, A., Kaplan, B.A. & **Sheehan, K.** (1992), Estimating population characteristics from sparse matrix samples of item responses, *Journal of Educational Measurement,* No. 29 (2), pp. 133-161.

Mislevy, R.J. & **Sheehan, K.M.** (1987), Marginal estimation procedures, in Beaton, A.E., Editor, 1987. *The NAEP 1983-84 technical report*, National Assessment of Educational Progress, Educational Testing Service, Princeton, pp. 293-360.

Mislevy, R.J. & **Sheehan, K.M.** (1989), Information matrices in latent-variable models, *Journal of Educational Statistics*, No. 14, pp. 335-350.

Mislevy, R.J. & **Sheehan, K.M.** (1989), The role of collateral information about examinees in item parameter estimation, *Psychometrika*, No. 54, pp. 661-679.

Monseur, C. & **Berezner, A.** (2007), The Computation of Equating Errors in International Surveys in Education, *Journal of Applied Measurement,* No. 8 (3), 2007, pp. 323-335.

Monseur, C. (2005), An exploratory alternative approach for student non response weight adjustment, *Studies in Educational Evaluation*, No. 31 (2-3), pp. 129-144.

Muthen, B. & **L. Muthen** (1998), [computer software], *Mplus* Los Angeles, CA: Muthen & Muthen.

Muthen, B., du Toit, S.H.C. & **Spisic, D.** (1997), *Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes*, unpublished manuscript.

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

OECD (2003), *Literacy Skills for the World of Tomorrow: Further results from PISA 2000*, OECD, Paris.

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

OECD (2005), *Technical Report for the OECD Programme for International Student Assessment 2003*, 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.

PISA Consortium (2006), *PISA 2006 Main Study Data Management Manual*, *https://mypisa.acer.edu.au/images/mypisadoc/opmanual/pisa2006_data_management_manual.pdf*

Rasch, G. (1960), Probabilistic models for some intelligence and attainment tests, Copenhagen: Nielsen & Lydiche.

Routitski A. & **Berezner, A.** (2006), Issues influencing the validity of cross-national comparisons of student performance. Data Entry Quality and Parameter Estimation. Paper presented at the Annual Meeting of the American Educational Research Association (AERA) in San Francisco, 7-11 April, *https://mypisa.acer.edu.au/images/mypisadoc/aera06routitsky_berezner.pdf*

Rust, K. (1985), Variance Estimation for Complex Estimators in Sample Surveys, *Journal of Official Statistics*, No. 1, pp. 381-397.

Rust, K.F. & **Rao, J.N.K.** (1996), Variance Estimation for Complex Surveys Using Replication Techniques, *Survey Methods in Medical Research*, No. 5, pp. 283-310.

Shao, J. (1996), Resampling Methods in Sample Surveys (with Discussion), *Statistics*, No. 27, pp. 203-254.

Särndal, C.-E., Swensson, B. & **Wretman, J.** (1992), *Model Assisted Survey Sampling*, New York: Springer-Verlag.

SAS® CALIS (1992), W. Hartmann [computer software], Cary, NC: SAS Institute Inc.

Scheerens, J. (1990), School effectiveness and the development of process indicators of school functioning, *School effectiveness and school improvement,* No. 1, pp. 61-80.

Scheerens, J. & **Bosker, R.J.** (1997), *The Foundations of School Effectiveness*, Oxford: Pergamon.

Schulz, W. (2002), Constructing and Validating the Questionnaire composites, in R.J. Adams and M.L. Wu (eds.), *PISA 2000 Technical Report*, OECD, Paris.

Schulz, W. (2004), Mapping Student Scores to Item Responses, in W. Schulz and H. Sibberns (eds.), *IEA Civic Education Study, Technical Report* (pp. 127-132), Amsterdam: IEA.

Schulz, W. (2006a), *Testing Parameter Invariance for Questionnaire Indices using Confirmatory Factor Analysis and Item Response Theory,* Paper presented at the Annual Meetings of the American Educational Research Association (AERA) in San Francisco, 7-11 April.

Schulz, W. (2006b), *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.

Thorndike, R.L. (1973), *Reading comprehension in fifteen countries,* New York, Wiley: and Stockholm: Almqvist & Wiksell.

Travers, K.J. & **Westbury, I.** (1989), *The IEA Study of Mathematics I: Analysis of Mathematics Curricula*, Oxford: Pergamon Press.

Travers, K.J., Garden R.A. & **Rosier, M.** (1989), Introduction to the Study, in Robitaille, D. A. and Garden, R. A. (eds), *The IEA Study of Mathematics II: Contexts and Outcomes of School Mathematics Curricula,* Oxford: Pergamon Press.

Verhelst, N. (2002), Coder and Marker Reliabilaity Studies, in R.J. Adams & M.L. Wu (eds.), *PISA 2000 Technical Report*. OECD, Paris.

Walberg, H.J. (1984), Improving the productivity of American schools, *Educational Leadership,* No. 41, pp. 19-27.

Walberg, H. (1986), Synthesis of research on teaching, in M. Wittrock (ed.), *Handbook of research on teaching* (pp. 214-229), New York: Macmillan.

Walker, M. (2006), *The choice of Likert or dichotomous items to measure attitudes across culturally distinct countries in international comparative educational research*. Paper presented at the Annual Meetings of the American Educational Research Association (AERA) in San Francisco, 7-11 April.

Walker, M. (2007), Ameliorating Culturally-Based Extreme Response Tendencies To Attitude items, *Journal of Applied Measurement,* No. 8, pp. 267-278.

Warm, T.A. (1989), Weighted Likelihood Estimation of Ability in Item Response Theory, *Psychometrika*, No. 54 (3), pp. 427-450.

Westat (2007), *WesVar® 5.1* Computer software and manual, Rockville, MD: Author (also see *http://www.westat.com/wesvar/*).

Wilson, M. (1994), Comparing Attitude Across Different Cultures: Two Quantitative Approaches to Construct Validity, in M. Wilson (ed.), *Objective measurement II: Theory into practice* (pp. 271-292), Norwood, NJ: Ablex.

Wolter, K.M. (2007), *Introduction to Variance Estimation*. Second edition, Springer: New York.

Wu, M.L., Adams, R.J. & **Wilson, M.R.** (1997), *ConQuest®: Multi-Aspect Test Software* [computer program manual], Camberwell, Vic.: Australian Council for Educational Research.

List of abbreviations – the following abbreviations are used in this report:

ACER	Australian Council for Educational Research
AGFI	Adjusted Goodness-of-Fit Index
BRR	Balanced Repeated Replication
CBAS	Computer Based Assessment of Science
CFA	Confirmatory Factor Analysis
CFI	Comparative Fit Index
CITO	National Institute for Educational Measurement, The Netherlands
CIVED	Civic Education Study
DIF	Differential Item Functioning
ENR	Enrolment of 15-year-olds
ESCS	PISA Index of Economic, Social and Cultural Status
ETS	Educational Testing Service
IAEP	International Assessment of Educational Progress
I	Sampling Interval
ICR	Inter-Country Coder Reliability Study
ICT	Information Communication Technology
IEA	International Association for the Evaluation of Educational Achievement
INES	OECD Indicators of Education Systems
IRT	Item Response Theory
ISCED	International Standard Classification of Education
ISCO	International Standard Classification of Occupations
ISEI	International Socio-Economic Index
MENR	Enrolment for moderately small school
MOS	Measure of size
NCQM	National Centre Quality Monitor
NDP	National Desired Population
NEP	National Enrolled Population
NFI	Normed Fit Index
NIER	National Institute for Educational Research, Japan
NNFI	Non-Normed Fit Index
NPM	National Project Manager
OECD	Organisation for Economic Cooperation and Development
PISA	Programme for International Student Assessment
PPS	Probability Proportional to Size
PGB	PISA Governing Board
PQM	PISA Quality Monitor
PSU	Primary Sampling Units
QAS	Questionnaire Adaptations Spreadsheet
RMSEA	Root Mean Square Error of Approximation
RN	Random Number
SC	School Co-ordinator
SE	Standard Error
SD	Standard Deviation
SEM	Structural Equation Modelling
SMEG	Subject Matter Expert Group
SPT	Study Programme Table
TA	Test Administrator
TAG	Technical Advisory Group
TCS	Target Cluster Size
TIMSS	Third International Mathematics and Science Study
TIMSS-R	Third International Mathematics and Science Study – Repeat
VENR	Enrolment for very small schools
WLE	Weighted Likelihood Estimates

Table of contents

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 –

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

LIST OF BOXES

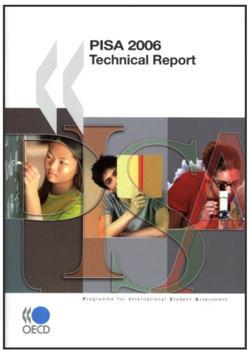
LIST OF FIGURES

LIST OF TABLES

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009



From:
PISA 2006 Technical Report

Access the complete publication at:
https://doi.org/10.1787/9789264048096-en

Please cite this chapter as:

OECD (2009), “Survey Weighting and the Calculation of Sampling Variance”, in *PISA 2006 Technical Report*, OECD Publishing, Paris.

DOI: https://doi.org/10.1787/9789264048096-9-en