



---

4

# Sample Design

<b>Target population and overview of the sampling design</b> .....	58
<b>Population coverage, and school and student participation rate standards</b> .....	58
<b>Main study school sample</b> .....	62

## TARGET POPULATION AND OVERVIEW OF THE SAMPLING DESIGN

The desired base PISA target population in each country consisted of 15-year-old students attending educational institutions in grades 7 and higher. This meant that countries were to include:

- 15-year-olds enrolled full-time in educational institutions;
- 15-year-olds enrolled in educational institutions who attended only on a part-time basis;
- students in vocational training programmes, or any other related type of educational programmes; and
- students attending foreign schools within the country (as well as students from other countries attending any of the programmes in the first three categories).

It was recognised that no testing of 15-year-olds schooled in the home, workplace or out of the country would occur and therefore these 15-year-olds were not included in the international target population.

The operational definition of an age population directly depends on the testing dates. The international requirement was that the assessment had to be conducted during a 42-day period, referred to as the testing period, between 1 March 2009 and 31 August 2009, unless otherwise agreed.

Further, testing was not permitted during the first six weeks of the school year because of a concern that student performance levels may have been lower at the beginning of the academic year than at the end of the previous academic year, even after controlling for age.

The 15-year-old international target population was slightly adapted to better fit the age structure of most of the Northern Hemisphere countries. As the majority of the testing was planned to occur in April, the international target population was consequently defined as all students aged from 15 years and 3 completed months to 16 years and 2 completed months at the beginning of the assessment period. This meant that in all countries testing in April 2009, the target population could have been defined as all students born in 1993 who were attending an educational institution as defined above.

A variation of up to one month in this age definition was permitted. This allowed a country testing in March or in May to still define the national target population as all students born in 1993. If the testing was to take place at another time until the end of August, the birth date definition had to be adjusted so that in all countries the target population was always students aged 15 years and 3 completed months to 16 years and 2 completed months at the time of testing, or a one month variation of this.

In all but one country, the Russian Federation, the sampling design used for the PISA assessment was a two-stage stratified sample design. The first-stage sampling units consisted of individual schools having 15-year-old students. Schools were sampled systematically from a comprehensive national list of all PISA-eligible schools, known as the school sampling frame, with probabilities that were proportional to a measure of size. The measure of size was a function of the estimated number of PISA-eligible 15-year-old students enrolled in the school. This is referred to as systematic Probability Proportional to Size (PPS) sampling. Prior to sampling, schools in the sampling frame were assigned to mutually exclusive groups based on school characteristics called explicit strata, formed to improve the precision of sample-based estimates.

The second-stage sampling units in countries using the two-stage design were students within sampled schools. Once schools were selected to be in the sample, a complete list of each sampled school's 15-year-old students was prepared. For each country a Target Cluster Size (*TCS*) was set, this value was typically 35 students although with agreement countries could use alternative values. From each list of students that contained more than the *TCS*, a sample of typically 35 students were selected with equal probability and for lists of fewer than the *TCS*, all students on the list were selected.

In the Russian Federation, a three-stage design was used. In this case, geographical areas were sampled first (first-stage units) using probability proportional to size sampling, and then schools (second-stage units) were selected within these sampled geographical areas. Students were the third-stage sampling units in this three-stage design and were sampled from the selected schools.

## POPULATION COVERAGE, AND SCHOOL AND STUDENT PARTICIPATION RATE STANDARDS

To provide valid estimates of student achievement, the sample of students had to be selected using established and professionally recognised principles of scientific sampling, in a way that ensured representation of the full target population of 15-year-old students in the participating countries.



Furthermore, quality standards had to be maintained with respect to (i) the coverage of the PISA international target population, (ii) accuracy and precision, and (iii) the school and student response rates.

### Coverage of the PISA international target population

National Project Managers (NPMs) might have found it necessary to reduce their coverage of the target population by excluding, for instance, a small, remote geographical region due to inaccessibility, or a language group, possibly due to political, organisational or operational reasons, or special education needs students. In an international survey in education, the types of exclusion must be defined consistently for all participating countries and the exclusion rates have to be limited. Indeed, if a significant proportion of students were excluded, this would mean that survey results would not be deemed representative of the entire national school system. Thus, efforts were made to ensure that exclusions, if they were necessary, were minimised according to the PISA 2009 Technical Standards (see Annex G).

Exclusion can take place at the school level (exclusion of entire schools) or at the within-school level (exclusion of individual students.) Areas deemed by the PGB to be part of a country (for the purpose of PISA), but which were not included for sampling, although this occurred infrequently, were designated as non-covered areas. Care was taken in this regard because, when such situations did occur, the national desired target population differed from the international desired target population.

International within-school exclusion rules for students were specified as follows:

- Intellectually disabled students are students who have a mental or emotional disability and who, in the professional opinion of qualified staff, are cognitively delayed such that they cannot be validly assessed in the PISA testing setting. This category includes students who are emotionally or mentally unable to follow even the general instructions of the test. Students were not to be excluded solely because of poor academic performance or normal discipline problems.
- Functionally disabled students are students who are permanently physically disabled in such a way that they cannot be validly assessed in the PISA testing setting. Functionally disabled students who could provide responses were to be included in the testing.
- Students with insufficient assessment language experience are students who need to meet all of the following criteria: i) are not native speakers of the assessment language(s), ii) have limited proficiency in the assessment language(s), and iii) have received less than one year of instruction in the assessment language(s). Students with insufficient assessment language experience could be excluded.
- Students not assessable for other reasons as agreed upon. A nationally-defined within-school exclusion category was permitted if agreed upon by the PISA Consortium. A specific subgroup of students (for example, students with dyslexia, dysgraphia, or dyscalculia) could be identified for whom exclusion was necessary but for whom the previous three within-school exclusion categories did not explicitly apply, so that a more specific within-school exclusion definition was needed.
- Students taught in a language of instruction for the main domain for which no materials were available. Standard 2.1 notes that the PISA test is administered to a student in a language of instruction provided by the sampled school to that sampled student in the major domain of the test. Thus if no test materials were available in the language in which the sampled student is taught, the student was excluded.

A school attended only by students who would be excluded for intellectual, functional or linguistic reasons was considered a school-level exclusion.

It was required that the overall exclusion rate within a country (i.e. school-level and within-school exclusions combined) be kept below 5% of the PISA Desired Target Population. Guidelines for restrictions on the level of exclusions of various types were as follows:

- School-level exclusions for inaccessibility, feasibility or other reasons were to cover fewer than 0.5% of the total number of students in the international target population for participating countries. Schools on the school sampling frame which had only one or two PISA-eligible students were not allowed to be excluded from the frame. However, if, based on the frame, it was clear that the percentage of students in these small schools would not cause a breach of the 0.5% allowable limit, then such schools could be excluded in the field at that time of the assessment, if they still only had one or two PISA-eligible students.
- School-level exclusions for intellectually or functionally disabled students, or students with insufficient assessment language experience, were to cover fewer than 2% of students.

- Because definitions of within-school exclusions could vary from country to country, NPMs were asked to adapt the international definitions to make them workable in their country but still to code them according to the PISA international coding scheme. Within-school exclusions for intellectually disabled or functionally disabled students, or students with insufficient assessment language experience, or students nationally-defined and agreed upon for exclusion were expected to cover fewer than 2.5% of students. Initially, this could only be an estimate. If the actual percentage was ultimately greater than 2.5%, the percentage was re-calculated without considering students excluded because of insufficient assessment language experience since this is known to be a largely unpredictable part of each country's PISA-eligible population, not under the control of the education system. If the resulting percentage was below 2.5%, the exclusions were regarded as acceptable.

### Accuracy and precision

A minimum of 150 schools had to be selected in each country, if a participating country had fewer than 150 schools, then all schools were selected. Within each participating school, a predetermined number of students, denoted as *TCS* (usually 35 students), were randomly selected with equal probability, or in schools with fewer than *TCS* eligible students, all students were selected. In total, a minimum sample size of 4 500 assessed students was to be achieved, or the full population if it was less than this size. It was possible to negotiate a *TCS* that differed from 35 students, but if it was reduced then the sample size of schools was increased beyond 150, so as to ensure that at least 4 500 students would be assessed. The *TCS* selected per school had to be at least 20 students, so as to ensure adequate accuracy in estimating variance components within and between schools – a major analytical objective of PISA.

NPMs were strongly encouraged to identify available variables to use for defining the explicit and implicit strata for schools to reduce the sampling variance. See later section on stratification for other benefits.

For countries that had participated in PISA 2006 that had larger than anticipated sampling variances associated with their estimates, recommendations were made about sample design changes that would possibly help to reduce the sampling variances for PISA 2009. These included modifications to stratification variables, and increases in the required sample size.

### School response rates

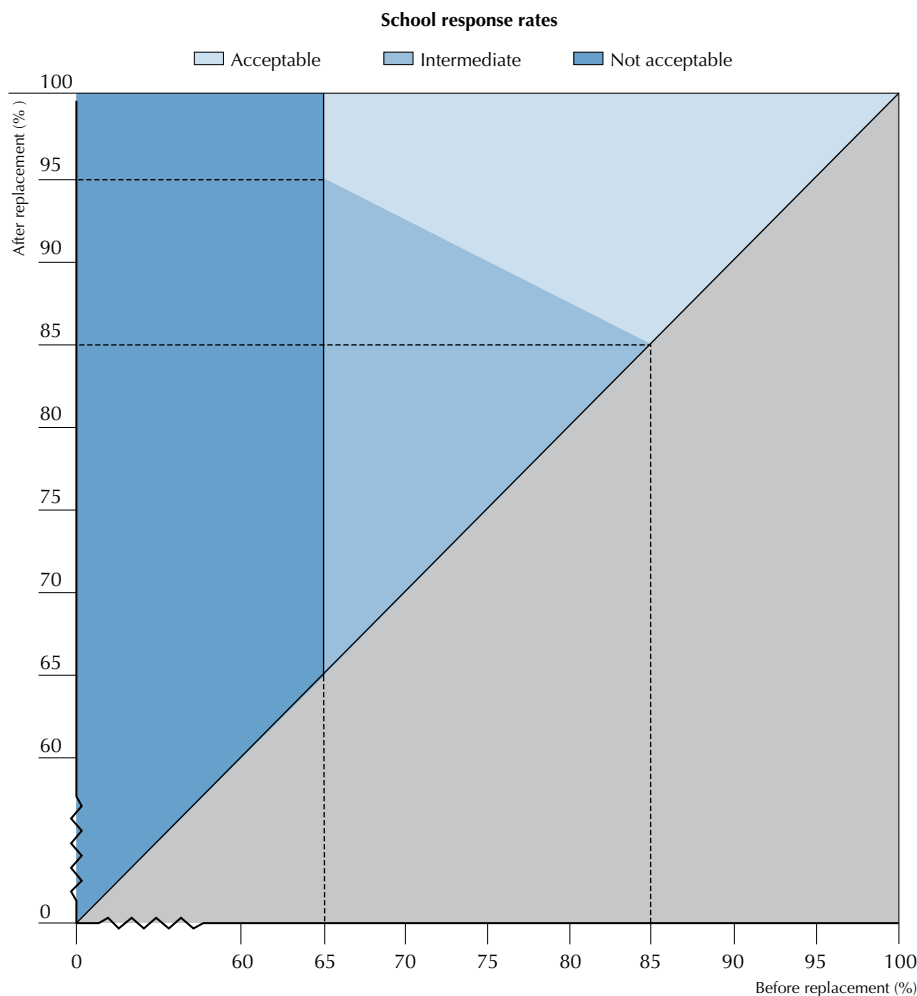
A response rate of 85% was required for initially selected schools. If the initial school response rate fell between 65% and 85%, an acceptable school response rate could still be achieved through the use of replacement schools. Figure 4.1 provides a summary of the international requirements for school response rates. To compensate for a sampled school that did not participate, where possible, two potential replacement schools were identified. Furthermore, a school with a student participation rate between 25% and 50% was not considered as a participating school for the purposes of calculating and documenting response rates.<sup>1</sup> However, data from such schools were included in the database and contributed to the estimates included in the initial PISA international report. Data from schools with a student participation rate of less than 25% were not included in the database, and such schools were regarded as non-respondents.

The rationale for this approach was as follows. There was concern that, in an effort to meet the requirements for school response rates, a national centre might accept participation from schools that would not make a concerted effort to have students attend the assessment sessions. To avoid this, a standard for student participation was required for each individual school in order that the school be regarded as a participant. This standard was set at a minimum of 50% student participation. However, there were a few schools in many countries that conducted the assessment without meeting that standard. Thus a judgement was needed to decide if the data from students in such schools should be used in the analyses, given that the students had already been assessed. If the students from such schools were retained, non-response bias would possibly be introduced to the extent that the students who were absent could have been different in achievement from those who attended the testing session, and such a bias is magnified by the relative sizes of these two groups. If one chose to delete all assessment data from such schools, then non-response bias would be introduced to the extent that the school was different from others in the sample, and sampling variance would be increased because of sample size attrition.

The judgement was made that, for a school with between 25% and 50% student response, the latter source of bias and variance was likely to introduce more error into the study estimates than the former, but with the converse judgement for those schools with a student response rate below 25%. Clearly the cut-off of 25% is arbitrary as one would need extensive studies to try to establish this cut-off empirically. However, it is clear that, as the student response rate decreases within a school, the possibility of bias from using the assessed students in that school will increase, while the loss in sample size from dropping all of the students in the school will be small.



■ Figure 4.1 ■  
**School response rate standards**



These PISA standards applied to weighted school response rates. The procedures for calculating weighted response rates are presented in Chapter 8. Weighted response rates weight each school by the number of students in the population that are represented by the students sampled from within that school. The weight consists primarily of the enrolment size of 15-year-old students in the school, divided by the selection probability of the school. Because the school samples were selected with probability proportional to size, in most countries most schools contributed approximately equal weights, as a consequence the weighted and unweighted school response rates were similar. Exceptions could occur in countries that had explicit strata that were sampled at very different rates. Details as to how the PISA participants performed relative to these school response rate standards are included in Chapters 11 and 14.

### Student response rates

An overall response rate of 80% of selected students in participating schools was required. A student who had participated in the original or follow-up cognitive sessions was considered to be a participant. A minimum student response rate of 50% within each school was required for a school to be regarded as participating: the overall student response rate was computed using only students from schools with at least a 50% student response rate. Again, weighted student response rates were used for assessing this standard. Each student was weighted by the reciprocal of his/her sample selection probability.

## MAIN STUDY SCHOOL SAMPLE

### Definition of the national target population

NPMs were first required to confirm their dates of testing and age definition with the PISA Consortium. Once these were approved, NPMs were alerted to avoid having the possible drift in the assessment period lead to an unapproved definition of the national target population.

Every NPM was required to define and describe their country's target population and explain how and why it might deviate from the international target population. Any hardships in accomplishing complete coverage were specified, discussed and approved or not, in advance. Where the national target population deviated from full coverage of all PISA-eligible students, the deviations were described and enrolment data provided to measure how much coverage was reduced. The population, after all exclusions, corresponded to the population of students recorded on each country's school sampling frame. Exclusions were often proposed for practical reasons such as increased survey costs or complexity in the sample design and/or difficult test conditions. These difficulties were mainly addressed by modifying the sample design to reduce the number of such schools selected rather than to exclude them. Schools with students that would all be excluded through the within-school exclusion categories could be excluded up to a maximum of 2% as previously noted. Otherwise, countries were instructed to include the schools but to administer the PISA UH booklet, consisting of a subset of the PISA assessment items, deemed more suitable for students with special education needs.

Within participating schools, all PISA-eligible students (i.e. born within the defined time period and in grades 7 or higher) were to be listed. From this, either a sample of *TCS* students was randomly selected or all students were selected if there were fewer than *TCS* students. The lists had to include students deemed to meet any of the categories for exclusion, and a variable maintained to briefly describe the reason for exclusion. This made it possible to estimate the size of the within-school exclusions from the sample data.

It was understood that the exact extent of within-school exclusions would not be known until the within-school sampling data were returned from participating schools, and sampling weights computed. Participating country projections for within-school exclusions provided before school sampling were known to be estimates.

NPMs were made aware of the distinction between within-school exclusions and nonresponse. Students who could not take the PISA achievement tests because of a permanent condition were to be excluded and those with a temporary impairment at the time of testing, such as a broken arm, were treated as non-respondents along with other absent sampled students.

Exclusions by country are documented in Chapter 11.

### The sampling frame

All NPMs were required to construct a school sampling frame to correspond to their national defined target population. The school sampling frame was defined by the *School Sampling Preparation Manual*<sup>2</sup> as a frame that would provide complete coverage of the national defined target population without being contaminated by incorrect or duplicate entries or entries referring to elements that were not part of the defined target population. It was expected that the school sampling frame would include any school that could have 15-year-old students, even those schools which might later be excluded, or deemed ineligible because they had no PISA-eligible students at the time of data collection. The quality of the sampling frame directly affects the survey results through the schools' probabilities of selection and therefore their weights and the final survey estimates. NPMs were therefore advised to be diligent and thorough in constructing their school sampling frames.

All but one country used school-level sampling frames as their first stage of sample selection. The *School Sampling Preparation Manual* indicated that the quality of sampling frames for both two and three-stage designs would largely depend on the accuracy of the approximate enrolment of 15-year-olds available (*ENR*) for each first-stage sampling unit. A suitable *ENR* value was a critical component of the sampling frames since selection probabilities were based on it for both two and three-stage designs. The best *ENR* for PISA was the number of currently enrolled 15-year-old students. Current enrolment data, however, were rarely available at the time of school sampling, which meant using alternatives. Most countries used the first-listed available option from the following list of alternatives:

- student enrolment in the target age category (15-year-olds) from the most recent year of data available;
- if 15-year-olds tend to be enrolled in two or more grades, and the proportions of students who are aged 15 in each grade are approximately known, the 15-year-old enrolment can be estimated by applying these proportions to the corresponding grade-level enrolments;



- the grade enrolment of the modal grade for 15-year-olds; and
- total student enrolment, divided by the number of grades in the school.

The *School Sampling Preparation Manual* noted that if reasonable estimates of *ENR* did not exist or if the available enrolment data were out of date, schools might have to be selected with equal probabilities which might require an increased school sample size. However, no countries needed to use this option.

Besides *ENR* values, NPMs were instructed that each school entry on the frame should include at minimum:

- school identification information, such as a unique numerical national identification, and contact information such as name, address and phone number; and
- coded information about the school, such as region of country, school type and extent of urbanisation, which could possibly be used as stratification variables.

As noted, a three-stage design and an area-level (geographic) sampling frame could be used where a comprehensive national list of schools was not available and could not be constructed without undue burden, or where the procedures for administering the test required that the schools be selected in geographic clusters. As a consequence, the area-level sampling frame introduced an additional stage of frame creation and sampling (first stage) before actually sampling schools (second stage with the third stage being students). Although generalities about three-stage sampling and using an area-level sampling frame were outlined in the *School Sampling Preparation Manual* (for example that there should be at least 80 first-stage units and at least 40 needed to be sampled), NPMs were also informed that the more detailed procedures outlined there for the general two-stage design could easily be adapted to the three-stage design. The NPM using a three-stage design was also asked to notify the PISA Consortium and received additional support in constructing and using an area-level sampling frame. The only country that used a three-stage design was the Russian Federation, where a national list of schools was not available. The use of the three-stage design allowed for school lists to be obtained only for those areas selected in stage one rather than for the entire country.

## Stratification

Prior to sampling, schools were to be ordered, or stratified, in the sampling frame. Stratification consists of classifying schools into *like* groups according to selected variables referred to as stratification variables. Stratification in PISA was used to:

- improve the efficiency of the sample design, thereby making the survey estimates more reliable;
- apply different sample designs, such as disproportionate sample allocations, to specific groups of schools, such as those in states, provinces, or other regions;
- ensure all parts of a population were included in the sample; and
- ensure adequate representation of specific groups of the target population in the sample.

There were two types of stratification utilised: explicit and implicit. Explicit stratification consists of grouping schools into strata that will be treated independently from one another or as if they were separate school sampling frames. Examples of explicit stratification variables could be states or regions of a country. Implicit stratification consists essentially of sorting the schools uniquely within each explicit stratum by a set of designated implicit stratification variables. Examples of implicit stratification variables could be type of school, urbanicity, or minority composition. This type of stratification is a way of ensuring a strictly proportional sample allocation of schools across all implicit strata. It can also lead to improved reliability of survey estimates, provided that the implicit stratification variables being considered are correlated with PISA achievement at the school level (Jaeger, 1984). Guidelines were provided in the *School Sampling Preparation Manual* on choosing stratification variables that would possibly improve the sampling.

Table 4.1 provides the explicit stratification variables used by each country, as well as the number of explicit strata found within each country. For example, Australia had eight explicit strata using states/territories which were then further delineated by three sectors and also had one explicit stratum for certain selections, so that there were 25 explicit strata in total. Variables used for implicit stratification and the respective number of levels can also be found in Table 4.1.

As countries were requested to sort the sampling frame by school size, school size was also an implicit stratification variable, though it is not listed in Table 4.1. The use of school size as an implicit stratification variable provides a degree of control over the student sample size so as to possibly avoid the sampling of too many relatively large schools or too many relatively small schools. A variable used for stratification purposes is not necessarily included in the PISA data files.

[Part 1/2]

Table 4.1 Stratification variables used in PISA 2009

	Explicit stratification variables	Number of explicit strata	Implicit stratification variables
<b>Albania</b>	Region (3); Urban/Rural (2); School Type (2); Certainty Selections	11	Public/Private (2); ISCED2/Mixed (2)
<b>Argentina</b>	Area (5)	5	Public/Private (2); School Type (35); Location (3); Orientation (3)
<b>Australia</b>	State/Territory (8); Sector (3); Certainty Selections	25	Geographic Zone (3); School Gender Composition (3); SEIFA (10)
<b>Austria</b>	Region (3); School Type (17); Certainty Selections	32	Province (7); School Type (17); Percentage of Girls (5)
<b>Azerbaijan</b>	School Type (4); Ministry Type (2); Public/Private (2); Language (2); Certainty Selections	14	Urbanicity (4); Education Department or Private (5); Region/District/City (77)
<b>Belgium</b>	Region (3); Form of Education - Flanders (5), French Community (3), German Community (2); Public/Private for Flanders (2) and French Community (4)	23	Flanders - ISCED (4); Retention Rate (5); Vocational/Special Education (2); Percentage of Girls (4); French Community - National/International School (2); Retention Rate (5); Vocational-Special Education/Other (2); German Community - Public/Private (2)
<b>Brazil</b>	State (27); Grade 9 status (3); Certainty Selections	82	Maintenance (3); Urban/Rural (2); HDI Level (3)
<b>Bulgaria</b>	Broad school type (3); Region (11)	32	Type of School (5); Size of Settlement(5); Funding (3)
<b>Canada</b>	Province (10); Language (3); School Size (6); Certainty Selections	45	Public/Private (2); Urban/Rural/Unknown (3)
<b>Chile</b>	Funding type (3); School level (3); School track (4)	18	% Girls (5); Urbanicity(2); Region (4)
<b>Colombia</b>	Region (11); Certainty Selections	12	Urbanicity (2); Funding (2); Weekend school or not (2)
<b>Croatia</b>	Dominant Programme Type (6); Certainty Selections	7	Urbanicity (3); County (21)
<b>Czech Republic</b>	Programmes (6); Region for Programmes 1 and 2 (14); School Size (4)	78	Region for Programmes 3, 4, 5, 6 (14); School Gender Composition for Programmes 4 and 5 (3)
<b>Denmark</b>	Minority Enrollment (4); Certainty Selections	5	School Type (5); Region (5)
<b>Dubai (UAE)</b>	Curriculum (7); Funding (2); Language (2)	9	School Level (3); School Gender (3)
<b>Estonia</b>	Language (3); Certainty Selections	4	School Type (3); Urbanicity (2); County (15)
<b>Finland</b>	Region (5); Urban/Rural (2); Language (3); School Types (5)	12	School Type (5)
<b>France</b>	School Type (4); School Size (3)	6	None
<b>Germany</b>	School Type (3); State (16)	18	Schulart/School Type (7)
<b>Greece</b>	Region (14); Public/Private (2); Evening Schools (1)	17	School Type (3); Public/Private (2) for Evening Schools Stratum
<b>Hong Kong-China</b>	Funding (4)	4	Student Academic Intake (4)
<b>Hungary</b>	School Type (4)	4	Region (7); Reading Performance (5)
<b>Iceland</b>	Region (9)	9	School Size (4)
<b>Indonesia</b>	Indonesia (1)	1	Province (28); Funding (2); School Type and Level (5); Criteria (3)
<b>Ireland</b>	School Type (3); School Size (3)	9	Socio-Economic Status Category (4); School Gender Composition Category (4)
<b>Israel</b>	School Type (2); Language (2); School Orientation (3); Subsectors for Arabic (3); Gender (3)	11	Group Size (2); SES (3); District (6)
<b>Italy</b>	Region (21); Study Programme (5); Certainty Selections	99	Public/Private (2)
<b>Japan</b>	Public/Private (2); School Type (2)	4	Levels of proportion of students taking University/College Entrance Exams (4)
<b>Jordan</b>	School Type / Funding (4)	4	Location (2); Gender (3); Level (2); Shift (2)
<b>Kazakhstan</b>	Region (16); Language (3)	48	Location (2); Level (3); Programme (2); Funding (2)
<b>Korea</b>	School Level (2); School Type for Upper Secondary (2)	3	Urbanicity Level (3); School Gender Composition (3)
<b>Kyrgyzstan</b>	Region (9); Urbanicity (3)	17	Language (7); Type and Level of School (5)





[Part 2/2]  
Table 4.1 Stratification variables used in PISA 2009

	Explicit stratification variables	Number of explicit strata	Implicit stratification variables
<b>Latvia</b>	Urbanicity (4); Certainty Selections	5	School Type and Level (6)
<b>Liechtenstein</b>	Liechtenstein (1)	1	Funding (2)
<b>Lithuania</b>	Location (4); School Type (4)	16	Funding (2)
<b>Luxembourg</b>	School Type (6)	6	School Gender Composition (2)
<b>Macao-China</b>	School Type (3); Programme (2); Language (5)	10	School Orientation (2); Gender (3)
<b>Mexico</b>	State (32); School Size (3); Certainty Selections	97	School Level (2); School Programme (7); Public/Private (2); Urban/Rural (2)
<b>Montenegro</b>	School Type (4); Region (3)	11	None
<b>Netherlands</b>	Limburg/Rest of Netherlands (2); Netherlands School Track (2)	4	Programme Category (6)
<b>New Zealand</b>	Certainty/Non-Certainty (2)	2	Socio-Economic Status Category (3); Public/Private (2); School Gender Composition (3); Urban/Rural (2)
<b>Norway</b>	School Level (3)	3	None
<b>Panama</b>	Urbanicity (2); Funding (2); Certainty Selections	4	Region (12); Orientation (2)
<b>Peru</b>	Funding (2); Urbanicity (2)	4	Region (26); Gender (3); School Type (4)
<b>Poland</b>	School Type (3); Public/Private (2) for Gymnasia	4	School Subtype (5); Public/Private (3) for Lycea and Vocational Schools; Locality (4)
<b>Portugal</b>	Geographic Region (30); Certainty Selections	31	Island (10); ISCED (3); Public/Private (2); Urbanicity (3)
<b>Qatar</b>	School Type (8)	8	Gender (3); Level (5); Funding (2)
<b>Romania</b>	Programme (3)	3	Language (3); Urbanicity (2)
<b>Russian Federation</b>	Region (45)	45	Location (9); School Type (8); School Sub-type (5);
<b>Serbia</b>	Region (8); School Type (8); Certainty Selections	58	None
<b>Shanghai-China</b>	School Level (3); Programme (2); Selectivity (2); Certainty Selections	7	Track (2); Funding (2); Location (2)
<b>Singapore</b>	Funding (2); Level (2); Certainty Selections	4	Gender (3)
<b>Slovak Republic</b>	Region (8); School Type (3)	24	Programme (9); Language (3); Grade Repetition Level (112)
<b>Slovenia</b>	Programme and Level (7)	7	Location (5)
<b>Spain</b>	Region (18); Public/Private (2); Teaching Modality for Basque (3); Certainty Selections	41	Postal Code for all
<b>Sweden</b>	Public/Private (2); School Level (2); Urbanicity (5) for Lower Secondary Schools	12	Geographic LAN (21) for Upper Secondary schools; School Type (3) for Upper Secondary schools; Income Quartiles (4) for Lower Secondary schools
<b>Switzerland</b>	School has Grade 9 or not (2); Language (3); Canton for adjudicated regions with Grade 9 oversample/Rest of Switzerland (13); Public/Private (3); School Type (4) within Upper Secondary schools; Certainty Selections	30	School Type (29)
<b>Chinese Taipei</b>	School type (7); Funding (2); Location (2); Certainty Selections	29	County/City area (25); School Gender (3)
<b>Thailand</b>	Administration (6); School Type (3); Certainty Selections	15	Local area (9)
<b>Trinidad and Tobago</b>	Districts (8); Management (3)	23	Gender (3); Programme (2); Level (2); Location (2)
<b>Tunisia</b>	Geographical Area (4); Level (3); Funding (2)	16	% Repeaters (3)
<b>Turkey</b>	Region (12); Programme (3)	36	Turkey School Type (17); Urban/Rural (2); Public/Private (2)
<b>United Kingdom</b>	Country (4); School Type (3) for England, Northern Ireland, and Wales; Region -- England (4), Northern Ireland (5), Wales (3); Certainty Selections; Scotland -- School Attainment Level (6)	36	England -- School Attainment Level (6); School Gender Composition (3); Local Authority; Northern Ireland -- School Gender Composition (3); Wales -- School Gender Composition (3); Local Authority; Scotland -- Area Type (6)
<b>United States</b>	Public/Private (2); Region (4)	8	Grade Span (5); Urbanicity (4); Minority Status (2); 3-digit Postal Code
<b>Uruguay</b>	Funding (2); School Type (3); Region (7); Certainty Selections	22	Level (3); Evening Shift/Not (2)

## Assigning a measure of size to each school

For the probability proportional to size sampling method used for PISA, a Measure of Size (*MOS*) derived from *ENR* was established for each school on the sampling frame. *MOS* was constructed as:  $MOS = \max(ENR, TCS)$ .

Thus, the measure of size was equal to the enrolment estimate (*ENR*), unless enrolment was less than the *TCS*, in which case the measure of size was set equal to the target cluster size. In most countries, the *MOS* was equal to *ENR* or 35 students, whichever was larger.

As schools were sampled with probability proportional to size, setting the measure of size of small schools to 35 students was equivalent to drawing a simple random sample of small schools. That is, small schools had an equally likely chance of being selected to participate.

## School sample selection

### School sample allocation over explicit strata

The total number of schools to be sampled in each country needed to be allocated among the explicit strata so that the expected proportion of students in the sample from each explicit stratum was approximately the same as the population proportions of PISA-eligible students in each corresponding explicit stratum. There were two exceptions. If very small schools required under-sampling, students in them had smaller percentages in the sample than in the population. To compensate for the resulting loss of sample, the large schools had slightly higher percentages in the sample than the corresponding population percentages. The other exception occurred if only one school was allocated to any explicit stratum. In this case, two schools were allocated for selection in the stratum to aid with variance estimation.

### Sorting the sampling frame

The *School Sampling Preparation Manual* indicated that, prior to selecting schools, schools in each explicit stratum were to be sorted by variables chosen for implicit stratification and finally by the *ENR* value within each implicit stratum. The schools were first to be sorted by the first implicit stratification variable, then by the second implicit stratification variable within the levels of the first implicit stratification variable, and so on, until all implicit stratification variables were used. This gave a cross-classification structure of cells, where each cell represented one implicit stratum on the school sampling frame. The sort order was alternated between implicit strata, from high to low and then low to high, etc., through all implicit strata within an explicit stratum.

### Determining which schools to sample

The PPS-systematic sampling method used in PISA first required the computation of a sampling interval for each explicit stratum. This calculation involved the following steps:

- recording the total measure of size,  $S$ , for all schools in the sampling frame for each specified explicit stratum;
- recording the number of schools,  $D$ , to be sampled from the specified explicit stratum, which was the number allocated to the explicit stratum;
- calculating the sampling interval,  $I$ , as follows:  $I = S/D$ ; and
- recording the sampling interval,  $I$ , to four decimal places.

Next, a random number had to be generated for each explicit stratum. The generated random number (*RN*) was from a uniform distribution between zero and one and was to be recorded to four decimal places.

The next step in the PPS selection method in each explicit stratum was to calculate selection numbers – one for each of the  $D$  schools to be selected in the explicit stratum. Selection numbers were obtained using the following method:

- Obtaining the first selection number by multiplying the sampling interval,  $I$ , by the random number, *RN*. This first selection number was used to identify the first sampled school in the specified explicit stratum.
- Obtaining the second selection number by adding the sampling interval,  $I$ , to the first selection number. The second selection number was used to identify the second sampled school.
- Continuing to add the sampling interval,  $I$ , to the previous selection number to obtain the next selection number. This was done until all specified line numbers (1 through  $D$ ) had been assigned a selection number.

Thus, the first selection number in an explicit stratum was  $RN \times I$ , the second selection number was  $(RN \times I) + I$ , the third selection number was  $(RN \times I) + I + I$ , and so on.

Selection numbers were generated independently for each explicit stratum, with a new random number generated for each explicit stratum.

### Identifying the sampled schools

The next task was to compile a cumulative measure of size in each explicit stratum of the school sampling frame that assisted in determining which schools were to be sampled. Sampled schools were identified as follows.

Let  $Z$  denote the first selection number for a particular explicit stratum. It was necessary to find the first school in the sampling frame where the cumulative  $MOS$  equalled or exceeded  $Z$ . This was the first sampled school. In other words, if  $C_s$  was the cumulative  $MOS$  of a particular school  $S$  in the sampling frame and  $C_{(s-1)}$  was the cumulative  $MOS$  of the school immediately preceding it, then the school in question was selected if:  $C_s$  was greater than or equal to  $Z$ , and  $C_{(s-1)}$  was strictly less than  $Z$ . Applying this rule to all selection numbers for a given explicit stratum generated the original sample of schools for that stratum.

#### Box 4.1 Illustration of probability proportional to size (PPS) sampling

To illustrate these steps, suppose that in an explicit stratum in a participant country, the PISA-eligible student population is 105 000, then:

- the total measure of size,  $S$ , for all schools is 105 000;
- the number of schools,  $D$ , to be sampled is 150;
- calculating the sampling interval,  $I$ ,  $105\ 000/150 = 700$ ;
- generate a random number,  $RN$ , 0.3230;
- the first selection number is  $700 \times 0.3230 = 226$ . This first selection number is used to identify the first sampled school in the specified explicit stratum; and
- the second selection number is  $226 + 700 = 926$ . The second selection number was used to identify the second sampled school.

The third selection number is  $926 + 700 = 1\ 626$ . The third selection number was used to identify the third sampled school, and so on until the end of the school list is reached. This will result in a school sample size of 150 schools.

The table below also provides these example data. The school that contains the generated selection number within its cumulative enrolment is selected for participation.

School	MOS	Cumulative MOS ( $C_s$ )	Selection Number	
001	550	550	226	Selected
002	364	914		
003	60	974	926	Selected
004	93	1 067		
005	88	1 155		
006	200	1 355		
007	750	2 105	1 626	Selected
008	72	2 177		
009	107	2 284		
010	342	2 626	2 326	Selected
011	144	2 770		

### Identifying replacement schools

Each sampled school in the main survey was assigned two replacement schools from the school sampling frame, if possible, identified as follows. For each sampled school, the schools immediately preceding and following it in the explicit stratum, which was ordered within by the implicit stratification, were designated as its replacement schools. The school immediately following the sampled school was designated as the first replacement and labelled  $R_1$ , while



the school immediately preceding the sampled school was designated as the second replacement and labelled  $R_2$ . The *School Sampling Preparation Manual* noted that in small countries, there could be problems when trying to identify two replacement schools for each sampled school. In such cases, a replacement school was allowed to be the potential replacement for two sampled schools (a first replacement for the preceding school, and a second replacement for the following school), but an actual replacement for only one school. Additionally, it may have been difficult to assign replacement schools for some very large sampled schools because the sampled schools appeared close to each other in the sampling frame. There were times when it was only possible to assign a single replacement school, or even none, when two consecutive schools in the sampling frame were sampled. That is, no unsampled schools existed between sampled schools.

Exceptions were allowed if a sampled school happened to be the last school listed in an explicit stratum. In this case the two schools immediately preceding it were designated as replacement schools. Similarly, for the first school listed in an explicit stratum, in which case the two schools immediately following it were designated as replacement schools.

### **Assigning school identifiers**

To keep track of sampled and replacement schools in the PISA database, each was assigned a unique, three-digit school code and two-digit stratum code (corresponding to the explicit strata) sequentially numbered starting with one within each explicit stratum. For example, if 150 schools are sampled from a single explicit stratum, they are assigned identifiers from 001 to 150. First replacement schools in the main survey are assigned the school identifier of their corresponding sampled schools, incremented by 300. For example, the first replacement school for sampled school 023 is assigned school identifier 323. Second replacement schools in the main survey are assigned the school identifier of their corresponding sampled schools, but incremented by 600. For example, the second replacement school for sampled school 136 took the school identifier 736.

### **Tracking sampled schools**

NPMs were encouraged to make every effort to confirm the participation of as many sampled schools as possible to minimise the potential for non-response biases. They contacted replacement schools after all contacts with sampled schools were made. Each sampled school that did not participate was replaced if possible. If both an original school and a replacement participated, only the data from the original school were included in the weighted data provided that at least 50% of the PISA-eligible, non-excluded students had participated. If this was not the case, it was permissible for the original school to be labelled as a nonrespondent and the replacement school as the respondent, provided that the replacement school had at least 50% of the PISA-eligible, non-excluded students as participants.

## **Special school sampling situations**

### **Treatment of small schools**

In PISA, schools were classified as very small, moderately small or large. A school was classified as large if it had an *ENR* above the *TCS* (35 students in most countries). A moderately small school had an *ENR* in the range of one-half the *TCS* to *TCS* (18 to 35 students in most countries). A very small school had an *ENR* less than one-half the *TCS* (17 students or fewer in most countries). Unless they received special treatment in the sampling, the occurrence of small schools in the sample will reduce the sample size of students for the national sample to below the desired target because the within-school sample size would fall short of expectations. A sample with many small schools could also be an administrative burden with many testing sessions with few students. To minimise these problems, procedures were devised for managing small schools in the sampling frame.

To balance the two objectives of selecting an adequate sample of small schools but not too many small schools so as to hurt student yield, a procedure was recommended that assumed the underlying idea of under-sampling the very small schools by a factor of two and to proportionally increasing the number of large schools to sample. Rather than create a stratum for very small schools and/or a stratum for moderately small schools for PISA 2009, the number of very small schools was controlled in the sample by assigning a measure of size to these schools equal to the . In effect, they were under-sampled by a factor of two (school probability of selection reduced by half), without explicitly stratifying them. This was accomplished as follows.

The sample had to be proportional to the number of students in the participating country and not to the number of schools. Suppose that 10% of students attend moderately small schools, 10% very small schools and the remaining 80% attend large schools. In the sample of 5 250, 4 200 students would be expected to come from large schools (i.e. 120 schools with 35 students), 525 students from moderately small schools and 525 students from very small schools.



If moderately small schools had an average of 25 students, then it would be necessary to include 21 moderately small schools in the sample. If the average size of very small schools was 10 students, then 52 very small schools would be needed in the sample and the school sample size would be equal to 193 schools rather than 150.

The formulae below assume a target school sample size of 150 and a target student sample size of 5 250.

- Step 1: From the complete sampling frame, find the proportions of total *ENR* that come from very small schools (*P*), moderately small schools (*Q*), and large schools (*R*). Thus,
- Step 2: Calculate the value *L*, where  $L = 1.0 + (P/2)$ . Thus *L* is a positive number slightly more than 1.0.
- Step 3: The minimum sample size for large schools is equal to  $150 \times R \times L$ , rounded to the nearest integer. It may need to be enlarged because of national considerations, such as the need to achieve minimum sample sizes for geographic regions or certain school types.
- Step 4: Calculate the mean value of *ENR* for moderately small schools (*MENR*), and for very small schools (*VENR*). *MENR* is a number in the range of *TCS*/2 to *TCS*, and *VENR* is a number no greater than *TCS*/2.
- Step 5: The number of schools that must be sampled from the moderately small schools is given by:  $(5\,250 \times Q \times L)/(MENR)$ .
- Step 6: The number of schools that must be sampled from the very small schools is given by:  $(2\,625 \times P \times L)/(VENR)$ .

To illustrate the steps, suppose that in a participant country, the *TCS* is equal to 35 students, with 10% of the total enrolment of 15-year-olds each in moderately small schools and in very small schools. Suppose that the average enrolment in moderately small schools is 25 students, and in very small schools it is 10 students.

- Step 1: The proportions of total *ENR* from very small schools is  $P = 0.1$ , moderately small schools is  $Q = 0.1$ , and large schools is  $R = 0.8$ . It can be shown that  $0.1 + 0.1 + 0.8 = 1.0$ .
- Step 2: Calculate the value *L*.  $L = 1.0 + (0.1/2)$ . Thus  $L = 1.05$ .
- Step 3: The minimum sample size for large schools is equal to  $150 \times 0.8 \times 1.05 = 126$ . That is, at least 126 of the large schools must be sampled.
- Step 4: The mean value of *ENR* for moderately small schools (*MENR*) is given in this example as 25, and for very small schools (*VENR*) as 10.
- Step 5: The number of schools that must be sampled from the moderately small schools is given by  $(5\,250 \times 0.1 \times 1.05)/25 = 22.1$ . At least 22 (rounded to the nearest integer) moderately small schools must be sampled.
- Step 6: The number of schools that must be sampled from the very small schools is given by  $(2\,625 \times 0.1 \times 1.05)/10 = 27.6$ . At least 28 (rounded to the nearest integer) very small schools must be sampled.

Combining these different sized school samples gives a total sample size of  $126 + 22 + 28 = 176$  schools, rather than just 150, or 193 as calculated above. Before considering school and student non-response, the larger schools will yield an initial sample of approximately  $126 \times 35 = 4\,410$  students. The moderately small schools will give an initial sample of approximately  $22 \times 25 = 550$  students, and very small schools will give an initial sample size of approximately  $28 \times 10 = 280$  students. The total initial sample size of students is therefore  $4\,410 + 550 + 280 = 5\,240$ .

This procedure, called small school analysis, was done not just for the entire school sampling frame, but new for 2009 for each individual explicit stratum. An initial allocation of schools to explicit strata provided the starting number of schools and students to project for sampling in each explicit stratum. The small school analysis for a single unique explicit stratum indicated how many very small schools (assuming under-sampling by 2, if needed), moderately small schools and large schools would be sampled in that stratum. Together, these provided the final sample size, *n*, of schools to select in the stratum. Based on the stratum sampling interval and random start, large, moderately small, and very small schools were sampled in the stratum, to a total of *n* sampled schools. Because of the random start, it was possible to have more or less than expected of the very small schools, of the moderately small schools, and of the large schools. The total number of sampled schools however was fixed at *n*, and the number of expected students to be sampled was always approximate to what had been projected from the unique stratum small school analysis.

### **Sampling for the Digital Reading Assessment (DRA) component**

Nineteen countries and economies participated in the Digital Reading Assessment (DRA): Australia, Austria, Belgium, Chile, Colombia, Denmark, France, Hong Kong-China, Hungary, Iceland, Ireland, Japan, Korea, Macao-China, New Zealand, Norway, Poland, Spain and Sweden. When a country participated in DRA, it was expected that DRA student sampling would occur in every PISA sampled and participating school.

The overall sample size requirement was 1 200 assessed DRA students. The recommended DRA Target Cluster Size (DTCS) was 14 students per sampled school. While 14 students for each of the 150 schools (the typical number of PISA schools per participating country) would potentially yield 2 100 students, the large DTCS was chosen to account for the fact that some schools would not have adequate computer resources. The DTCS of 14 students also accounted for the loss in the DRA sample that would accrue from prior losses in the PISA sample. It was a requirement that all DRA students also participate in the main paper-based PISA assessment. The DRA student sample was selected at the same time that the PISA student sample was selected in each school by the student sampling software, *KeyQuest*. Therefore, any PISA student also sampled for DRA who did not participate in paper-based PISA assessment was an automatic loss for the DRA student sample. There would also be additional loss for DRA due to refusals, or other absences. Setting the DTCS at 14 students guarded against these losses. It was possible to vary the DTCS if more than the usual number of schools were sampled for PISA.

The actual DRA student sample size at each school was calculated with *KeyQuest*, as the minimum of the DTCS, and the number of sampled PISA students. Arrangements had to be made at the school level to either bring in laptops, or to have extra sessions to alleviate any computer resource problems.

If a country had a large PISA school sample and wished to subsample the PISA sampled schools where DRA student sampling would be done, this became an additional national option. Only two DRA countries, Colombia and Spain, chose to have schools subsampled for DRA from their large national school sample.

The schools for DRA were subsampled with equal probability from sampled schools in each explicit stratum. The number to subsample for DRA in each stratum was based on how many schools would have been needed from each explicit stratum for a school sample of 150 schools. Any schools selected with certainty for the large national school sample and placed in their own stratum, were added back to their original strata for the subsampling of DRA schools.

### **PISA and ICCS overlap control**

The main studies for PISA 2009 and the 2009 International Civic and Citizenship Education Study (ICCS) were to occur at approximately the same time in some participating countries. Because of the potential for increased burden, an overlap control procedure was used for twelve countries (Belgium [Flemish], England and Northern Ireland from the United Kingdom, Finland, Ireland, Latvia, the Netherlands, New Zealand, Norway, the Slovak Republic, Sweden and Chinese Taipei) who requested for there to be a minimum incidence of the same schools being sampled for both PISA and ICCS. This overlap control procedure required that the same school identifiers be used on the PISA and ICCS school frames for the schools in common across the two assessments.

The ICCS samples were usually selected before the PISA samples. Thus, for countries requesting overlap control, the ICCS International Study Centre supplied the PISA Consortium with their school frames, school IDs, each school's probability of selection, and an indicator showing which schools had been sampled for the ICCS study.

Sample selections for PISA and ICCS could totally avoid overlap of schools if schools which would have been selected with high probability for either study had their selection probabilities capped at 0.5. Such an action would make each study's sample slightly less than optimal, but this might be deemed acceptable when weighed against the possibility of low response rates due to the burden of participating in two assessments. This was requested only by Ireland. In the other countries, if any schools had probabilities of selection greater than 0.5 on either study frame, these schools had the possibility to be selected to be in both studies.

To control overlap of schools between PISA and ICCS, the sample selection of schools for PISA adopted a modification of an approach due to Keyfitz (1951), based on Bayes Theorem. To use PISA and ICCS in an example of the overlap control approach, suppose that *PROBP* is the PISA probability of selection and *PROBI* is the ICCS probability of selection. Then a conditional probability of a school's selection into PISA (*CPROB*) is determined as follows:

$$4.1 \quad CPROB = \begin{cases} \max \left[ 0, \left( \frac{PROBI + PROBP - 1}{PROBI} \right) \right] & \text{if the school was an ICCS school} \\ \min \left[ 1, \frac{PROBP}{(1 - PROBI)} \right] & \text{if the school was not an ICCS school} \\ PROBP & \text{if the school was not an ICCS eligible school} \end{cases}$$



Then a conditional *CMOS* variable was created to coincide with these conditional probabilities as follows:

$$CMOS = CPROB \times \text{stratum sampling interval}$$

The PISA school sample was then selected using the line numbers created as usual (see earlier section), but applied to the cumulated *CMOS* values (as opposed to the cumulated *MOS* values). Note that it was possible that the resulting PISA sample size could be slightly lower or higher than the originally assigned PISA sample size, but this was deemed acceptable.

Luxembourg also requested to have minimal overlap with ICCS but since a census of PISA students is usually taken, this presented a unique challenge. It was agreed that although ICCS usually sampled grade 8 classes, grade 8 students of PISA age would be listed and ICCS would take half of the students for their study while PISA would take the other half of students for their study.

### Monitoring school sampling

For PISA 2009, as in the previous two cycles, it was a strong recommendation that the PISA Consortium select the school samples rather than the participating countries. This was incorporated into the 2009 procedures to alleviate the weighting difficulties caused by receiving school sampling frame files in many different formats. Japan was the only participant that selected their own school sample, doing so for reasons of confidentiality.

Sample selection for Japan was replicated by the PISA Consortium to ensure quality in this case. All other participating countries school samples were selected by and checked in detail by the PISA Consortium. To enable this, all countries were required to submit sampling information on forms associated with the following various sampling tasks:

- time of testing and age definition for both the field trial and main study were captured on Sampling Task 1 at the time of the field trial, with updates being possible before the main study;
- information about stratification for the field trial and for the main study was recorded on Sampling Task 2;
- forms or data associated with Sampling Tasks 3, 4, 5 and 6 were all for the field trial;
- the national desired target population information for the main study was captured on the form associated with Sampling Task 7a;
- information about the defined national target population was recorded on the form associated with Sampling Task 7b;
- the description of the sampling frame was noted on the form associated with Sampling Task 8a; and
- the school sampling frame was created in one spreadsheet and the list of any excluded schools in a second spreadsheet associated with Sampling Task 8b.

The PISA Consortium completed and returned other information (small school analyses, school allocation, and sample selection) along with a spreadsheet that countries could use for tracking school participation. In some cases, countries also submitted other sets of information for approval. Table 4.2 provides a summary of the information required for each sampling task and the timetables (which depended on national assessment periods).

Once received from each participating country, each set of information was reviewed and feedback was provided to the country. Forms were only approved after all criteria were met. Approval of deviations was only given after discussion and agreement by the PISA Consortium. In cases where approval could not be granted, countries were asked to make revisions to their sample design and sampling forms and resubmit.

Checks that were performed in the monitoring of each set of information follow. All entries were observed in their own right but those below were additional matters explicitly examined.

As part of the initial pre-form checks, all special situations known about the participating country were verified with the country. Such special situations included, TCS values different from 35 students, whether or not the Digital Reading Assessment was being conducted, whether or not overlap control procedures with ICCS were required, whether or not there was any regional or other type of oversampling, whether or not the UH booklet would be used, and whether or not any grade or other type of student sampling would be used. Additionally, any countries with fewer than 4 500 or just over 4 500 assessed students in either PISA 2003 or 2006 had increased school sample sizes discussed and agreed upon. Finally, any countries with effective student sample sizes less than 400 in PISA 2006 also had increased school sample sizes discussed and agreed upon.

Table 4.2 Schedule of school sampling activities

Activity	Submit to Consortium	Due date
Update time of testing and age definition of population to be tested	Sampling Task 1 – time of testing and age definition	Update what was submitted at the time of the FT, two months before the school sample is to be selected
Finalise explicit and implicit stratification variables	Sampling Task 2 – stratification and other information	Update what was submitted at the time of the FT, two months before the school sample is to be selected
Define national desired target population	Sampling Task 7a – national desired target population	Submit two months before the school sample is to be selected
Define national defined target population	Sampling Task 7b – national defined target population	Submit two months before the school sample is to be selected
Create and describe sampling frame	Sampling Task 8a – sampling frame description	Submit two months before the school sample is to be selected
Submit sampling frame	Sampling Task 8b – sampling frame (in one Excel® sheet), and excluded schools (in another Excel® sheet)	Submit two months before the school sample is to be selected
Decide how to treat small schools	Treatment of small schools	The Consortium will complete and return this information to the NPM about one month before the school sample is to be selected
Finalise sample size requirements	Sampling Task 9 – sample allocation by explicit strata	The Consortium will complete and return this information to the NPM about one month before the school sample is to be selected
Describe population within strata	Population counts by strata	The Consortium will complete and return this information to the NPM when the school sample is sent to the NPM
Select the school sample	Sampling Task 10 – school sample selection	The Consortium will return the sampling frame to the NPM with sampled schools and their replacement schools identified and with PISA IDs assigned when the school sample is selected
Review and agree to the sampling form required as input to <i>KeyQuest</i>	Sampling Task 11 – reviewing and agreeing to the Sampling Form for <i>KeyQuest</i> (SFKQ)	Countries had one month after their sample was selected to agree to their SFKQ
Submit sampling data	Sampling Task 12 – school participation information and data validity checks	Submit within one month of the end of the data collection period

### Sampling Task 1: Time of testing and age definition

- Assessment dates had to be appropriate for the selected target population dates.
- Assessment dates could not cover more than a 42-day period unless agreed upon.
- Assessment dates could not be within the first six weeks of the academic year.
- If assessment end dates were close to the end of the target population birth date period, NPMs were alerted not to conduct any make-up sessions beyond the date when the population births dates were valid.

### Sampling Task 2: Stratification (and other information)

- Since explicit strata are formed to group similar schools together to reduce sampling variance and to ensure representativeness of students in various school types, using variables that might be related to outcomes, each participating country's choice of explicit stratification variables was assessed. If a country was known to have school tracking or distinct school programmes and these were not among the explicit stratification variables, a suggestion was made to include this type of variable.
- Levels of variables and their codes were checked for completeness.
- If no implicit stratification variables were noted, suggestions were made about ones that might be used. In particular, if a country had single gender schools and school gender was not among the implicit stratification variables, a suggestion was made to include this type of variable to ensure no sample gender imbalances. Similarly, if there were ISCED school level splits, the ISCED school level was also suggested as an implicit stratification variable.
- Without overlap control there is nearly as good control over the sample whether explicit or implicit strata are used. With overlap control some control is lost when using implicit strata, but not when using explicit strata. For countries which wanted overlap control with ICCS, as many as possible of their implicit stratification variables were made explicit stratification variables.
- A new requirement for PISA 2009 was that there could only be one student sampling option per explicit stratum. Checks were done to ensure this.





### **Sampling Task 7a: National desired target population**

- The total national number of 15-year olds of participating countries was compared with those from previous cycles. Differences, and any kind of trend, were queried.
- Large deviations between the total national number of 15-year-olds and the enrolled number of 15-year-olds were questioned.
- Large increases or decreases in enrolled population numbers compared to those from previous PISA cycles were queried, as were increasing or decreasing trends in population numbers since PISA 2000.
- Any population to be omitted from the international desired population was noted and discussed, especially if the percentage of 15-year-olds to be excluded was more than 0.5% or if it was not noted for PISA 2006.
- Calculations did not have to be verified as in previous cycles as such data checks were built into the form.
- For any countries using a three-stage design, a Sampling Task 7a form also needed to be completed for the full national desired population as well as for the population in the sampled regions.
- For countries having adjudicated regions, a Sampling Task 7a form was needed for each region.
- If websites were provided with an English page option, the submitted data was verified against those sources.

### **Sampling Task 7b: National defined target population**

- The population value in the first question needed to correspond with the final population value on the form for Sampling Task 7a. This was accomplished through built-in data checks.
- Reasons for excluding schools for reasons other than special education needs were checked for appropriateness (i.e. some operational difficulty in assessing the school). In particular, school-level language exclusions were closely examined to check correspondence with what had been noted about language exclusions on Sampling Task 2.
- Exclusion types and extents were compared to those recorded for PISA 2006 and previous cycles. Differences were queried.
- The number and percentage of students to be excluded at the school level and whether the percentage was less than the guideline for maximum percentage allowed for such exclusions were checked.
- Reasonableness of assumptions about within-school exclusions was assessed by checking previous PISA coverage tables. If there was an estimate noted for “other”, the country was queried for reasonableness about what the “other” category represented. New for PISA 2009 was a within-school exclusion category for “no tests in the student’s language of instruction”. It was necessary to have estimates for this type of within-school exclusion if it was known the country would have such students.
- Form calculations were verified through built-in data checks, and the overall coverage figures were assessed.
- New for PISA 2009, if it was noted that there was a desire to exclude schools with only one or two PISA-eligible students at the time of contact, then the school sampling frame was checked for the percentage of population that would be excluded. If countries had not met the 2.5% school-exclusion guideline and if these schools would account for not more than 0.5% and if within-school exclusions looked similar to the past and were within 2.5%, then the exclusion of these schools at the time of contact was agreed upon.
- The population figures on this form after school-level exclusions were compared against the aggregated school sampling frame enrolment. Differences were queried.
- For any countries using a three-stage design, a Sampling Task 7b form also needed to be completed for the full national defined population as well as for the population in the sampled regions.
- For countries having adjudicated regions, a Sampling Task 7b form was needed for each region.
- If websites were provided with an English page option, the submitted data was verified against those sources.

### **Sampling Task 8a: Sampling frame description**

- Special attention was given to countries who reported on this form that a three-stage sampling design was to be implemented and additional information was sought from countries in such cases to ensure that the first-stage sampling was done adequately.
- The type of school-level enrolment estimate and the year of data availability were assessed for reasonableness.
- New for PISA 2009, countries were asked to provide information for each of various school types,<sup>3</sup> whether those schools were included on or excluded from the sampling frame, or the country did not have any of such schools. The information was matched to the different types of schools containing PISA students noted on Sampling Task 2. Any discrepancies were queried.
- Any school types noted as being excluded were verified as school-level exclusions on the Sampling Task 7b form. Any discrepancies were queried.



### **Sampling Task 8b: Sampling frame**

- On the spreadsheet for school-level exclusions, the number of schools and the total enrolment figures, as well as the reasons for exclusion, were checked to ensure correspondence with values reported on the Sampling Task 7b form detailing school-level exclusions. It was verified that this list of excluded schools did not have any schools which only had one or two PISA-eligible students, as these schools were not to be excluded from the school sampling frame. Checks were done to ensure that excluded schools did not still appear on the other spreadsheet containing the school sampling frame.
- All units on the school sampling frame were confirmed to be those reported on the Sampling Task 2 as sampling frame units. The sampling unit frame number was compared to the corresponding frame for PISA 2006 as well as previous cycles. Differences were queried.
- NPMs were queried about whether or not they had included schools with grades 7 or 8, or in some cases those with grades 10 or higher, that could potentially have PISA-eligible students at the time of assessment even if the school currently did not have any.
- NPMs were queried about whether they had included vocational or apprenticeship schools, schools with only part-time students, international or foreign schools or schools not under the control of the Ministry of Education or any other irregular schools that could contain PISA-eligible students at the time of the assessment, even if such schools were not usually included in other national surveys.
- The frame was checked for all required variables: a national school identifier with no duplicated values, a variable containing the school enrolment of PISA-eligible students, and all the explicit and implicit stratification variables and all related levels as noted on Sampling Task 2, and that none had missing values.
- Any additional school sampling frame variables were assessed for usefulness. In some instances other variables were noted on the school frame that might also have been useful for stratification.
- The frame was checked for schools with only one or two PISA-eligible students. If no schools were found with extremely low counts, but the country's previous sampling frames had some, this was queried.
- The frame was checked for schools with zero enrolment. If there were none, this was assessed for reasonableness. If some existed, it was verified with the NPM that these schools could possibly have PISA-eligible students at the time of the assessment.

### **Treatment of small schools**

- All calculations were verified.
- It was verified that separate small school analyses were done for adjudicated or non-adjudicated oversampled regions (if these were different from explicit strata).

### **Sampling Task 9: Sample allocation by explicit strata**

- All explicit strata had to be accounted for on the form for Sampling Task 9.
- All explicit strata population entries were compared to those determined from the sampling frame.
- The calculations for school allocation were checked to ensure that schools were allocated to explicit strata based on explicit stratum student percentages and not explicit stratum school percentages, that all explicit strata had at least two allocated schools, and that no explicit stratum had only one remaining non-sampled school.
- It was verified that the allocation matched the results of the explicit strata small school analyses, with allowances for random deviations in the numbers of very small, moderately small, and large schools to be sampled in each explicit stratum.
- The percentage of students in the sample for each explicit stratum had to be approximate to the percentage in the population for each stratum (except in the case of oversampling).
- The overall number of schools to be sampled was checked to ensure that at least 150 schools would be sampled.
- The overall number of students to be sampled was checked to ensure that at least 5 250 students would be sampled.
- Previous PISA response rates were reviewed and if deemed necessary, sample size increases were suggested.

### **Population counts by strata**

- Population counts by strata were compared to counts arising from the frame.

### **Sampling Task 10: School sample selection**

- All calculations were verified, including those needed for ICCS overlap control.
- Particular attention was paid to the required four decimal places for the sampling interval and the generated random number.



- The frame was checked for proper sorting according to the implicit stratification scheme, for enrolment values, and the proper assignment of the measure of size value, especially for very small and moderately small schools. The assignment of replacement schools and PISA identification numbers were checked to ensure that all rules established in the *Sampling Preparation Manual* were adhered to.

### **Sampling Task 11: Reviewing and agreeing to the Sampling Form**

- The form for Sampling Task 11 was prepared as part of the sample selection process. After the PISA Consortium verified that all entries were correct, NPMs had one month to perform the same checks and to agree to the content in this form.

### **Sampling Task 12: School participation and data validity checks**

- Extensive checks were completed on Sampling Task 12 data since it would inform the weighting process. Checks were done to ensure that school participation statuses were valid, that student participation statuses had been correctly assigned, and that all student sampling data required for weighting were available and correct for all student sampling options. Quality checks also highlighted schools having only one grade with PISA-eligible students, only one gender of PISA-eligible students, or schools which had noticeable differences in enrolled student counts than expected based on sampling frame enrolment information. Such situations were queried.
- Large differences in overall grade and gender distributions compared to unweighted 2006 data were queried.
- These data also provided initial unweighted school and student response rates. Any potential response rate issues were discussed with NPMs if it seemed likely that a non-response bias report might be needed.
- Large differences in response rates compared to PISA 2006 were queried.
- Participating countries doing DRA were expected to have data for DRA related variables. Any expected DRA data entries were queried.

## **Student samples**

Student selection procedures in the main study were the same as those used in the field trial. Student sampling was generally undertaken using the PISA Consortium software, *KeyQuest*, at the national centres from lists of all PISA-eligible students in each school that had agreed to participate. These lists could have been prepared at national, regional, or local levels as data files, computer-generated listings, or by hand, depending on who had the most accurate information. Since it was important that the student sample be selected from accurate, complete lists, the lists needed to be prepared slightly in advance of the testing period and had to list all PISA-eligible students. It was suggested that the lists be received one to two months before the testing period so that the NPM would have adequate time to select the student samples.

Eight countries (Brazil, Chile, Germany, Iceland, Japan, Liechtenstein, Slovenia and Switzerland) chose student samples that included students aged 15 and/or enrolled in a specific grade (e.g. grade 10). Thus, a larger overall sample, including 15-year-old students and students in the designated grade (who may or may not have been aged 15) was selected. The necessary steps in selecting larger samples are noted where appropriate in the following details:

- Brazil, Iceland, Liechtenstein, Slovenia and Switzerland used the standard method of direct student sampling described here.
- For Iceland and Japan, the sample constituted a de facto grade sample because nearly all of the students in the grade to be sampled were PISA-eligible 15-year-olds.
- In the case of Iceland, the few additional grade 10 students in the country were added to the sample, so that there was a census of both PISA-eligible students and grade 10 students.
- Germany supplemented the standard sampling method with an additional sample of grade-eligible students which was selected by first selecting grade 9 classes within PISA sampled schools that had this grade.
- In Chile, the standard method was supplemented with additional grade-eligible students from a sample of grade 10 classes within PISA sampled schools that had this grade; Mexico also selected a grade 12 sample but accomplished this by having a completely separate sample of schools containing grade 12 students.

### **Preparing a list of age-eligible students**

Each school drawing an additional grade sample was to prepare a list of *age* and *grade-eligible* students that included all PISA-eligible students in the designated grade (e.g. grade 10); and all other 15-year-old students (using the appropriate 12-month age span agreed upon for each participating country) currently enrolled in other grades. This form was referred to as a student listing form. The following were considered important:

- Age-eligible students were all students born in 1993 (or the appropriate 12-month age span agreed upon for the participating country).



- The list was to include students who might not be tested due to a disability or limited language proficiency.
- Students who could not be tested were to be excluded from the assessment after the student sample was selected. It was stressed that students were to be excluded after the students sample was drawn, not prior.
- It was suggested that schools retain a copy of the student list in case the NPM had to contact the school with questions.
- Student lists were to be up-to-date at the time of sampling rather than a list prepared at the beginning of the school year. Students were identified by their unique student identification numbers.

### **Selecting the student sample**

Once NPMs received the list of PISA-eligible students from a school, the student sample was to be selected and the list of selected students (i.e. the student tracking form) returned to the school. NPMs were required to use *KeyQuest*, the PISA Consortium sampling software, to select the student samples unless otherwise agreed upon. Only Germany did not use the PISA Consortium software for selecting the student sample for reasons including extra student demographic data that could not fit in the available columns on the student tracking form produced by *KeyQuest*.

### **Preparing instructions for excluding students**

PISA was a timed assessment administered in the instructional language(s) of each participating country and designed to be as inclusive as possible. For students with limited assessment language(s) experience or with physical, mental, or emotional disabilities who could not participate, PISA developed instructions in cases of doubt about whether a selected student should be assessed. NPMs used the guidelines to develop any additional instructions; school co-ordinators and test administrators needed precise instructions for exclusions. The national operational definitions for within-school exclusions were to be clearly documented and submitted to the PISA Consortium for review before testing.

### **Sending the student tracking form to the school co-ordinator and test administrator**

The school co-ordinator needed to know which students were sampled in order to notify students, parents and teachers to update information and to identify students to be excluded. The student tracking form was therefore sent approximately two weeks before the testing period. It was recommended that a copy of the tracking form be kept at the national centre and the NPM send a copy of the form to the test administrator in case the school copy was misplaced before the assessment day. The test administrator and school co-ordinator manuals (see Chapter 6) both assumed that each would have a copy.

In the interest of ensuring PISA was as inclusive as possible, student participation and reasons for exclusion were separately coded in the student tracking form. This allowed for students with Special Education Needs (SEN) to be included when their SEN was not severe enough to be a barrier to their participation. The participation status could therefore detail, for example, that a student participated and was not excluded for SEN reasons even though the student was noted with a special education need. Any student whose participation status indicated they were excluded for SEN reasons had to have an SEN code that explained the reason for exclusion. It was important that these criteria be followed strictly for the study to be comparable within and across participating countries. When in doubt, the student was included. The instructions for excluding students are provided in the PISA Technical Standards.

### **Definition of school**

Although the definition of a “school” is difficult, PISA generally aims to sample whole schools as the first stage units of selection, rather than programmes or tracks or shifts within schools, so that the meaning of “between school variance” is more comparable across countries.

There are exceptions to this, such as when school shifts are actually more like separate schools than part of the same overall school. However, in some countries with school shifts this is not the case and therefore whole schools are used as the primary sampling unit. Similarly, many countries have schools with different tracks/programs but generally it is recommended again that the school as a whole should be used as the primary sampling unit. There are some exceptions, such as the schools being split for sampling in previous PISA cycles (trends would be affected if the same practice was not continued), or if there is a good reason for doing so (such as to improve previously poor response rates, differential sampling of certain tracks or programs is desired, etc).

Sampling units to be used on school-level frames have been discussed with each country before the field trial. Table 4.3 presents the comments from NPMs, in cases where “school” was not the unit of sampling. Where the Sampling Unit column indicates SFRUNITS, this means that the school was the sampling unit. Where it shows SFRUNITO then something else was used, as described in the comments. Table 4.3 shows the extent to which countries do not select schools in PISA, but rather something else.



[Part 1/2]  
Table 4.3 Sampling frame unit

	Sampling unit school / Other	Comment
Albania	SFRUNITS	
Argentina	SFRUNITO	Schools are indicated by location.
Australia	SFRUNITO	A very small percentage of schools in Australia have more than one campus. However, from experience, we have found not all schools with more than one campus are recorded as a separate record in the sampling frame (which is compiled from data provided by the federal and state governments).
Austria	SFRUNITO	We sample separately for programmes within each school, because different programmes award different certificates (the programmes vary in their ISCED levels).
Azerbaijan	SFRUNITS	
Belgium	SFRUNITO	Belgium: a combination of whole schools and "implantations" Flanders: "implantations" - Tracks/ programmes taught on a single address/location (administrative address) (same unit as used in PISA 2003 and PISA 2006). French and German Speaking Community: "whole schools" - pedagogical-administrative units, which may include different tracks, programmes, and which may include distinct geographical units (same unit as used in PISA 2003 and PISA 2006). One variation to the sampling in the previous PISA-cycles = the part-time vocational schools in the French Speaking Community. In contrast to the situation in PISA 2003 and PISA 2006 these schools are now no longer linked to a regular school so they are no longer automatically selected together with a corresponding regular school. In PISA 2009 the French part-time vocational schools will be considered as separate administrative schools.
Brazil	SFRUNITS	
Bulgaria	SFRUNITS	
Canada	SFRUNITS	
Chile	SFRUNITS	
Colombia	SFRUNITS	
Croatia	SFRUNITO	In Main Survey 2009 the sampling units will be school locations. Namely, there are some primary schools that are central but have their branch schools. These branch schools are situated on other locations but have the same school administration as their central school.
Czech Republic	SFRUNITO	Basic school - whole school special and practical school - whole school gymnasium - pseudo schools according to the length of study (4 year gymnasium and 6 or 8 year gymnasium) upper secondary vocational - pseudo schools (schools with leaving exam, schools without leaving exam).
Denmark	SFRUNITS	
Dubai (UAE)	SFRUNITO	Schools with mixed genders that have two separate campuses will be split into two schools with the same ID but differentiated with M for males and F for females.
Estonia	SFRUNITS	
Finland	SFRUNITS	
France	SFRUNITS	
Germany	SFRUNITO	Most schools will be sampled as whole schools. But some schools have different school types/ tracks together in the same school, there we would see the school types as implicit strata. In 2003 this was 20% of PISA students in the sample.
Greece	SFRUNITS	
Hong Kong-China	SFRUNITS	
Hungary	SFRUNITO	Tracks within schools which are located on different campuses. In the last few years Hungary has seen a transformation of schools into large institutes with multiple complexes/buildings, sometimes not even on the same campus and that often have a completely separate teaching staff.
Iceland	SFRUNITS	
Indonesia	SFRUNITS	
Ireland	SFRUNITS	
Israel	SFRUNITS	
Italy	SFRUNITS	
Japan	SFRUNITO	The units on our school sampling frame are programmes.
Jordan	SFRUNITS	
Kazakhstan	SFRUNITS	
Korea	SFRUNITS	
Kyrgyzstan	SFRUNITS	

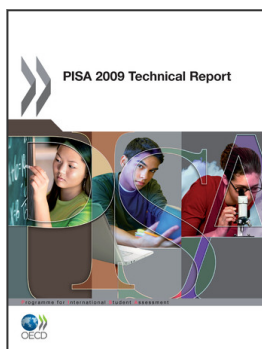
[Part 2/2]  
Table 4.3 **Sampling frame unit**

	Sampling unit school/ Other	Comment
Latvia	SFRUNITS	
Liechtenstein	SFRUNITS	
Lithuania	SFRUNITS	
Luxembourg	SFRUNITS	
Macao-China	SFRUNITS	
Mexico	SFRUNITS	
Montenegro	SFRUNITS	
Netherlands	SFRUNITS	Our proposed units are the same as in PISA 2003/2006: locations of (parts of) schools. These are often parts of a larger managerial unit.
New Zealand	SFRUNITS	
Norway	SFRUNITS	
Panama	SFRUNITS	
Peru	SFRUNITS	
Poland	SFRUNITS	
Portugal	SFRUNITS	
Qatar	SFRUNITS	
Republic of Moldova	SFRUNITS	
Romania	SFRUNITS	Since in Romania the combination of school programmes (GIM/SAM/LIC) represents an important characteristic of the school system organisation, the sampling frame for both FT and MS will be by school programme.
Russian Federation	SFRUNITS	
Scotland	SFRUNITS	
Serbia	SFRUNITS	
Shanghai-China	SFRUNITS	
Singapore	SFRUNITS	
Slovak Republic	SFRUNITS	
Slovenia	SFRUNITS	The preferred approach to sampling in Slovenia is by study programme. Many programmes share the same school building, however they operate largely independently from each other, sometimes even having different school principals, and in most cases a vice-principal for each programme. 15-year-olds attend ISCED 2 or ISCED 3 programmes. ISCED 2 is a part of compulsory elementary general education (only 5% of 15-year-olds). Upper secondary education (ISCED 3) consists of the following school types - study programmes: GIMg = gymnasia general; GIMs = gymnasia specialist; STSI = technical educational programmes; SPI = vocational of medium duration; NPI = vocational of short duration. Due to the conceptual differences between the ISCED 3 study programmes it is essential for the quality of national analysis to sample students separately from each of the programmes.
Spain	SFRUNITS	Whole School is the option selected for Spain. Only in the Basque Country (5% of Spanish population) the same school can be divided into three, one for each linguistic model (A, B, D).
Sweden	SFRUNITS	
Switzerland	SFRUNITS	
Chinese Taipei	SFRUNITS	
Thailand	SFRUNITS	
Trinidad and Tobago	SFRUNITS	
Tunisia	SFRUNITS	
Turkey	SFRUNITS	
United Kingdom (excl. Scotland)	SFRUNITS	
United States	SFRUNITS	
Uruguay	SFRUNITS	



## Notes

1. Students were deemed participants if they gave at least one response to the cognitive assessment, or they responded to at least one student questionnaire item and either they or their parents provided the occupation of a parent or guardian (see Annex G).
2. Available at [www.pisa.oecd.org](http://www.pisa.oecd.org) > what PISA produces > PISA 2009 > PISA 2009 manuals and guidelines.
3. These include schools with multiple languages of reading instruction, vocational schools, technical schools, agriculture schools, and schools with only part-time students, schools with multiple shifts and so on.



**From:**  
**PISA 2009 Technical Report**

**Access the complete publication at:**  
<https://doi.org/10.1787/9789264167872-en>

**Please cite this chapter as:**

OECD (2012), "Sample Design", in *PISA 2009 Technical Report*, OECD Publishing, Paris.

DOI: <https://doi.org/10.1787/9789264167872-5-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).