Chapter 5: SAMPLE DESIGN

INTRODUCTION

This chapter presents information about the PISA-D Strand C Main Survey (MS) sample design and selection plans. The method of data collection involved in-person interviews, mainly in households or other similar locations. Generally, costs associated with in-person interviewing are much higher than those of school-based administration of tests and questionnaires, especially when the target population is a rare and geographically disperse group, as is the case for PISA-D Strand C. Therefore, a critical component of this pilot project was to arrive at sampling plans that minimised costs to countries, while ensuring the coverage and reporting goals of PISA-D Strand C were met.

The international contractor (Westat) submitted to countries a general sampling plan, a summary of options, and for the recommended option, a sample size worksheet and a sample design summary. These documents were meant to serve as guidance so that final country-specific MS sampling plans could be developed after individual consultations with the countries.

The target population for PISA-D Strand C consists of 14 to 16-year-olds who are out of school or in school but in grade 6 or below. In the first section of this chapter we provide more detail on the PISA-D Strand C target population and give an overview of the PISA-D Strand C sample design for the representative and limited representative samples. We then present country-specific information, such as sources of country sampling frames and their coverage of the target population; definition of sampling units and sample selection methods; stratification; national sample designs; descriptions for the limited representative sample; sampling quality control checks; and respondent incentives. We also provide information about lessons learnt and in the last section we give recommendations for possible future cycles of PISA involving this out-of-school population.

TARGET POPULATION AND OVERVIEW OF THE PISA-D STRAND C SAMPLE DESIGN

A clear and precise definition of the target population is necessary to ensure that the population of interest is adequately covered by each participating country and to maintain consistency and comparability across countries. The target population for PISA-D Strand C consists of 14- to 16- year-olds who are out of school, or in school but grade 6 or below. The target population definition for the PISA-D Strand C population, combined with that of the Strand A/B target population (15- year-olds in grade 7 or above), is close to covering the entire population of 15- year-olds in a country. Although references are made to 15-year-olds for the PISA-D Strand C population, the age range for the sample was broadened to include 14- to 16-year-olds to increase the size of the target population. The term “out-of-school 15-year-olds” continues to be used throughout this report in reference to the PISA-D Strand C target population, even though it is technically 14- to 16-year-olds and includes in-school youth in grade 6 or below.

The general sampling plan served as the basis for national samples and followed the two main objectives for the PISA-D Strand C data collection:

- To yield a sample size large enough to evaluate the psychometric characteristics of items within each of the participating countries, and in the aggregate determine whether linking to PISA and PISA-D Strand A is possible.
- To explore various approaches and evaluate various options to arrive at a recommendation for identifying and assessing a nationally representative sample of out- of-school 14- to 16-year-olds in future cycles of PISA, possibly beginning with the PISA 2021 assessment cycle.

The challenge with selecting a representative sample of out-of-school 15-year-olds is that a substantial amount of screening is required to locate this relatively small and dispersed subgroup of the population. When designing the sample, the international contractor (Westat) drew from the lessons presented in the 2014 International Advisory Group workshop in Paris, France and the 2014 Technical Workshop on out-of-school 15-year-olds in Montreal, Canada, as well as from experiences relating to surveys conducted on out-of-school youth through the joint UNESCO Institute of Statistics (UIS)/UNICEF Global Initiative on Out-of-School Children (OOSCI) initiative. The following general sample design was aimed at reducing the cost of locating and interviewing the required number of out-of-school 15-year-olds, while at the same time evaluating various options for developing representative samples for future cycles of PISA-D.

To satisfy the above objectives, the international contractors developed general guidance toward a core sample design consisting of two major strata defined as high and low concentration strata of the target population and two components—representative and limited representative— within each stratum. The core guidance was to design the sample to have two strata: i) stratum H, including the areas of the country that have high concentrations of out-of-school youth (e.g. rural areas, urban slums, etc.); and ii) stratum L consisting of areas with low concentrations. The creation of the two strata facilitated the use of variable sampling rates across strata. Higher sampling rates were used in stratum H to reduce the cost of locating out-of-school 15-year-olds in the probability sample.

The stratified sample was designed to arrive at a minimum of 1 600 completed cases. In general, a completed case is one that contains responses to key background questions and a sufficient number of assessment items. The completed cases for analysis include all cases that received a final weight and are included in the analysis file. They consist of youth that completed the Youth Interview (YI) and those who could not respond to the Youth Interview for a literacy-related reason (language barrier, or learning or mental disability).

The 1 600 was split into at least 1 200 from a representative (R) sample and at most 400 from a limited representative (LR) sample. The R sample, described in a later section, would adequately represent the target population and would consist of at least 600 cases from probability-selected households (seeds), with the remaining cases coming from one or two waves of link-traced cases.

In addition to the 1 200 R sample cases, at most 400 cases could be added to help boost sample size in a cost-efficient manner. These 400 cases could be from an LR sample or from the R sample as described above. Both the R and LR samples could be included in the psychometric testing, as well as used to represent the target population. Ideally, the LR cases would be selected in a way to allow representation of a limited portion of the target population, but at minimum the LR sample could be weighted to be self-representing only. The LR options are described in a later section.

Representative sample

The Representative (R) sample component was to be comprised of a probability sample with an option to collect data through link-tracing from the probability-based households.

Probability sample

Applying probability sampling with higher rates in the high concentration areas can greatly reduce the cost of locating out-of-school 15-year-olds. The trade-off is an increase of the sampling variance for estimates from the sample. Generally, the most efficient sample design for a household survey with in-person interviews is a self-weighted multi-stage stratified cluster design, but the sampling plan was meant to be flexible and adaptive to each country's best sampling scenario.

In Strand A, the sample designs were essentially standard, where a sampling frame of schools was used to select schools, and then students within schools were selected following standard procedures. For household studies, sampling frames are sometimes not readily available or are more difficult to create. Because of these differences, sampling approaches needed to be variable across participating countries. Therefore, prior to specifying the details of the design, the international contractor (Westat) gathered relevant input from the national authorities in participating countries. At the same time, sampling statisticians assigned to PISA-D Strand C communicated with National Project Managers (NPMs) to gather information about the country's approaches to probability sampling and to gain an understanding of the geographic spread to identify areas with large concentrations of out-of-school 15-year-olds. The final sample design was tailored to national requirements and constraints while satisfying the comparability requirements of PISA-D Strand C.

In some countries, Census Enumeration Areas (EAs) were already formed and had information available that could be useful for sampling purposes. We consulted with the national statistical institutes about their census data and proposed an approach that was tailored to each country's unique situation.

Once Primary Sampling Units (PSUs) or area clusters were selected, either a mini-census was conducted to list persons in the target population in the areas, or a listing of dwelling units (DUs) or collective dwelling units was conducted. Variations were expected since the type of sampling frames within countries could range from having person or address registries to not having any complete lists for use in area sampling. The general approach was to allow for flexibility in the

sample design, conduct a thorough assessment of the quality of sampling frames through a traditional approach and prepare to be adaptive to each country's situation.

In general, conditions that are conducive to conducting a mini-census include small-sized PSUs (around 50 DUs), no DU list to sample from and outdated maps. Conditions conducive to selecting DUs (with probability less than one) and conducting link-tracing, are when a list of addresses exists or when a listing (with selection of DUs and link-tracing) can be considered a more affordable option than a mini-census.

The key to success in sampling from households is the construction of a DU frame of high quality. To achieve a high-quality frame of high coverage, the list of DUs must be accurate and current. If a list of addresses is not available, then a listing procedure can be conducted. As an alternative, maps with indications of existing DUs can be used.

The sample design options for the R sample were provided as follows. Once the DU frame is established, within PSU selection rates for DUs should be assigned such that the overall DU probability of selection (P_{hij} for DU j in PSU i of stratum h) is equal within stratum h (H or L).

$$P_{hij} = P_{hi} \times P_{j|hi,}$$

where,

P_{hi}= probability of selection of the PSU i in stratum h

$P_{j|hi}$= probability of selection for DU j conditional on PSU i in stratum h (H, L) being selected.

For a mini-census, it is optimal to select an equal probability sample of PSUs within stratum h such that

$$Phij = Phi \times Pj|hi$$

$$= P_{hi} \times 1 = n_h/N_h$$

where n_h is the number of PSUs selected in stratum h and N_h is the number of PSUs in stratum h. An improvement is to select an equal probability systematic sample from a sorted list of PSUs (sorted on a variable related to the survey outcome), which would have a progressively decreasing impact on resulting sample variances.

If a sample of DUs is conducted in lieu of a mini-census, then PSUs should be selected proportionate to size within each stratum. One consideration is to purposively select a small percentage of PSUs on the sampling frame with P_{hi} = 1 (those with a very high concentration of the target population), and then remove them from the sampling frame for the selection of the rest of the PSUs. With a mini-census (or a probability-based sample of DUs) in each PSU, this results in the entire sample aligned as a probability sample. While this approach may reduce cost of screening, it introduces more variation into the survey results as the number of purposively-selected PSUs increases.

The planned sample design for each country was such that DUs are assigned rates to result in overall equal probabilities of selection within each stratum. To do so, PSUs were assigned the following probabilities:

$$P_{hi} = N_h \times X_{hi} / \sum X_{hi}$$

where, X_{hi} = the measure of size for PSU i. For example, the measure of size could be assigned as the size of the target population, if available.

Then DU probabilities were assigned as:

$$P_{j|hi} = r_h / P_{hi}$$

where r_h = the overall rate for the DUs, that is, the overall number of DUs to select within stratum h divided by the total population of DUs.

Then the overall probability of selection would be:

$$P_{hij} = P_{hi} \times P_{j|hi}$$

$$= P_{hi} \times r_h / P_{hi}$$

$$= r_h$$

A screener questionnaire was used to list persons within the dwelling, to identify persons in the target population, and to select the eligible youth for interview and assessment. Once selected during the screening, the ideal scenario was to interview and assess the eligible youth in the same visit, if possible.

Link-tracing sample

Nonprobability samples are used especially in Field Trials to capture data at a substantially lower cost than their probability-based counterpart. The disadvantage of nonprobability designs is that sampling theory does not hold for making generalisations to the population. Examples of purposive designs are respondent-driven sampling (RDS) (Heckathorn, 2007) and snowball sampling (Goodman, 1961). Under some strict assumptions, Heckathorn and Goodman claim that RDS and snowballing can produce national estimates. However, these assumptions are rarely satisfied in real-life situations. Therefore, a probability-based link-tracing (PBLT) approach was developed to reach the required number of completed cases in PISA-D Strand C while producing a representative sample.

As mentioned above, the requirement for the R sample consisted of at least 600 cases from probability-selected households (seeds). The remaining cases in the R sample were to come from one or two waves of link-traced cases.

In the first wave, referrals from the probability-selected household could be used as an efficient way to increase the sample yield at low cost. That is, after the Youth Interview, or after the screener if the respondent is not an eligible youth, the respondent would be asked for location information about two eligible youth in the neighborhood, which we refer to as referrals. In the first wave, link-tracing was triggered by the Youth Interview or screener questionnaire in all selected and cooperating households from the probability component, whether or not they included any individuals in the PISA-D Strand C target population, to gather information needed to locate and interview other eligible individuals in the area. The goal was to identify eligible youth within the same cluster. However, the respondent was likely unfamiliar with cluster boundaries and so could be asked about the neighbourhood. The relevant screener questions are shown below in Box 5.1.

Information about each referral was to be sent to the home office for reconciliation with other household seeds or other referral cases and eliminate potential duplicate cases, and then referrals that passed the home office process would be assigned to the interviewer.

The link-tracing could be expanded into a second wave through the use of recruitment, as described in the next section. Limiting link-tracing to two waves had the advantage of reducing the amount of duplication, in other words, reducing the number of times that the same youth was identified through multiple sources. It also decreased the number of cases obtained through link-tracing so that the sample yield did not exceed available resources. The disadvantage of limiting to two waves was that it was less likely that full coverage of the target population would be obtained in the sampled area.

Box 5.1 Referral questions from the PISA-D Strand C screener

Do you know any {other} 14 - 16 year olds who are out of school or at grade 6 or below in your neighborhood?
[HELP SCREEN: BY NEIGHBORHOOD WE MEAN {COUNTRY DEFINITION}]
YES
NO [GO TO END]
980 DON'T KNOW [GO TO Q.END]
990 REFUSED [GO TO END]
I need to collect their name and address so we can contact them to participate in the study.
Please tell me the youth's name _____.
Please give me their address or directions to where the youth lives.
_____
Any others?
RECORD NAME AND ADDRESS FOR EACH REFERRAL (UP TO TWO) ON SEPARATE BLANK HOUSEHOLD FOLDERS
980 DON'T KNOW
990 REFUSED

In the second wave, referrals would be provided five coupons to hand to others (called recruits) that they knew in the target population. The recruits then would call the home office to arrange an appointment. Countries were encouraged to provide incentives for the referrals and the recruits. With a small number of referrals in an enumeration area, there could be more potential for success in a second wave of recruiting as youth may be more comfortable with recruiting

rather than referring others to interviewers. A similar process to identify duplicates and eliminate them, as above, was applied to the recruits sample.

Because the link-tracing waves were linked to a probability-based household, the process is referred to as representative.

Limited representative sample

The purpose of the limited representative (LR) sample was to add at most 400 completed cases in a cost-efficient manner. Ideally, the allocation between out-of-school youth and students at or below grade 6 were to be aligned with their proportions in the population. Four LR options are described below.

School administrative (SCAD) records of students at or below grade 6

This option refers to the selection of schools to identify students at or below grade 6 and the subsequent selection of a sample from the identified students. The schools needed to at least offer grades 5 and 6, and were asked to provide a list (including address and other contact information) of all 14- to 16-year-old students attending their school in grade 6 or below. This would become the relevant sampling frame for selection of 14- to 16-year-olds in grades 6 or below.

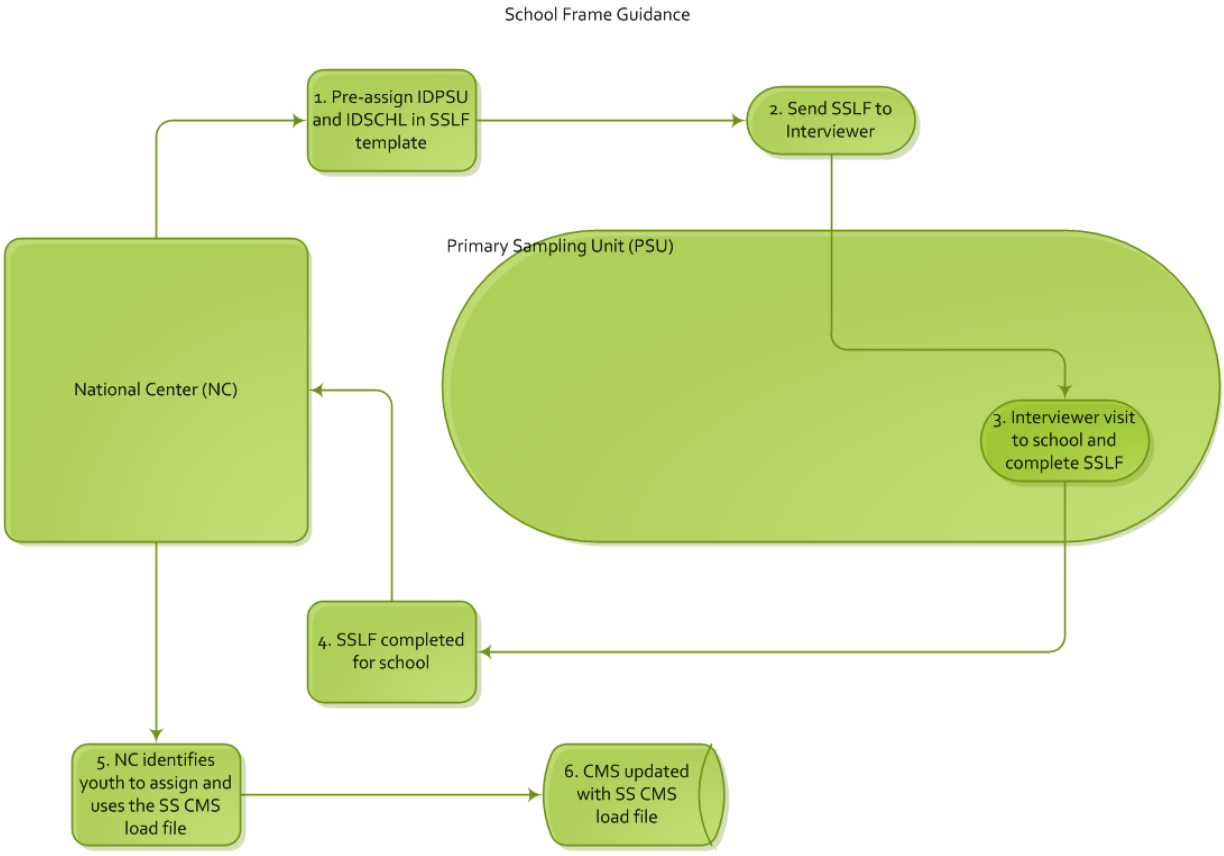
The School Sample List Form (SSLF) was used to capture the contact information for potentially eligible youth within schools, as part of the LR sample component.

The general steps were as follows:

- National Centre (NC) provides the SSLF template to interviewers. It was beneficial for the NC to pre-assign and pre-fill the IDPSU (start with 001) and IDSCHL (start with 901 within each IDPSU) columns before providing to interviewers.
- Interviewers contact the school and complete the SSLF and send back to the NC.
- The NC then identifies the youth to assign to interviewers. Not all youth need to be assigned in order to balance the number of completed cases with the probability sample component.
- The NC then completes the *School Sample CMS Load File* template to help load the assigned school frame cases into the Case Management System (CMS).

These steps are shown graphically in Figure 5.1 below.

Figure 5.1 Flow of sample frame creation from school frame approach



Link-tracing from students and teachers (LTST)

The teachers and/or students in the same schools, as mentioned above, were asked to provide the names of out-of-school youth in the neighborhood. This was done by administering a short questionnaire in the classrooms. The referred-to-youth were checked by the home office for duplication and then assigned to be interviewed through the process described above using the SSLF.

Location sampling (LOCA)

In this option, a location and time was advertised for youth to meet to conduct the interview and administer the assessment. A more sophisticated approach is following the time and space sampling as described in Kalton (1991). However, for PISA-D Strand C, eligible youth were asked to meet at a specific location or interviewers went to potentially relevant locations to recruit youth, such as local programmes that serve youth in the target population.

Street children

Another useful LR sample approach was to select street children under the LR component. Sampling street children was encouraged for gathering information on the group; but their inclusion in the sample depended on their distribution and proportion in each country.

The link-tracing approach was useful in identifying youth who resided in places other than households. While, another identification approach based on special operations conducted in known high-concentration locations was also used.

Population coverage and participation rate standards (R sample)

In the Main Survey, a large portion of the sample had to come from a probability-based sample. Furthermore, quality standards had to be maintained with respect to i) the coverage of the PISA-D Strand C international target population; ii) accuracy and precision; and iii) response rates.

Coverage of the PISA-D Strand C international target population

The sampling frame needed to include 95% or more of the core PISA-D Strand C target population. That means, the undercoverage rate, combined over all stages of sampling, was not to exceed 5%.

Accuracy and precision

As mentioned above, the minimum sample size requirement was 1 600 completed cases with at least 1 300 that pass the core cognitive assessment module and continue to take the main assessment items. With a total of 1 600 completed cases, there would be an average of 650 responses per main assessment item. The 1 600 was split into at least 1 200 from a representative sample and at most 400 from a limited representative sample. Countries with multiple languages needed to achieve a minimum complete sample size of 1 600 respondents for their main language, with additional completed cases needed as proportionate to the population speaking other languages.

Response rates

A minimum overall response rate of 70% is the standard. The overall response rate is computed as the product of the response rates for the stages included in data collection. All response rates must be weighted by the household base weight, in the case of a screener response rate, and by the person base weight, in the case of a BQ or assessment response rate.

MAIN SURVEY SAMPLE

Definition of the national target population

The definitions of the international and national target populations are identical. The operational definition of an age population directly depends on the testing dates. For the representative sample, age was determined when the screener was conducted. For the limited representative sample, age was determined at the time of the Youth Interview.

Sampling frames and their coverage (R sample)

Sampling frames

The sampling frame is the source from which the sample is selected at the given stage of sampling, and so the quality of the sampling frame affects the quality of the sample.

Therefore, sampling frames must meet a minimum set of quality standards to ensure that adequate and accurate information is available for carrying out sampling, data collection, weighting and nonresponse bias analyses. It is also important that exclusions be clearly specified and limited as much as possible so that no extensive biases are introduced as a result of undercoverage of the population. Exclusions need to be documented thoroughly and transparently to assess the representativeness of the sample in PISA-D Strand C.

Multi-stage sample designs require a sampling frame for each stage of selection. See Table 5.1 for the full list of sampling frames employed by countries.

Table 5.1 Sampling frames

Country	Stage 1	Stage 2	Stage 3
Guatemala	List of areas from the National Census, 2002	List of dwelling units (DUs) from the National Census, 2002	Field enumeration
Honduras	Enumeration areas (EAs) from Statistics National Institute (INE) Population and Household National Census, combined to meet minimum size, 2013	List of dwelling units (DUs) in the EAs based on a canvas of the area, 2018	Field enumeration
Panama	Primary Sampling Units (PSUs) from the National Institute and Statistics (INEC) of Panama Census of Population and Housing, 2010	Addresses drawn from the cartographic information used for the Census of Population and Housing, 2010	Field enumeration
Paraguay	Updated cartographic information provided by the Direccion General de Estadisticas, Encuestas, y Censos of Paraguay, 2012	List of all dwelling units (DUs) in the enumeration area, based on a canvas of the area, 2018	Field enumeration
Senegal	List of census districts from the General Population and Housing Census, 2013	List of dwelling units (DUs) from the General Population and Housing Census, 2013	Field enumeration

Undercoverage of the target population

As mentioned earlier, the undercoverage rate for PISA-D Strand C, combined over all stages of sampling, may not exceed 5% (standard 1.3), and thus the sampling frames for each country were required to include 95% or more of the standard PISA-D Strand C target population. All exclusions to the PISA-D Strand C target population, whether or not they exceed the threshold, were reviewed by the contractor. Exclusions are acceptable only if they occur because of operational or resource considerations such as excluding persons in hard-to-reach areas. The international contractor (Westat) asked that each country identify, to the extent possible, exclusions before sample selection. Adjustments for any undercoverage of the target population in each country

were made through benchmarking during the weighting process. A list of exclusions is presented in Table 5.2. Note the undercoverage rate in the table accounts for excluded subpopulations such as mobile, nomadic and/or pastoralist populations. Other exclusions that will occur as a natural part of the survey process are not included in the expected undercoverage rate.

Table 5.2 Portion of target population not covered by Main Survey sampling frames

Country	Percentage of target population not covered	Group not covered
Guatemala	No data available	Mobile, nomadic, and/or pastoralist populations
Honduras	5%	Remote areas, not easily accessible
Panama	0.0%	No exclusions from the frame for rural and indigenous areas
Paraguay	2.18%	Population located in two departments: Boquerón, Alto Paraguay
Senegal	No data available	Mobile, nomadic, and/or pastoralist populations

Sampling units and sampling selection methods (R sample)

Details regarding sampling units and sampling selection methods are presented in Tables 5.3 and 5.4, respectively.

Table 5.3 Main Survey sampling units

Country	Stage 1	Stage 2	Stage 3
Guatemala	PSUs – Census districts	DUs	Youth
Honduras	PSUs – Enumeration areas or combined enumeration areas	DUs	Youth
Panama	PSUs	DUs	Youth
Paraguay	PSUs – Enumeration areas or combined enumeration areas	DUs	Youth
Senegal	PSUs – Census districts	DUs	Youth

Table 5.4 Main Survey selection methods

Country	Stage 1	Stage 2	Stage 3
Guatemala	Probability proportionate to size (number of DUs) from a sorted list within explicit strata	Census	Take all likely eligible youth
Honduras	Probability proportionate to size within the municipalities	Census	Take all likely eligible youth
Panama	Systematic probability proportional to size from a sorted list of PSUs	Systematic probability from a sorted list of DUs within the selected PSUs	Take all likely eligible youth
Paraguay	Systematic probability proportional to size from a sorted list within explicit strata. Select 60 PSUs that have high proportion of eligible population with certainty before selecting the other PSUs	Census	Take all likely eligible youth
Senegal	Probability proportionate to size (estimated number of 14- to 16-year-olds out of school) from a sorted list within explicit strata	Census	Simple random sample of one-third of likely ineligibles among 14- to 16-year-olds as programmed in the tablet.

Stratification (R sample)

Stratification combines sample units into homogeneous groups and reduces sampling variability between such groups, thus reducing the overall sampling variance associated with the resulting survey estimates. To maximise the benefit of stratification, stratification variables should be reliable and related to the survey outcome. The stratification and/or sorting variables are shown in Table 5.5.

Table 5.5 Main Survey stratification/sorting variables by country

Country	Description
Guatemala	Major strata (high vs. low concentration): High: Rural areas; Low: Urban areas Minor strata: Socioeconomic status (4 levels)
Honduras	Major strata (high vs. low concentration): High: Municipalities in rural areas that have a population concentration greater than or equal to 36%; Low: Municipalities in rural areas that have a population concentration less than 36% Within strata: sort by concentration

Country	Description
Panama	Major strata (high vs. low concentration): High: Indigenous areas; Low: Rural areas Within strata: sort by district, percentage of the 14-16 year old population out-of-school or in grade 6 and below, education level
Paraguay	Major strata (high vs. low concentration): High: administrative districts with proportion of PISA-D Strand C population one or more standard deviations above the mean; Low: administrative districts with proportion of PISA-D Strand C population less than 1 standard deviation above the mean Minor strata: PSU urbanisation (urban and rural) Within strata: sort by PSU geographic order by the Dirección General de Estadísticas, Encuestas y Censos
Senegal	Major strata (high vs. low concentration): High: seven regions with expected hit rate of 31% or more; Low: seven regions with less than 31% expected hit rate (Hit rate is defined as the number of sampled dwellings units required to obtain one completed assessment.) Minor strata: Regions Within strata : sort by Département/Arrondissement/Commune_Communauté rurale/Quartier village, likely eligible, likely ineligible

National sample designs

The international contractor (Westat) submitted to countries a general sampling plan, a summary of options, and for the recommended option, a sample size worksheet and a sample design summary. These documents were meant to serve as guidance so that final country-specific MS sampling plans could be developed after individual consultations with countries.

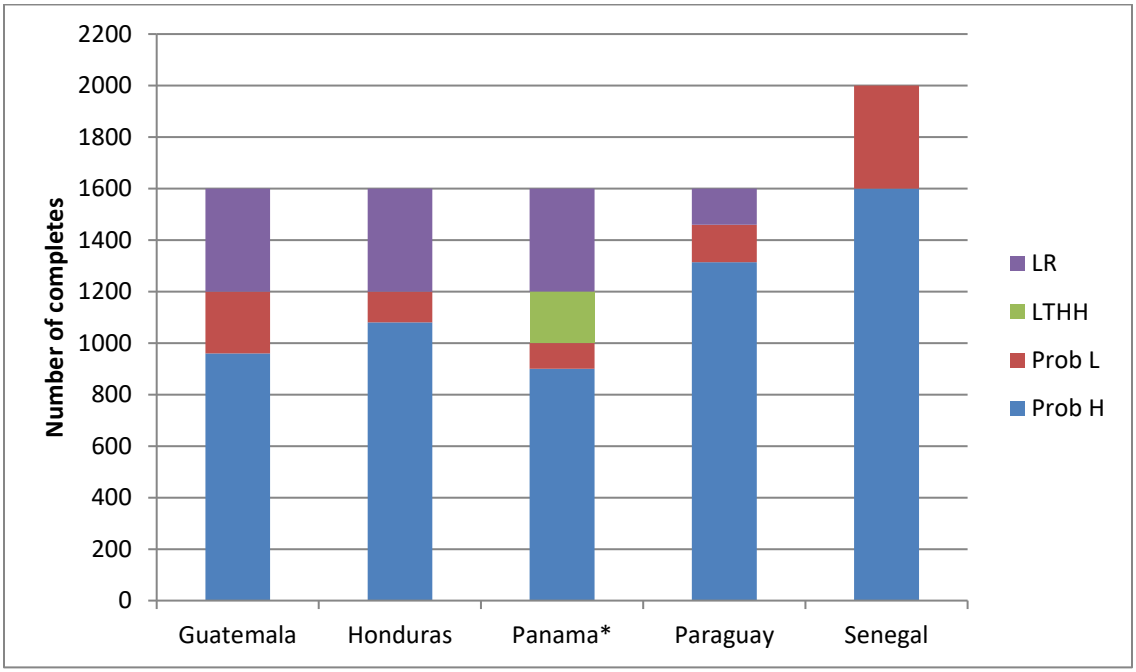
Table 5.6 shows information about the target sample sizes by sample type for each country. For Guatemala, Honduras and Panama, the representative sample was meant to yield 1 200 completed cases with the remainder 400 coming from the limited representative sample. For Paraguay, the repartition between the representative and limited representative samples was planned to be 1 460 and 140 respectively. Senegal elected to have a representative sample only with an expected number of completed cases of 2 000 accounting for two languages (Wolof and French). Figure 5.2 shows each country's target number of completed assessments for the representative and limited representative samples by major strata (i.e. high and low concentration of target population). Table 5.7 provides more details on the expected sample sizes. The rates used in this table were informed by Field Trial (FT) results.

Table 5.6 MS Target sample sizes by sample type

Country	Representative probability	Representative - nonprobability		Limited representative			
		Link-tracing through households	Link-tracing through households recruiting	School frame approach for out-of-school youth (OOS)	School frame approach for grade 6 or below	Location sampling	Special operation for street children
Guatemala*	1 200	NA	NA	230	120	NA	50
Honduras	1 200	NA	NA	NA	40	360	NA
Panama	1 000	200		100		300	NA
Paraguay	1 460	NA	NA	NA	140	NA	NA
Senegal	2 000	NA	NA	NA	NA	NA	NA

Note: * Guatemala implemented location sampling near the end of their data collection period because of a shortfall of expected cases from the school frame approach for OOS. A special operation for street children was planned, but not implemented.

Figure 5.2. Target number of completed assessments for MS, by country



Note: Prob H = representative probability sample high concentration stratum; Prob L = representative probability low concentration stratum; LTHH = link-tracing through households; LR = limited representative sample.
* Stratification for Panama was based on Indigenous/Rural areas rather than high/low concentrations of target population.

Table 5.7 Target sample sizes, expected eligibility rates and expected response rates for the MS representative probability sample, by country and major strata

Target counts and expected rates	Guatemala			Honduras (rural only)			Panama			Paraguay			Senegal		
	Total	H	L	Total	H	L	Total	Indigenous	Rural	Total	H	L	Total	H	L
Number of PSUs	101	59	42	217	169	48	531	414	117	488	386	102	80	56	24
Number of sampled dwelling units	20 621	10 210	10 411	19 019	14 865	4 154	17 163	13 454	3 709	24 880	19 675	5 205	8 020	5 363	2 657
Occupancy rate	81%	81%	81%	86%	86%	86%	89%	89%	89%	78%	78%	78%	98%	98%	98%
Screener response rate	80%	82%	70%	80%	80%	80%	80%	80%	80%	88%	88%	88%	84%	83%	84%
Average household size	5.0	6.0	4.0	4.4	4.4	4.4	3.8	3.8	3.8	3.9	3.9	3.9	10	10.5	9
Age eligibility rate (14-16 years old)	6.0%	6.0%	6.0%	5%	5%	5%	4.6%	10.1%	5.2%	6.3%	6.7%	6.3%	6.0%	6.0%	6.0%
Enrolment-related eligibility rate (OOS or grade 6 or below)	51%	70%	30%	65%	68%	27%	15%	30%	24%	28%	51%	23%	56%	79%	48%
Number of eligible sampled persons	2 134	1 709	425	1 697	1 528	170	1 232	1 109	123	1 991	1 792	199	2 738	2 171	567
YI response rate	71%	71%	67%	80%	80%	80%	92%	92%	92%	84%	84%	84%	94%	94%	90%
Assessment response rate	92%	92%	98%	98%	98%	98%	98%	98%	98%	97%	97%	97%	98%	98%	98%
Number of completed assessments	1 200	960	240	1 200	1 080	120	1 000	900	100	1 460	1 314	146	2 000	1 600	400

Limited representative sample

Different approaches to the LR sample were taken across countries. The following paragraphs discuss each country approach separately.

- **Guatemala** selected a convenience sample of 47 schools in the vicinity of probability-selected clusters (16 schools in the high concentration stratum and 31 in the low concentration stratum) using a frame from the Ministry of Education. Since a mini-census was conducted for the probability sample, schools were chosen far enough from the selected clusters to avoid duplication. Schools were also chosen to ensure representativeness across geographic and demographic characteristics. The school sample was expected to yield 120 completed cases for youth in grade 6 or below identified through school administrative lists; and 230 completed cases for out-of-school youth identified through the link-tracing referral approach by teachers or students in the selected schools. A special operation for assessing street children was also planned with an expected yield of 50 completed cases, but not implemented. Guatemala implemented location sampling near the end of their data collection period because of a shortfall of expected cases from the school frame approach for out-of-school youth (OOS). They identified non-sampled urban areas with a high concentration of the PISA-D Strand C target population.
- **Honduras** selected 17 schools in Tegucigalpa where they expected to find about 80 eligible students for an expected yield of 40 completed cases. They did not conduct the link-tracing referral approach by teachers or students in these schools. Honduras conducted data collection for the location sample also in Tegucigalpa. This approach was meant to identify around 400 eligible cases for an expected yield of 360 completed cases. Both of these approaches were intended to cover urban areas, but in the end Honduras was able to implement this in Tegucigalpa only.
- **Panama** conducted surveys of students attending school in urban areas. They targeted 100 completed cases through school administrative records, but were not able to implement data collection through the link-tracing referral approach by teachers or students. They were not able to conduct any location sampling in the urban areas, so they were not able to obtain the expected 300 completed cases.
- **Paraguay** selected 35 schools initially and expected to find 449 eligible students. They later excluded 6 schools due to either the schools being in the selected PSUs of the representative sample or inaccessibility. This approach was expected to yield 140 completed cases.

Sampling quality control checks

The international contractor (Westat) developed a comprehensive set of quality control checks to ensure that standards were followed so that survey results would be comparable in quality across countries. A final sample design summary was put together by the international contractor (Westat) and submitted to each respective country for review. The sample design summary was

one of the four submissions relating to quality control (QC) of the PISA-D Strand C sample design and selection. The other submissions that countries needed to provide were:

- sample selection quality control forms
- sample design international file (SDIF).

The sample selection quality control forms were to be submitted at the end of each stage of selection to help the international contractor (Westat) verify that the process was conducted accurately. Countries were also asked to conduct checks on the Survey Control File (SCF) after dwelling units were selected (or after PSUs were selected in the case of a mini-census) prior to uploading the data to the CMS. The sample was monitored on a weekly basis through the CMS and throughout data collection in order to identify potential shortfalls, problems in achieving the desired response rate and the potential for nonresponse bias in the collected sample. The SDIF was due after data collection and contained sampling information such as selection probabilities and linking IDs, which was used by the international contractor (Westat) to perform a final QC check on the sample. This file also served as the input file for weighting and nonresponse bias analysis.

Respondent incentives

Respondent incentives have been shown to be effective for improving response rates without affecting the respondent's performance. As a result, the use of incentives can potentially reduce bias in the estimates. As such, countries were permitted to offer modest incentives to obtain respondent cooperation, such as a monetary or nonmonetary incentive (e.g. pen, notepad, candy, mug, voucher or gift certificate). Three countries offered some form of incentive which is detailed in Chapter 7 of this Technical Report.

LESSONS LEARNT

The Field Trial and Main Survey of the PISA-D Strand C pilot provided the opportunity to conduct two experiments to better understand the data collection challenges involved in screening households to identify PISA-D Strand C eligible youth. The following provides brief descriptions of the experiments and their results.

Usual-residence/Slept-at-residence rules

Two alternative approaches are commonly used in listing household members during the screening stage in area surveys, usual-place of residence and slept-at-residence. An evaluation of the two rules was conducted to arrive at a recommendation for future cycles. The usual-place of residence rule is widely used and is less susceptible to giving individuals chances of selection from multiple households. The slept-at-residence rule is more susceptible to chances of selection from multiple households. Counts of usual residents and visitors (those who slept at residence but is not the usual residence) are provided in Table 5.8 for completed YIs. For example, for Guatemala, if the usual place of residence is used, among selected households with completed screeners, only four individuals completed the YI who did not sleep at the residence with a completion rate of only seven per cent. Higher completion rates were observed among the other

four countries. In contrast, in Guatemala there were 37 completed YIs who were visitors, and 90 per cent of visitors completed the survey among all who were visitors. This compares to 71% of the sampled youth who completed the survey overall. The other four countries showed similar results. Therefore, it seems advantageous to consider the slept-at-residence rule in order to maximise the number of completed YIs. However, because all interviews are not completed in the same day, employing the slept-at-residence rule is susceptible to allowing individuals multiple chances of selection. This aspect needs to be further examined in future cycles.

Table 5.8 Counts of complete YIs for each country by Usual-residence/Slept-at-residence status

Country	Total Slept at residence	Total Usual residents	Usual residents and Did not sleep at residence	Percentage of completes among Usual residents who Did not sleep at residence	Visitors (Not Usual residence but Slept at residence)	Percentage completes among all Visitors
Guatemala	1 254	1 221	4	7	37	90
Honduras	1 165	1 129	7	27	43	81
Panama	1 937	1 858	13	100	92	100
Paraguay	811	802	3	13	12	57
Senegal	2 094	2 050	11	25	55	87

Evaluation of the screener questionnaire's effectiveness in identifying eligibility

The Field Trial included an experiment to evaluate whether eligibility status was accurately reported by the household at the time of screening (since the head of household was likely to be someone other than the eligible youth). The main concern was the amount of false negatives, that is, eligible youth screening out because the head of household reported them as attending school in grade 7 or above (referred to as likely ineligible). Eligibility status derived from the Youth Interview data was evaluated against the likely ineligible flag variable which is a screener variable marking whether an interviewed youth was likely ineligible based on the completed background questionnaire. The experiment showed that for all countries, except Senegal, less than 5% of cases deemed likely ineligible at the screener stage turned out to be eligible at the Youth Interview stage. For Senegal however, about 27% of the likely ineligible cases turned out to be eligible at the Youth Interview stage. It was therefore decided to keep sampling one-third of likely ineligibles for Senegal's Main Survey, but not for the other four countries.

In the Main Survey, we compared eligibility status from the screener and the Youth Interview to obtain the rate of false positives. For Guatemala, Honduras, Paraguay and Senegal, about 15-17% of youth deemed eligible at the screener stage were not eligible at the Youth Interview stage. This value is about 6% for Panama. For Senegal, about 7% of cases deemed likely ineligible at the screener stage (false negatives) turned out to be eligible at the Youth Interview stage. This lower percentage, as compared to the Field Trial, might be due to wording changes for the screener between the Field Trial and the Main Survey.

RECOMMENDATIONS

Based on the Field Trial and Main Survey experience of PISA-D Strand C, the international contractor (Westat) provides the following sampling recommendations for future cycles of PISA- D Strand C:

- The National Centre relationship with the National Statistical Institute (NSI) is crucial for PISA-D Strand C success. It is critical for the country to have access to reliable data for stratification of small areas including the number of dwellings, population, target population, concentration classification (major strata), urbanisation and region. Furthermore, it is critical to have access to up-to-date area maps and lists of dwellings and/or develop a strategy to create a sampling frame of dwelling units within selected areas. Sample selection could also be done by the NSI.
- More training is needed so that countries can better understand the benefits (cost reduction, potential sampling/non-sampling error reduction) of using a probability sample of dwelling units along with the probability-based link-tracing (PBLT) approach as compared to a mini-census. For the link-tracing, an experiment in the FT could be done to compare optional use of coupons versus referrals only, or to compare mini-census versus PBLT (with coupon/referral options) to see what approach would work best for the Main Survey.
- The international contractor (or the NSI) should conduct sample selection as soon as it is determined to be beneficial to do so. For the pilot project, the international contractor (Westat) conducted a large number of capacity building activities but they did not seem to be effective in most situations. Most countries' Ministries of Education are not set up to conduct surveys requiring household sampling. In some cases, sample selection done by the international contractor (or the NSI) would improve:
    - processing efficiency for both contractors and countries
    - design efficiency in costs and variance and more efficient complex samples designed to be near-optimal
    - reduce non-sampling errors, such as errors reduced in ID assignments
    - knowledge of key quality indicators, such as undercoverage
    - gathering sample design information (e.g. relevant information needed on the sampling frame).
- An improved quality control (QC) process is needed for the mini-census and for the selection process of dwellings in case of a sample. GPS data and digitalised maps could potentially be used as QC or to assist in the creation of a sampling frame of dwelling units within selected areas.
- Countries in the lowest-end of the estimated percentage in the target population should be given the option of a reduced sample size while understanding the pros and cons (e.g. a reduced sample size could mean their sample will be combined with similar countries with the same assessment language for item response theory modeling).
- The international contractor (Westat) should provide more data cleaning tools and instructions for sampling-related data and more time should be allotted for data cleaning by countries.

- Depending on the situation, visits could be helpful during the sampling process (frame creation and selection), either as a workshop, to plan and assist in the process, or for verification purposes.

REFERENCES

Goodman, L. (1961), Snowball Sampling, Ann, *Math, Statist, 32, 148-170.*

Heckathorn, D. (2007). "Extensions of Respondent Driven Sampling: Analyzing Continuous Variables and Controlling for Differential Recruitment", *Social Methodology,* 37, 151-207.

Kalton, G. (1991), "Sampling flows of mobile human populations", *Survey Methodology,* 17, 183-194.