Measuring Proficiency Skills in Household Surveys

PIAAC and PISA-D

Leyla Mohadjer, Tom Krenzke
Westat

Measuring Learning through the Lifecycle in Multi-Topic National Household Surveys: Options for low-income countries

February 3-4, 2020
World Bank
Washington, DC
Agenda

Brief and general introduction

- PIAAC (Programme for the International Assessment of Adult Competencies)
- PISA (Programme for the International Student Assessment)
- PISA-D (PISA for Development)
Programme for the International Assessment of Adult Competencies (PIAAC)

- A multi-cycle international survey of assessment of adult skills and competencies, sponsored by the Organisation for Economic Co-operation and Development (OECD)
- Collects information on skills required in the workplace, educational background, professional attainment, and the ability to use information and communications technology
PIAAC – Sample Design and Data Collection Mode

› Household survey of adults 16-65 years old
› Sample design adaptive to each country’s best sampling scenario
› One or two adults selected per household
› In-person interviews conducted in respondents’ households
   • Background questionnaire
   • Direct assessment of cognitive skills to measure participants’ general levels of proficiency
› Main data collection for Cycle 2 in 2021-2022
Programme for the International Student Assessment (PISA)

❯ An international assessment of the knowledge and skills of 15-year-old students, sponsored by the Organisation for Economic Co-operation and Development (OECD)

❯ Conducted in schools

❯ Started in 2000

❯ Data collection every three years

❯ Includes over 80 countries and economies
  • Includes low-income countries
PISA for Development (PISA-D)

› Make PISA more accessible and relevant to middle- and low-income countries

› Includes both in-school and out-of-school data collection components
  • Out-of-school data collection conducted in households or similar locations
    – In-person interviews
      • Background questionnaire
      • Direct assessment of cognitive skills
The first PISA survey to collect data outside the schools

Goal: Development of robust sampling procedures

- Take account of the varied demographic profiles of PISA countries and economies
  - Ensure each sample adequately represents the target population of all 14- to 16-year-olds who are out-of-school, or in-school but in grades 6 or below
PISA-D Design and Challenges
Outline

PISA-D design and challenges
Solutions to reduce cost
Some sampling results
Generalisability
Within household sampling
Sustainable Development Goal -- ensure that all youth... achieve literacy and numeracy

Three “Strands”

- “A” in-school 15 year olds grade 7 or above
- “B” Contextual questionnaires
- “C” 15 year olds out-of-school or in-school grade 6 or below

— Pilot project (Field Trial and Main Survey)

In-person household youth assessment, with youth interview and screener

Five countries: Guatemala, Honduras, Panama, Paraguay, Senegal
Strand C Target Population

An illustration

- In-school, grade 7 or above (Strand A)
- Out-of-School (Strand C)
- In-school, grade 6 or below (Strand C)

Out-of-school (OOS) and below 7th grade

Broadened to 14-to-16 year olds
Percentage in the Target Population Among the Total Population

By Out-of-school (OOS) and In-school Below Grade 7 Strand C Countries
(Estimates derived from information provided by country prior to data collection)

Rare population:
The target population is about 1 to 3% of the total population
Dwelling Unit Sample Size, By Prevalence of Subpopulation of Interest, Under Simple Random Sampling, by Number of Completes (500, 1000, 1500)

Assumes household size = 5 and completion rate = 50%

Probability-based household samples:

- Screening amounts escalate at about 1 to 3% of the target population
- Large Increases to screening amounts depending on target number of completed cases
Cluster Sampling to Reduce Travel Costs

› Two-stage area probability sample
  • Stage 1: Select Primary Sampling Units (PSUs) to reduce cost of interviewer travel
  • Stage 2: Select Dwelling Units (DUs) to screen for eligibility status
    – Take all eligible youth ages 14-16
Reduce screening costs further

- Sampling rare populations -- Kalton and Anderson (1986)
  - Stratification by concentration (higher selection rate in higher concentration strata)
  - *Implemented in PISA-D Strand C*

Create DU sampling frame

- Probability-based options for Stage 2
  - Conduct traditional listing,
  - Use existing maps, or
  - Mini-census within areas -- Carr-Hill (2015)
  - *Implemented in most PISA-D Strand C countries*
A random route design is one that pre-assigns a skip interval to each interviewer as they walk systematically through the selected area following the instructions for the walking route.

The number of cases to select is typically pre-set for each area. This cost-effective approach was appealing because it did not rely on a DU sampling frame.

However, the approach is not a probability-based method due to potential coverage error issues, and error-prone reliance on interviewers to follow procedures closely without being left to judgment about the DUs to select (Bauer 2014).

Therefore, not considered as a design to provide adequate representation.
Non-probability Approaches Considered to Locate Youth

› Link-tracing
  • Respondent driven sampling -- Heckathorn (1997), and e.g., Lee (2009)

› Others considered, including location sampling

› Potential for coverage error and sampling bias
At Stage 2, select a probability sample of DUs

Enlist the respondent to help -- at the end of the screener ask for referrals, “Do you know any {other} 14-16 year olds who are out of school or below grade 7 in your neighborhood?”

• If so, obtain the name and address
• Send to home office and check for any duplication with probability sample of DUs, or other referrals
• Assign referrals to interviewers

Met the requirements of “Representative” sample – per PISA-D Strand C Technical Advisory Group (TAG)
Sample Size Requirements

- **n = 1,600** total completes, with objective to reduce costs
  - At least 1300 needing to pass the core cognitive assessment module
    - 650 responses per main assessment item
  - Representative sample (includes probability sample and probability-based link-tracing): **n = 1,200**
    - TAG approved at least 600 probability-based and up to 600 referrals
  - Limited Representative sample: **n = 400**
    - Various options provided
      - School-based sampling
        - Administrative records for in-school below 7th grade
        - Link-trace to out-of-school youth from students and teachers
      - Location sampling
      - Street children
      - Other
<table>
<thead>
<tr>
<th>Country</th>
<th>Overall</th>
<th>Representative probability</th>
<th>Link-tracing through households</th>
<th>Link-tracing through households recruiting</th>
<th>School frame approach for OOS</th>
<th>School frame approach for &lt;7th grade</th>
<th>Location sampling</th>
<th>Special operation for street children</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guatemala</td>
<td>1,749</td>
<td>1,250</td>
<td>NA</td>
<td>NA</td>
<td>129</td>
<td>181</td>
<td>189</td>
<td>0</td>
</tr>
<tr>
<td>Honduras</td>
<td>1,281</td>
<td>1,161</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>15</td>
<td>105</td>
<td>NA</td>
</tr>
<tr>
<td>Panama</td>
<td>2,055</td>
<td>1,902</td>
<td>79</td>
<td>2</td>
<td>0</td>
<td>72</td>
<td>0</td>
<td>NA</td>
</tr>
<tr>
<td>Paraguay</td>
<td>1,002</td>
<td>814</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>188</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Senegal</td>
<td>2,103</td>
<td>2,103</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>
## Completed Cases By Out-of-school and In-school < 7th Grade

<table>
<thead>
<tr>
<th>Country</th>
<th>Completed cases for analysis</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Overall</td>
<td>Out of school</td>
<td>In school &lt; 7th grade</td>
<td></td>
</tr>
<tr>
<td>Guatemala</td>
<td>1,749</td>
<td>1,031</td>
<td>718</td>
<td></td>
</tr>
<tr>
<td>Honduras</td>
<td>1,281</td>
<td>1,084</td>
<td>197</td>
<td></td>
</tr>
<tr>
<td>Panama</td>
<td>2,055</td>
<td>1,795</td>
<td>260</td>
<td></td>
</tr>
<tr>
<td>Paraguay</td>
<td>1,002</td>
<td>626</td>
<td>376</td>
<td></td>
</tr>
<tr>
<td>Senegal</td>
<td>2,103</td>
<td>1,272</td>
<td>831</td>
<td></td>
</tr>
</tbody>
</table>
## Representative Samples

### Sampling Information

<table>
<thead>
<tr>
<th>Country</th>
<th>Number of selected PSUs</th>
<th>Number of selected dwelling units</th>
<th>Number of completed assessments</th>
<th>Household size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guatemala</td>
<td>101</td>
<td>25,975</td>
<td>1,041</td>
<td>4.5</td>
</tr>
<tr>
<td>Honduras</td>
<td>217</td>
<td>18,582</td>
<td>1,070</td>
<td>4.1</td>
</tr>
<tr>
<td>Panama</td>
<td>531</td>
<td>6,081</td>
<td>1,902</td>
<td>3.9</td>
</tr>
<tr>
<td>Paraguay</td>
<td>498</td>
<td>28,709</td>
<td>669</td>
<td>3.8</td>
</tr>
<tr>
<td>Senegal</td>
<td>80</td>
<td>8,774</td>
<td>2,023</td>
<td>9.4</td>
</tr>
</tbody>
</table>

### Weighted Response Rates

<table>
<thead>
<tr>
<th>Country</th>
<th>Screener</th>
<th>YI</th>
<th>Assessment</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guatemala</td>
<td>89%</td>
<td>71%</td>
<td>95%</td>
<td>60%</td>
</tr>
<tr>
<td>Honduras</td>
<td>92%</td>
<td>81%</td>
<td>95%</td>
<td>71%</td>
</tr>
<tr>
<td>Panama</td>
<td>96%</td>
<td>94%</td>
<td>100%</td>
<td>90%</td>
</tr>
<tr>
<td>Paraguay</td>
<td>81%</td>
<td>83%</td>
<td>96%</td>
<td>64%</td>
</tr>
<tr>
<td>Senegal</td>
<td>80%</td>
<td>91%</td>
<td>100%</td>
<td>73%</td>
</tr>
</tbody>
</table>
Eligibility and Hit Rates

- Hit Rate = 100 x (number of completed assessments) / (number of dwelling units selected)

- Representative sample

<table>
<thead>
<tr>
<th>Country</th>
<th>Eligibility rate (%)</th>
<th>Expected hit rate (%)</th>
<th>Actual hit rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>H</td>
</tr>
<tr>
<td>Guatemala</td>
<td>2.1</td>
<td>5.8</td>
<td>9.4</td>
</tr>
<tr>
<td>Honduras</td>
<td>2.3</td>
<td>6.3</td>
<td>7.3</td>
</tr>
<tr>
<td>Panama</td>
<td>9.6</td>
<td>5.8</td>
<td>6.7</td>
</tr>
<tr>
<td>Paraguay</td>
<td>1.4</td>
<td>5.9</td>
<td>6.7</td>
</tr>
<tr>
<td>Senegal</td>
<td>3.4</td>
<td>25.6</td>
<td>30.6</td>
</tr>
</tbody>
</table>
## Summary of Weighting Steps

<table>
<thead>
<tr>
<th>Sample type</th>
<th>Step</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Guatemala</td>
</tr>
<tr>
<td><strong>Representative</strong></td>
<td>Base weights</td>
<td>Nation</td>
</tr>
<tr>
<td></td>
<td>Screener eligibility and nonresponse</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>adjustments</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Youth interview eligibility and</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>nonresponse adjustments</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trimming, calibration</td>
<td>x</td>
</tr>
<tr>
<td><strong>Limited representative</strong></td>
<td>Base weights</td>
<td>Nation</td>
</tr>
<tr>
<td></td>
<td>Calibration</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>Compositing, calibration</td>
<td>x</td>
</tr>
</tbody>
</table>
## Generalisability

<table>
<thead>
<tr>
<th>Country</th>
<th>Generalisability of estimates</th>
<th>Cautions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guatemala</td>
<td>National</td>
<td>Large underestimate in sum of weights prior to weight calibration step. Some caution is given due to some sample coming from non-probability methods.</td>
</tr>
<tr>
<td>Honduras</td>
<td>Rural population, urban respondents</td>
<td>Large underestimate in sum of weights prior to weight calibration step occurred for rural areas. The urban sample is purposively selected and very limited.</td>
</tr>
<tr>
<td>Panama</td>
<td>Rural and indigenous population, urban respondents</td>
<td>Large overestimate in sum of weights prior to weight calibration step for rural and indigenous areas. This is likely due to unexpected extreme hit rates especially in rural areas, and fewer DUs listed and selected than expected given the data provided. The urban sample is purposively selected and very limited.</td>
</tr>
<tr>
<td>Paraguay</td>
<td>National</td>
<td>Large underestimate in sum of weights prior to weight calibration step. Some caution is given due to some sample coming from non-probability methods. Two departments were excluded: Boqueron and Alto Paraguay.</td>
</tr>
<tr>
<td>Senegal</td>
<td>National</td>
<td>Low-to-moderate underestimate in sum of weights prior to weight calibration step. This is likely due to a short data collection period, and use of outdated DU listings.</td>
</tr>
</tbody>
</table>
Within Household Enumeration

› Usual residence vs Slept at residence (Guatemala, for example)

— Usual place of residence -- only 4 completed the interview who did not sleep at the residence with a completion rate of only 7%.

— 37 completes who were visitors, and 90% of visitors completed the survey among all who were visitors. This compares to 71% of the sampled youths who completed the survey overall.

• Slept at residence might be a better rule – but individuals are susceptible to have multiple chances of selection

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

› Field Trial
  • Took all eligible youth 14-16 (eligible based on age and enrollment)
  • Took 1/3 of screened ineligible youth 14-16 (based on enrollment response)
  • Results showed 4 of the 5 countries had negligible rates of misclassification of eligibility status (using the interview self-reports as the truth)

› Main Study
  • 1 of the 5 countries continued to take 1/3 of screened ineligible youth 14-16 to interview to be conservative for potential misclassification


References


Thank You

Leyla Mohadjer -- LeylaMohadjer@Westat.com

Tom Krenzke -- TomKrenzke@Westat.com