Sampling Notes - Peru

Brian Stacy

## Introduction

This file outlines the sampling strategy for Peru.

## General Notes on Sampling

The aim of the Global Education Policy Dashboard school survey is to produce nationally representative estimates, which will be able to detect changes in the indicators over time at a minimum power of 80% and with a 0.05 significance level. We also wish to detect differences by urban/rural location.

For our school survey, we will employ a two-stage random sample design, where in the first stage a sample of around 200 schools, based on local conditions, is drawn, chosen in advance by the Bank staff. In the second stage, a sample of teachers and students will be drawn to answer questions from our survey modules, chosen in the field. A total of 10 teachers will be sampled for absenteeism. Five teachers will be interviewed and given a content knowledge exam. Three 1st grade students will be assessed at random, and a classroom of 4th grade students will be assessed at random. Stratification will be based on the school’s urban/rural classification and based on region. When stratifying by region, we will work with our partners within the country to make sure we include all relevant geographical divisions.

For our Survey of Public Officials, we will sample a total of 200 public officials. Roughly 60 officials will be surveyed at the federal/regional level, while 140 officials will be surveyed at the district level. For selection of officials at the district level, we will employ a cluster sampling strategy, where 20 district offices are chosen at random. Within the districts seven officials will be sampled, including the head of organization, HR director, two division directors from finance and planning, and three randomly selected professional employees among the finance, planning, and one other service related department chosen at random. At the federal level, we will interview the HR director, finance director, planning director, and three randomly selected service focused departments. In addition to the directors of each of these departments, a sample of 9 professional employees will be chosen in each department at random on the day of the interview.

Sampling Approach for Global Education Policy Dashboard

This document will provide an overview of the sampling strategy used in the Global Education Policy Dashboard (GEPD) surveys, as well as remaining questions. New data for the dashboard will be collected using three main instruments: a School Survey, an Expert Survey, and a Survey of Public Officials. More information pertaining to each can be found below. The goal of the Global Education Policy Dashboard is to provide summary information at the national level on a set of 35 indicators and to allow countries to track progress on those indicators over a short time frame (every 2 years). Specifically, we aim to produce nationally representative estimates, which will be able to detect changes in the indicators over time at a minimum power of 80% and with a 0.05 significance level. We also wish to disaggregate by urban/rural.

School Survey: The School Survey will collect data primarily on Practices (the quality of service delivery in schools), but also on some de facto Policy and school-level Politics indicators. It will consist of streamlined versions of existing instruments—including SDI and SABER SD on teachers, 4th grade students, and inputs/infrastructure, TEACH on pedagogical practice, GECDD on school readiness of young children, and DWMS on management quality—together with new questions to fill gaps in those instruments. Though the number of modules is similar to the full version of SDI, the number of items within each module is significantly lower. In each country, this survey will be administered in a nationally representative sample of 200 schools, selected through clustered random sampling. As currently envisioned, the School Survey will include 8 short modules. Expert Survey: The Expert Survey will collect information to feed into the policy indicators. This survey will be filled out by key informants in each country, drawing on their knowledge to identify key elements of the policy framework (as in the SABER approach to policy-data collection that the Bank has used over the past 7 years). The survey will have 4 modules with each including approximately ten questions. Survey of Public Officials: The Survey of Public Officials will collect information about the capacity and orientation of the bureaucracy, as well as political factors affecting education outcomes. This survey will be a streamlined and education-focused version of the civil-servant surveys that the Bank’s Bureaucracy Lab has implemented recently in several countries, and the dashboard team is collaborating closely with DEC and Governance GP staff to develop this instrument. As currently envisioned, the survey will be administered to a random sample of about 200 staff serving in the central education ministry and district education offices. It will include questions about technical and leadership skills, work environment, stakeholder engagement, clientelism, and attitudes and behaviors.

## Peru Specific Comments

MELQO data was merged with the Peru school frame in order to optimally stratify. We stratified on the basis of urban/rual and department. There are 25 departments in Peru. In 2017, Peru conducted an examination of around 4,500 children between 5 and 8 years old, with a median age of 6. The MELQO exam is quite similar to our ECD examination module. We are able to use data from this 2017 survey to choose the number of schools in each province optimally by calculating means and standard deviations by province and feeding this information into the optimal stratification algorithm. See <https://cran.r-project.org/web/packages/SamplingStrata/vignettes/SamplingStrata.html>. Provinces with low standard deviations among students in terms of their MELQO development scores are allocated fewer schools compared to an allocation that is simply based on population, and provinces with high standard deviations are allocated more schools.

203 schools were chosen for our survey after optimally stratifying. The table below shows the allocation across department and by urban/rural.

Counts of Schools Sampled per Departamento and Urban/Rural Status

departamento

rural

N

AMAZONAS

FALSE

2

AMAZONAS

TRUE

6

ANCASH

FALSE

3

ANCASH

TRUE

4

APURIMAC

FALSE

4

APURIMAC

TRUE

4

AREQUIPA

FALSE

4

AREQUIPA

TRUE

1

AYACUCHO

FALSE

4

AYACUCHO

TRUE

4

CAJAMARCA

FALSE

6

CAJAMARCA

TRUE

16

CALLAO

FALSE

3

CUSCO

FALSE

6

CUSCO

TRUE

6

HUANCAVELICA

FALSE

2

HUANCAVELICA

TRUE

3

HUANUCO

FALSE

4

HUANUCO

TRUE

7

ICA

FALSE

3

JUNIN

FALSE

2

JUNIN

TRUE

3

LA LIBERTAD

FALSE

6

LA LIBERTAD

TRUE

8

LAMBAYEQUE

FALSE

2

LIMA

FALSE

16

LIMA

TRUE

2

LORETO

FALSE

5

LORETO

TRUE

14

MADRE DE DIOS

TRUE

1

PASCO

FALSE

2

PASCO

TRUE

2

PIURA

FALSE

9

PIURA

TRUE

8

PUNO

FALSE

4

PUNO

TRUE

5

SAN MARTIN

FALSE

6

SAN MARTIN

TRUE

6

TACNA

FALSE

2

TACNA

TRUE

2

TUMBES

TRUE

1

UCAYALI

FALSE

1

UCAYALI

TRUE

4

## Map of Peru Schools

## PhantomJS not found. You can install it with webshot::install\_phantomjs(). If it is installed, please make sure the phantomjs executable can be found via the PATH variable.

## Map of Peru Schools with Replacements