Peru 2007

1 Survey Description

Survey: Encuesta Nacional de Hogares sobre Condiciones de Vida y Pobreza (ENAHO), carried out by the National Institute of Statistics and Informatics (INEI) of Peru, for the year 2007.

Link to the document: http://webinei.inei.gob.pe/anda_inei/index.php/catalog/362

Sample: The survey was conducted between January and December 2007. The ENAHO collects information from 24 states and is nationally representative. The sample considered is restricted to adult individuals aged 19 to 60 with information about outcome and circumstances (39,634 individuals). Section 3 of this document describes the prevalence and pattern of missing data., 57,118 individuals in the total sample and 26,074 individuals in the analysis sample. Section 3 of this document describes the prevalence and pattern of missing data.

Weights: The weights are the inverse of the probability that the observation is included because of the sampling design. It could be understood as the number of individuals in the population that an individual represents in the sample according to the 2007 Population Census and demographic projections (*weights*).

Outcome: The outcome variables are annual equivalized household disposable total (*eq_iinc*) income in dollars PPP 2017.[1]

Circumstances:

- Sex (female, male)
- Race 'Ethnicity' (indigenous, afro-descendant and others, described in Table 1)
- Region of birth 'Birthplace' (25 provinces, described in Table 1)
- Father's edu. (years of education (maximum 15), described in Table 2)
- Mother's edu. (years of education (maximum 15), described in Table 2)

[1] Income variable was equivalized using the square root scale.

2 Descriptive Statistics

Table 1: Respondant's socio-demographics - 2007

	Analisis sample	Total sample
	(N=26,074)	(N=57,118)
Gender		
0 Female	7,890 (30.3%)	29,194 (51.1%)
1 Male	18,184 (69.7%)	27,924 (48.9%)
Ethnicity		
1 White	730 (2.8%)	1,602 (2.8%)
2 Mix-race	13,473 (51.7%)	27,277 (47.8%)
3 Indigenous	10,713 (41.1%)	20,567 (36.0%)
4 Afro-descendant	461 (1.8%)	837 (1.5%)
5 Other	697 (2.7%)	1,443 (2.5%)
Missing	0 (0%)	5,392 (9.4%)
Region of birth		
1 Amazonas-San Martin-Ucayali	2,181 (8.4%)	4,819 (8.4%)
2 Anchash	1,218 (4.7%)	2,796 (4.9%)
3 Apurimac	893 (3.4%)	1,930 (3.4%)
4 Arequipa	1,011 (3.9%)	2,156 (3.8%)
5 Ayacucho	1,201 (4.6%)	2,652 (4.6%)
6 Cajamarca	1,937 (7.4%)	4,291 (7.5%)
7 Cusco-Madre de Dios	680 (2.6%)	1,472 (2.6%)
8 Huancavelica	1,294 (5.0%)	2,789 (4.9%)
9 Huánuco	1,019 (3.9%)	2,188 (3.8%)
10 Ica	1,068 (4.1%)	2,543 (4.5%)
11 Junin-Pasco	1,830 (7.0%)	3,583 (6.3%)
12 La Libertad	1,121 (4.3%)	2,534 (4.4%)
13 Lambayeque	1,060 (4.1%)	2,348 (4.1%)
14 Lima	988 (3.8%)	2,152 (3.8%)
15 Callao	2,580 (9.9%)	5,689 (10.0%)
16 Loreto	1,104 (4.2%)	2,848 (5.0%)
17 Moquegua-Tacna	958 (3.7%)	1,927 (3.4%)
18 Piura-Tumbes	2,365 (9.1%)	4,886 (8.6%)
19 Puno	1,566 (6.0%)	3,424 (6.0%)
Missing	0 (0%)	91 (0.2%)

Table 2: Parental education - 2007

	N	Mean	SD	Median	Min	Max	Missing
Analisis sample - Mother's edu	26,074	4.610	4.699	3	0	15	0
Analisis sample - Father's edu	26,074	6.187	4.803	6	0	15	0
Total sample - Mother's edu	57,118	4.673	4.742	3	0	15	24,863
Total sample - Father's edu	57,118	6.245	4.805	6	0	15	28,739

Table 3: Respondant's income - 2007

	N	Mean	SD	Median	Min	Max	Missing
Analisis sample	26,074	8,301	9,358	5,921	42.62	191,782	0
Total sample	57,118	8,533	10,062	5,995	42.62	197,124	0

3 Missing data analysis

3.1 Missing patterns

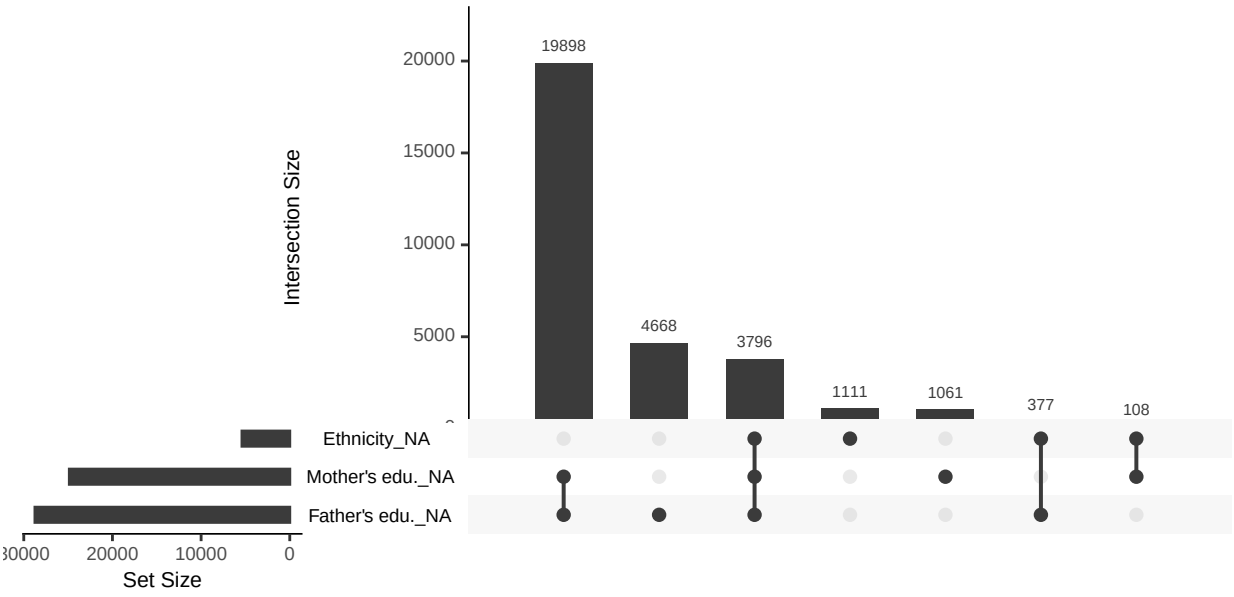

Figure 1: Missing patterns: *Left:* Marginal distribution of missing observations per variable. *Right:* Combination of missingness across cases

3.2 Differences in expected total equivalized household income between samples

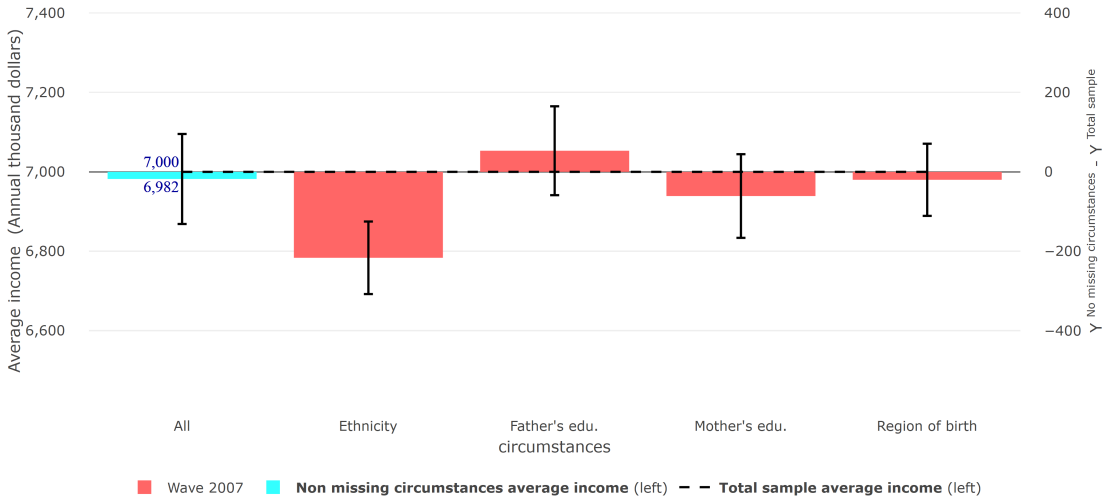

Figure 2: Differences in expected equivalized household income between the sample with non-missing circumstances and the total sample

3.3 Gini coefficient

Table 4: Gini coefficient in analysis sample and total sample

Wave	Sample	Gini	Lower bound	Upper bound	Average income
Wave 2007	Analysis sample	0.468	0.458	0.469	6,982
Wave 2007	Total sample	0.475	0.463	0.471	7,000

3.4 Differences in Gini coefficient between samples

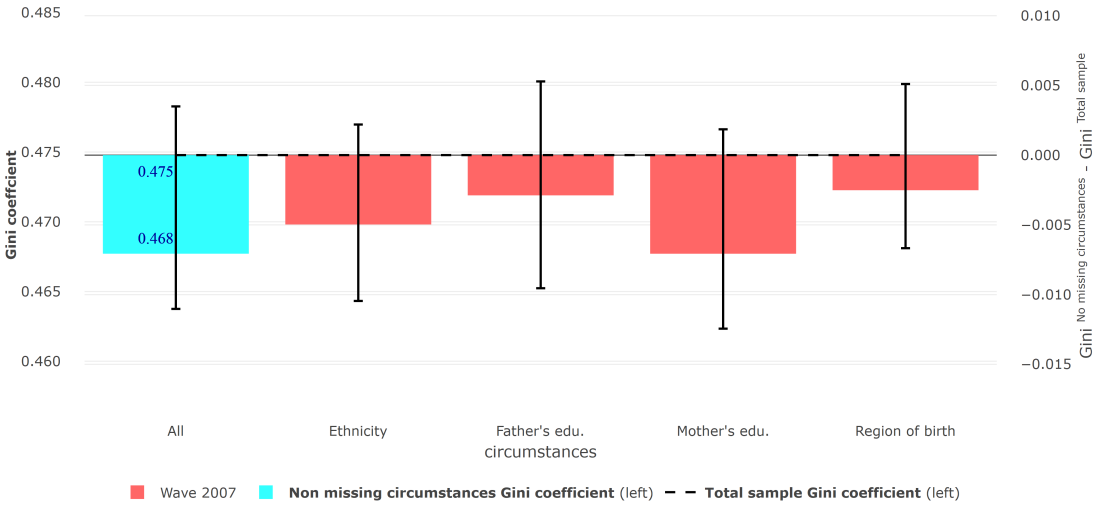

Figure 3: Differences in Gini coefficient between the sample with non-missing circumstances and the total sample