

Peru 2008

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 2008.

Link to the document: http://webinei.inei.gob.pe/anda_inei/index.php/catalog/364

Sample: The survey was conducted between January and December 2008. The ENAHO collects information from 24 states and is nationally representative. The sample considered is restricted to adult individuals aged 19 to 61 with information about outcome and circumstances (38,724 individuals). Section 3 of this document describes the prevalence and pattern of missing data., 55,486 individuals in the total sample and 25,287 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)

¹Income variable was equivalized using the square root scale.

2 Descriptive Statistics

Table 1: Respondant's socio-demographics - 2008

	Analysis sample	Total sample
	(N=25,287)	(N=55,486)
Gender		
0 Female	7,765 (30.7%)	28,338 (51.1%)
1 Male	17,522 (69.3%)	27,148 (48.9%)
Ethnicity		
1 White	730 (2.9%)	1,665 (3.0%)
2 Mix-race	12,495 (49.4%)	25,643 (46.2%)
3 Indigenous	10,633 (42.0%)	20,485 (36.9%)
4 Afro-descendant	340 (1.3%)	610 (1.1%)
5 Other	1,089 (4.3%)	2,378 (4.3%)
Missing	0 (0%)	4,705 (8.5%)
Region of birth		
1 Amazonas-San Martin-Ucayali	2,137 (8.5%)	4,698 (8.5%)
2 Ancash	1,277 (5.1%)	2,741 (4.9%)
3 Apurimac	862 (3.4%)	1,831 (3.3%)
4 Arequipa	1,065 (4.2%)	2,248 (4.1%)
5 Ayacucho	1,166 (4.6%)	2,633 (4.7%)
6 Cajamarca	1,726 (6.8%)	4,041 (7.3%)
7 Cusco-Madre de Dios	678 (2.7%)	1,462 (2.6%)
8 Huancavelica	1,254 (5.0%)	2,757 (5.0%)
9 Huánuco	990 (3.9%)	2,132 (3.8%)
10 Ica	1,008 (4.0%)	2,449 (4.4%)
11 Junin-Pasco	1,693 (6.7%)	3,393 (6.1%)
12 La Libertad	1,159 (4.6%)	2,564 (4.6%)
13 Lambayeque	998 (3.9%)	2,317 (4.2%)
14 Lima	918 (3.6%)	2,080 (3.7%)
15 Callao	2,507 (9.9%)	5,430 (9.8%)
16 Loreto	1,137 (4.5%)	2,739 (4.9%)
17 Moquegua-Tacna	971 (3.8%)	1,908 (3.4%)
18 Piura-Tumbes	2,203 (8.7%)	4,629 (8.3%)
19 Puno	1,538 (6.1%)	3,331 (6.0%)
Missing	0 (0%)	103 (0.2%)

Table 2: Parental education - 2008

	N	Mean	SD	Median	Min	Max	Missing
Analysis sample - Mother's edu	25,287	4.721	4.735	3	0	15	0
Analysis sample - Father's edu	25,287	6.287	4.810	6	0	15	0
Total sample - Mother's edu	55,486	4.766	4.782	3	0	15	24,097
Total sample - Father's edu	55,486	6.319	4.818	6	0	15	28,061

Table 3: Respondant's income - 2008

	N	Mean	SD	Median	Min	Max	Missing
Analysis sample	25,287	8,566	8,569	6,442	10.57	137,890	0
Total sample	55,486	8,576	9,423	6,377	10.57	350,861	0

3 Missing data analysis

3.1 Missing patterns

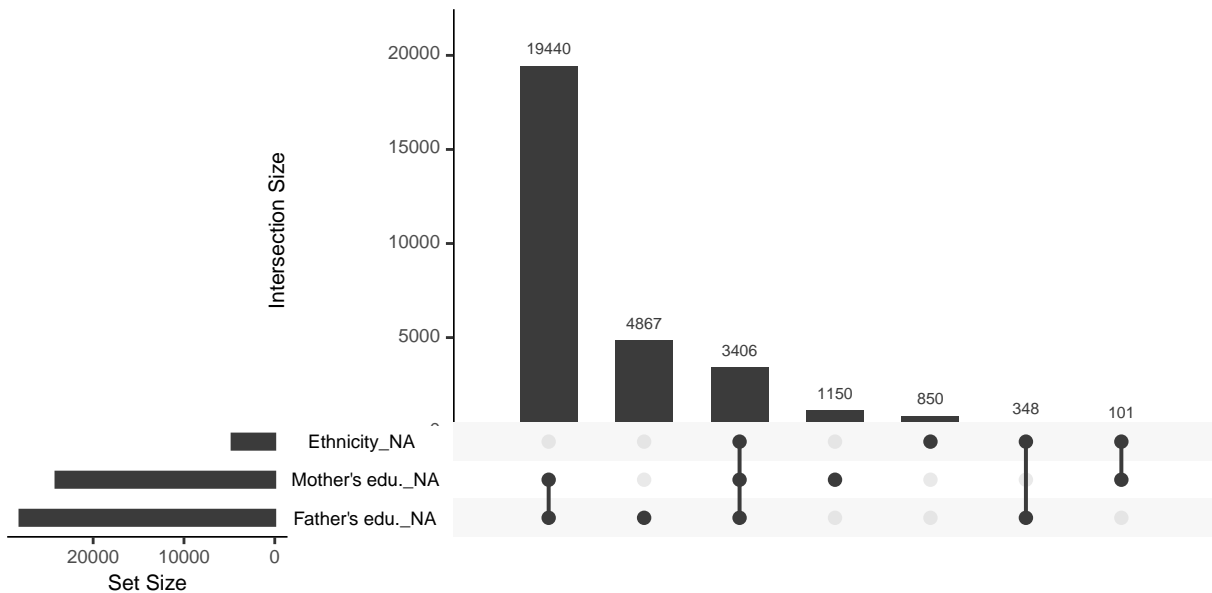


Figure 1: Missing patterns: *Left:* Marginal distribution of missing observations per variable. *Right:* Combination of missingness across cases

3.2 Differences in expected total equivalized household income between samples

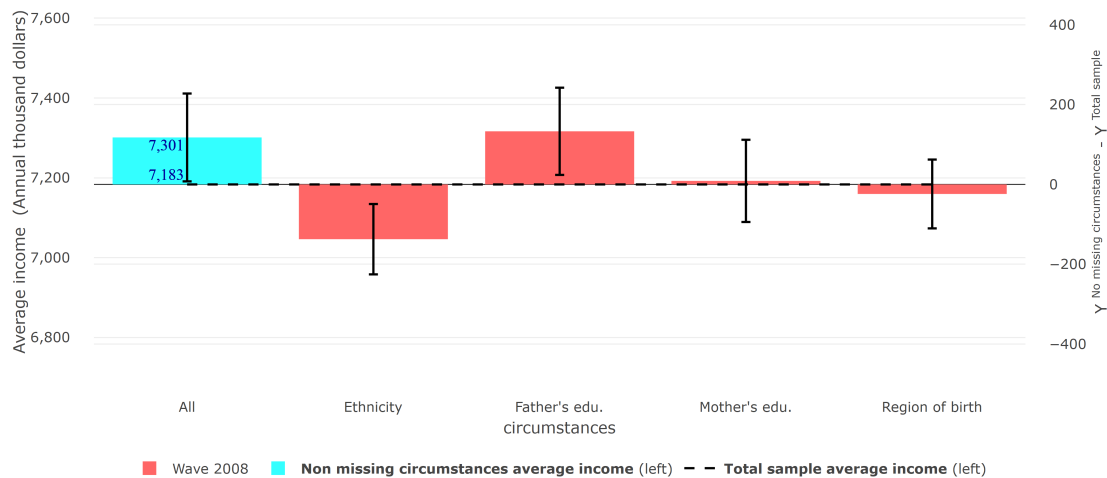


Figure 2: Differences in expected total equivalized household income between the sample with non-missing circumstances and the total sample

3.3 Gini coefficient

Table 4: Gini coefficient in analysis sample and total sample

Wave	Sample	Gini	Lower bound	Upper bound	Average income
Wave 2008	Analysis sample	0.441	0.439	0.45	7,301
Wave 2008	Total sample	0.444	0.442	0.45	7,183

3.4 Differences in Gini coefficient between samples

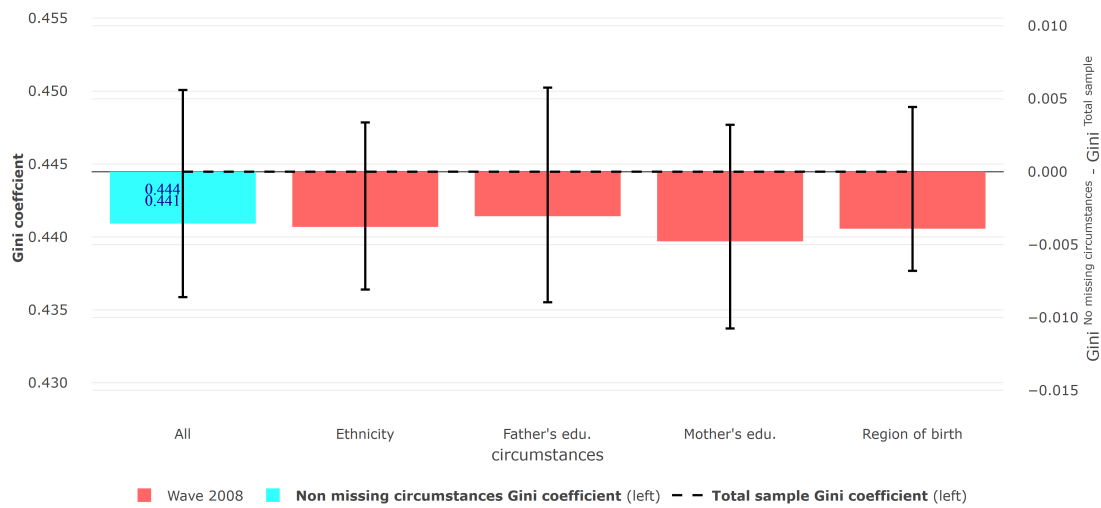


Figure 3: Differences in Gini coefficient between the sample with non-missing circumstances and the total sample