

India 2005

1 Survey Description

Survey: The India Human Development Survey (IHDS), carried out by researchers from the University of Maryland and the National Council of Applied Economic Research

Link to the document: <https://ihds.umd.edu/data>

Sample: The India Human Development Survey (IHDS) is a nationally representative, multi-topic panel survey of 41,554 households in 1503 villages and 971 urban neighborhoods across India. The first round of interviews were completed in 2004-2005.. There are 131,670 individuals in the total sample and 40,074 individuals in the analysis sample. Section 3 of this document describes the prevalence and pattern of missing data.

Weights: Weights adjust for differential sampling proportions from rural districts and from urban towns and cities, and for the probability of villages or towns/cities being sampled.

Outcome: The outcome variables are annual equivalized household disposable total (*INCOME*) income in dollars PPP 2017.¹

Circumstances:

- Sex (female, male, Table 1)
- Ethnicity (several ethnic and caste categories, described in Table 1)
- Birth Area (places of birth, described in Table 1)
- Fathers's edu. (levels of education, described in Table 2)
- Mother's edu. (levels of education, described in Table 2)
- Father's occ. (8 categories, from NOC-1968 code, described in Table 3)

¹Income variable was equivalized using the square root scale.

2 Descriptive Statistics

Table 1: Respondant's socio-demographics - 2005

	Analysis sample	Total sample
	(N=40,074)	(N=131,670)
Gender		
0 Female	6,604 (16.5%)	65,552 (49.8%)
1 Male	33,470 (83.5%)	66,118 (50.2%)
Ethnicity		
2 Forward caste	10,149 (25.3%)	36,502 (27.7%)
3 Other Backward castes (OBC)	12,789 (31.9%)	42,376 (32.2%)
4 Dalit	8,435 (21.0%)	24,359 (18.5%)
5 Adivasi	3,054 (7.6%)	9,245 (7.0%)
6 Muslim	4,727 (11.8%)	15,179 (11.5%)
7 Christian, Sikh, Jain	920 (2.3%)	4,009 (3.0%)
Birth Area		
1 Same state as current residence	38,817 (96.9%)	126,324 (95.9%)
2 Another state	1,040 (2.6%)	4,466 (3.4%)
3 Overseas	217 (0.5%)	880 (0.7%)

Table 2: Parental education - 2005

	Analysis sample	Total sample
	(N=40,074)	(N=131,670)
Father's education (levels)		
0 None	23,114 (57.7%)	55,987 (42.5%)
1 Incomplete primary	4,539 (11.3%)	10,348 (7.9%)
2 Complete primary	2,961 (7.4%)	8,007 (6.1%)
3 Incomplete secondary	4,411 (11.0%)	11,724 (8.9%)
4 Complete secondary	2,759 (6.9%)	8,074 (6.1%)
5 Higher secondary	860 (2.1%)	2,530 (1.9%)
6 Post-secondary or higher	1,430 (3.6%)	3,280 (2.5%)
Missing	0 (0%)	31,720 (24.1%)
Mother's education (levels)		
0 None	31,202 (77.9%)	60,799 (46.2%)
1 Incomplete primary	2,438 (6.1%)	5,431 (4.1%)
2 Complete primary	2,052 (5.1%)	4,791 (3.6%)
3 Incomplete secondary	2,689 (6.7%)	6,365 (4.8%)
4 Complete secondary	1,054 (2.6%)	2,623 (2.0%)
5 Higher secondary	312 (0.8%)	699 (0.5%)
6 Post-secondary or higher	327 (0.8%)	683 (0.5%)
Missing	0 (0%)	50,279 (38.2%)

Table 3: Parental occupation - 2005

	Analysis sample	Total sample
	(N=40,074)	(N=131,670)
Father's occupation		
1 Professionals and technicians	2,628 (6.6%)	3,904 (3.0%)
2 Administratives and managers	881 (2.2%)	1,223 (0.9%)
3 Clerks and related	2,826 (7.1%)	3,785 (2.9%)
4 Sales workers	2,273 (5.7%)	3,791 (2.9%)
5 Services workers	2,512 (6.3%)	3,755 (2.9%)
6 Agricultural, Forestry and fishery workers	21,064 (52.6%)	34,709 (26.4%)
7 Production and related workers, transport operators and labourers	7,488 (18.7%)	9,793 (7.4%)
8 Unemployed	402 (1.0%)	587 (0.4%)
Missing	0 (0%)	70,123 (53.3%)

Table 4: Respondant's income - 2005

	N	Mean	SD	Median	Min	Max	Missing
Analysis sample	40,074	6,452	7,810	4,077	1.003	213,985	0
Total sample	131,670	6,976	9,141	4,326	1.003	438,255	0

3 Missing data analysis

3.1 Missing patterns

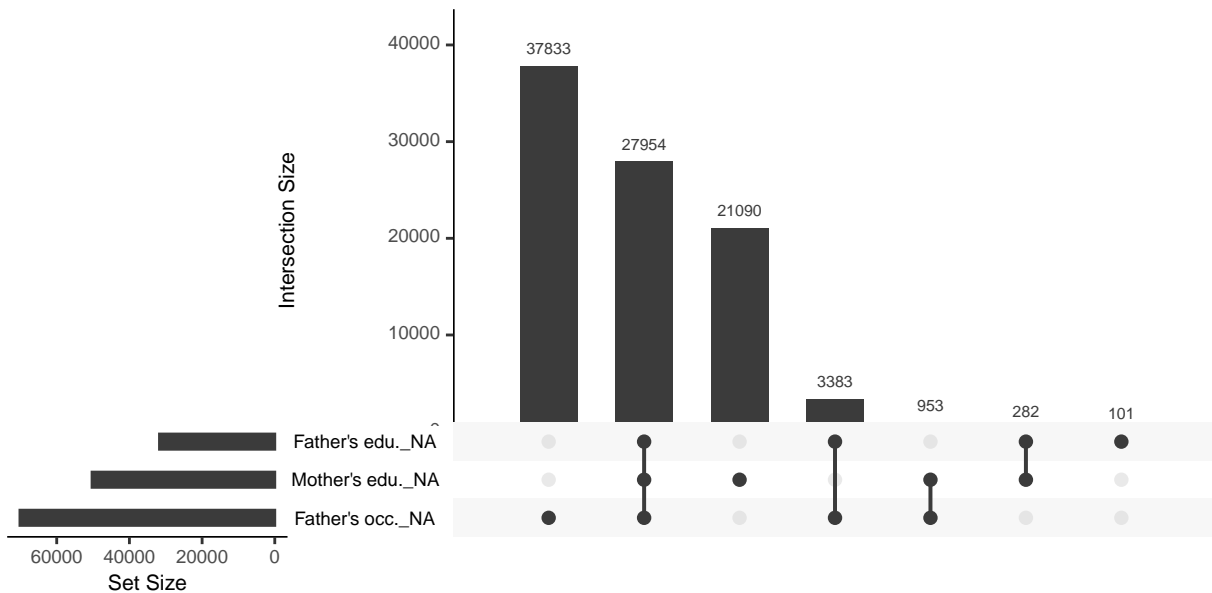


Figure 1: Missing patterns: *Left:* Marginal distribution of missing observations per variable. *Right:* Combination of missingness across cases

3.2 Differences in expected total equivalized household income between samples

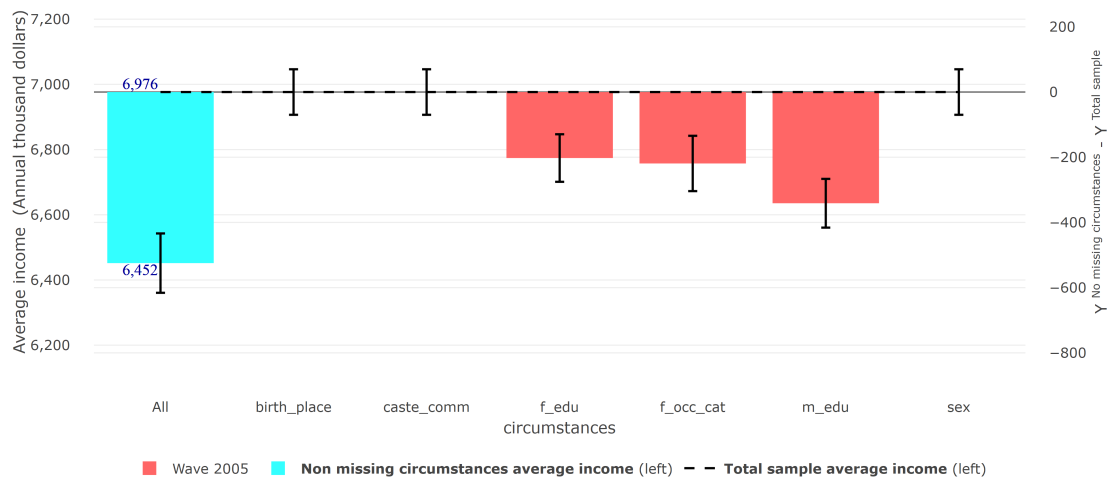


Figure 2: Differences in expected total equivalized household income between the sample with non-missing circumstances and the total sample

3.3 Gini coefficient

Table 5: Gini coefficient in analysis sample and total sample

Wave	Sample	Gini	Lower bound	Upper bound	Average income
Wave 2005	Analysis sample	0.492	0.481	0.494	6,452
Wave 2005	Total sample	0.505	0.497	0.505	6,976

3.4 Differences in Gini coefficient between samples

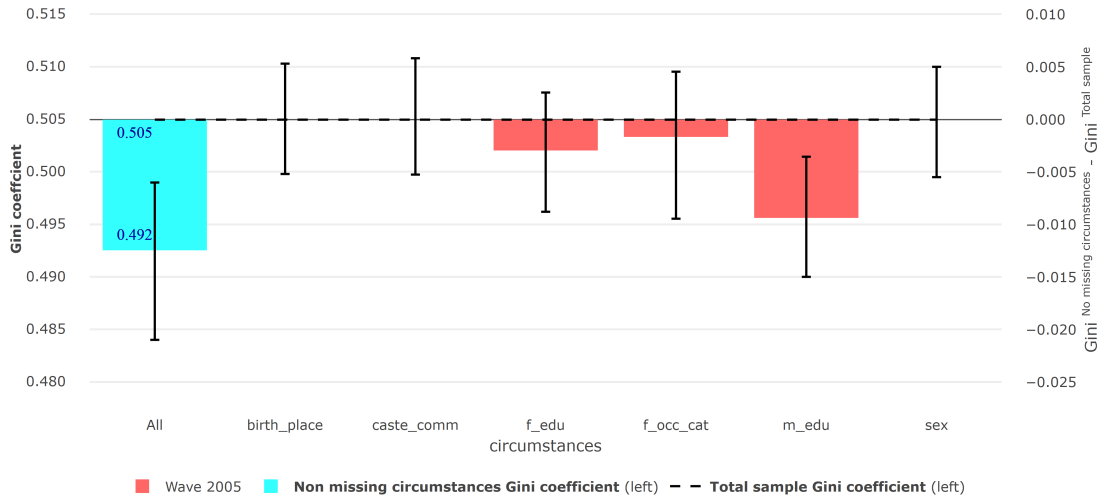


Figure 3: Differences in Gini coefficient between the sample with non-missing circumstances and the total sample