China 2010

1 Survey Description

Survey: The China Family Panel Studies (CFPS), carried out by the Institute of Social Science Survey (ISSS) of Peking University.

Link to the document: https://www.isss.pku.edu.cn/cfps/en/

Sample: The China Family Panel Studies (CFPS) is a nationally representative, biennual longitudinal survey designed to collect individual-, family-, and community-level longitudinal data in contemporary China. The studies focus on the economic, as well as the non-economic, wellbeing of the Chinese population. All members over age 9 in a sampled household are interviewed. The sample for the 2010 CFPS baseline survey through a multi-stage probability is drawn with implicit stratification. It is designed to be multi-stage so as both to reduce the operational cost of the survey and to allow for studies of social contexts. Each subsample in the CFPS study is drawn through three stages: county (or equivalent), then village (or equivalent), then household. . There are 30,671 individuals in the total sample and 17,402 individuals in the analysis sample. Section 3 of this document describes the prevalence and pattern of missing data.

Weights: The national full sample weight is the combined weights of five large provinces and 25 small provinces. Weight calculations take into account sampling design weights, non-response adjustment weights, post-hoc stratification adjustment weights, and trimming of the weights.

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 categories, described in Table 1)
- Birth Area (several provinces, 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. (11 categories, 10 from 1-Digit ISCO + one category including death/unknown/unemployed, described in Table 3)
- Mother's occ. (11 categories, 10 from 1-Digit ISCO + one category including death/unknown/unemployed, described in Table 3)

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

2 Descriptive Statistics

Table 1: Respondant's socio-demographics - 2010

	Analisis sample	Total sample
	(N=17,402)	(N=30,671)
Gender		
0 Female	9,000 (51.7%)	15,776 (51.4%)
1 Male	8,402 (48.3%)	14,895 (48.6%)
Ethnicity		
1 Han	16,003 (92.0%)	28,111 (91.7%)
2 Mongol	61 (0.4%)	92 (0.3%)
3 Hui	132 (0.8%)	251 (0.8%)
4 Tibetan	5 (0.0%)	117 (0.4%)
6 Miao	282 (1.6%)	526 (1.7%)
7 Yi	319 (1.8%)	405 (1.3%)
8 Zhuang	98 (0.6%)	233 (0.8%)
9 Bouyei	116 (0.7%)	219 (0.7%)
10 Korean	3 (0.0%)	13 (0.0%)
11 Manchu	233 (1.3%)	393 (1.3%)
99 Other	150 (0.9%)	284 (0.9%)
Missing	0 (0%)	27 (0.1%)
Birth Area		
13 Hebei	1,122 (6.4%)	1,676 (5.5%)
14 Shanxi	837 (4.8%)	1,436 (4.7%)
21 Liaoning	1,380 (7.9%)	2,738 (8.9%)
22 Jilin	270 (1.6%)	687 (2.2%)
23 Heilongjiang	434 (2.5%)	1,012 (3.3%)
31 Shanghai	1,150 (6.6%)	2,204 (7.2%)
32 Jiangsu	417 (2.4%)	788 (2.6%)
33 Zhejiang	304 (1.7%)	516 (1.7%)
34 Anhui	405 (2.3%)	693 (2.3%)
35 Fujian	152 (0.9%)	355 (1.2%)
36 Jiangxi	312 (1.8%)	552 (1.8%)
37 Shandong	994 (5.7%)	1,548 (5.0%)
41 Henan	2,028 (11.7%)	3,324 (10.8%)
42 Hubei	279 (1.6%)	578 (1.9%)
43 Hunan	515 (3.0%)	920 (3.0%)
44 Guangdong	1,301 (7.5%)	2,490 (8.1%)
45 Guangxi Zhuang Autonomous Region	333 (1.9%)	657 (2.1%)
51 Sichuan	990 (5.7%)	1,795 (5.9%)
52 Guizhou	523 (3.0%)	1,024 (3.3%)
53 Yunnan	683 (3.9%)	912 (3.0%)
61 Shaanxi	456 (2.6%)	716 (2.3%)
62 Gansu	2,124 (12.2%)	3,288 (10.7%)
90 Other	393 (2.3%)	704 (2.3%)
Missing	0 (0%)	58 (0.2%)

Table 2: Parental education - 2010

	Analisis sample	Total sample
	(N=17,402)	(N=30,671)
Father's education (levels)		
1 Illiterate/Semi-literate	8,876 (51.0%)	12,000 (39.1%)
2 Primary school	4,796 (27.6%)	6,776 (22.1%)
3 Junior high school	2,315 (13.3%)	3,890 (12.7%)
4 Senior high school/secondary school/technical school/vocational senior school	1,100 (6.3%)	1,959 (6.4%)
5 3-year college	167 (1.0%)	297 (1.0%)
6 4-year college/Bachelor's degree	140 (0.8%)	242 (0.8%)
7 Master's degree	4 (0.0%)	6 (0.0%)
8 Doctoral degree	4 (0.0%)	9 (0.0%)
Missing	0 (0%)	5,492 (17.9%)
Mother's education (levels)		
1 Illiterate/Semi-literate	12,464 (71.6%)	18,382 (59.9%)
2 Primary school	3,117 (17.9%)	4,769 (15.5%)
3 Junior high school	1,226 (7.0%)	2,297 (7.5%)
4 Senior high school/secondary school/technical school/vocational senior school	500 (2.9%)	967 (3.2%)
5 3-year college	57 (0.3%)	115 (0.4%)
6 4-year college/Bachelor's degree	36 (0.2%)	61 (0.2%)
7 Master's degree	1 (0.0%)	1 (0.0%)
8 Doctoral degree	1 (0.0%)	12 (0.0%)
9 No need to go to school	0 (0%)	1 (0.0%)
Missing	0 (0%)	4,066 (13.3%)

Table 4: Respondant's income - 2010

	N	Mean	SD	Median	Min	Max	Missing
Analisis sample	17,402	4,878	8,505	3,291	0.7989	276,736	0
Total sample	30,671	5,069	8,627	3,376	0.7989	276,736	0

Table 3: Parental occupation - 2010

	Analisis sample	Total sample
	(N=17,402)	(N=30,671)
Father's occupation		
0 Armed forces	82 (0.5%)	109 (0.4%)
1 Managers	744 (4.3%)	955 (3.1%)
2 Professionals	898 (5.2%)	1,135 (3.7%)
3 Technicians and Associate professionals	270 (1.6%)	333 (1.1%)
4 Clerks	149 (0.9%)	198 (0.6%)
5 Services and Sales workers	464 (2.7%)	586 (1.9%)
6 Agricultural, Forestry and Fishery workers	11,776 (67.7%)	13,187 (43.0%)
7 Craft and trade workers	1,218 (7.0%)	1,537 (5.0%)
8 Plant and machine opperators and assemblers	868 (5.0%)	1,090 (3.6%)
9 Elementary occupations	721 (4.1%)	891 (2.9%)
10 Unemployed	212 (1.2%)	274 (0.9%)
Missing	0 (0%)	10,376 (33.8%)
Mother's occupation		
0 Armed forces	3 (0.0%)	3 (0.0%)
1 Managers	163 (0.9%)	180 (0.6%)
2 Professionals	265 (1.5%)	307 (1.0%)
3 Technicians and Associate professionals	106 (0.6%)	123 (0.4%)
4 Clerks	80 (0.5%)	96 (0.3%)
5 Services and Sales workers	248 (1.4%)	311 (1.0%)
6 Agricultural, Forestry and Fishery workers	13,797 (79.3%)	15,161 (49.4%)
7 Craft and trade workers	547 (3.1%)	633 (2.1%)
8 Plant and machine opperators and assemblers	208 (1.2%)	241 (0.8%)
9 Elementary occupations	492 (2.8%)	543 (1.8%)
10 Unemployed	1,493 (8.6%)	1,691 (5.5%)
Missing	0 (0%)	11,382 (37.1%)

3 Missing data analysis

3.1 Missing patterns

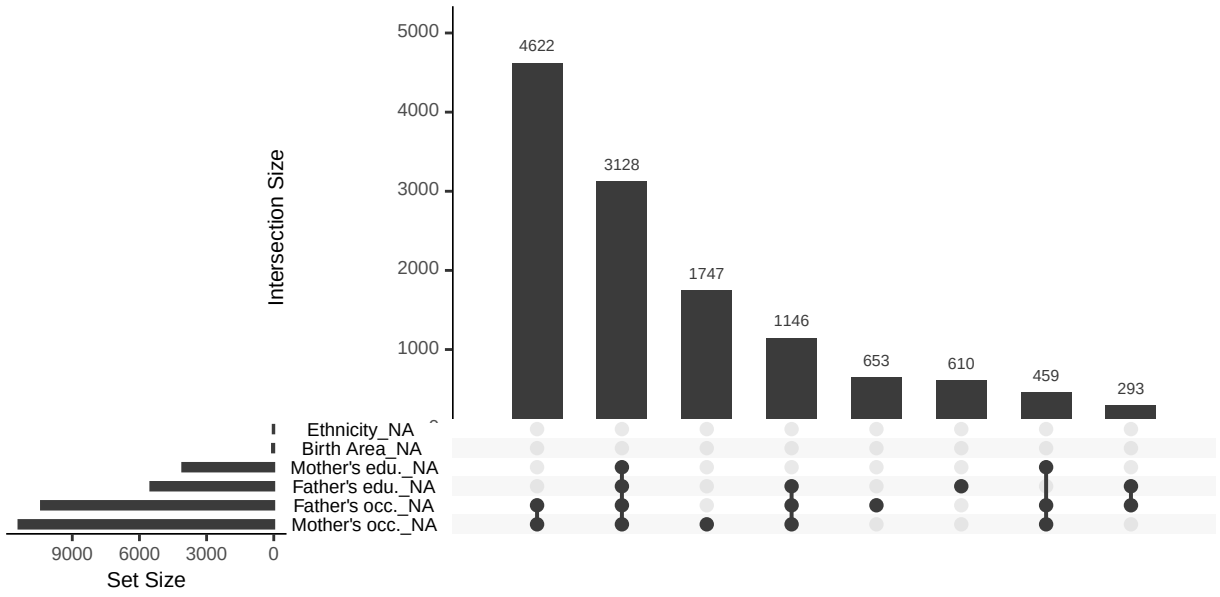

Figure 1: Missing patterns: *Left:* Marginal distribution of missing observations per variable. *Right:* Combination of missingness across cases

3.2 Differences in expected total equivalized household income between samples

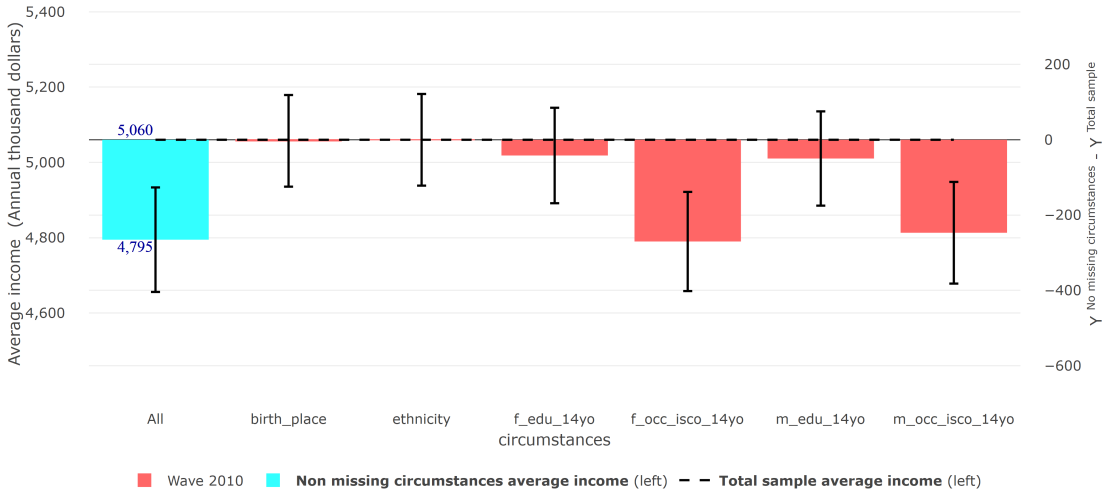

Figure 2: Differences in expected equivalized household income between the sample with non-missing circumstances and the total sample

3.3 Gini coefficient

Table 5: Gini coefficient in analysis sample and total sample

Wave	Sample	Gini	Lower bound	Upper bound	Average income
Wave 2010	Analysis sample	0.491	0.483	0.507	4,795
Wave 2010	Total sample	0.499	0.496	0.514	5,060

3.4 Differences in Gini coefficient between samples

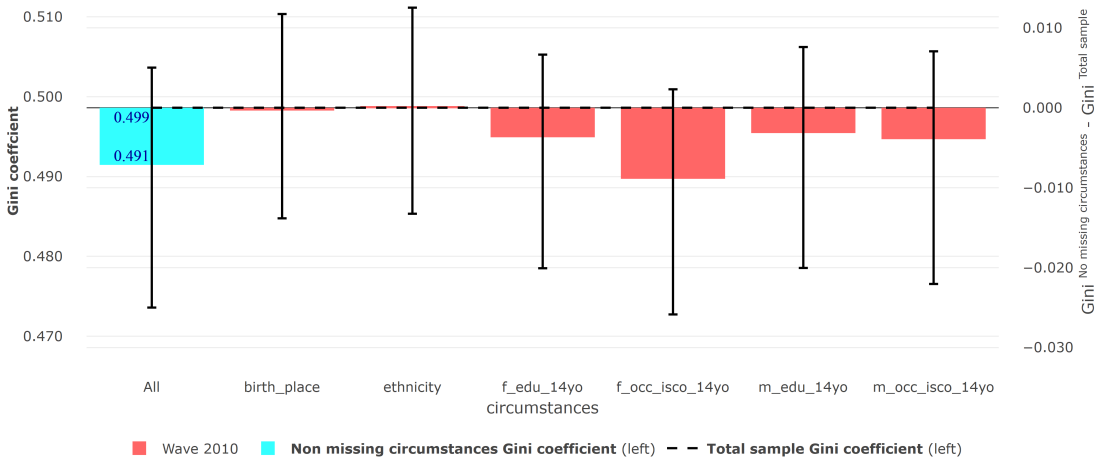

Figure 3: Differences in Gini coefficient between the sample with non-missing circumstances and the total sample