

United States 1998

1 Survey Description

Survey: The Panel Study of Income Dynamics (PSID) is the longest longitudinal household survey in the world. The Study started in 1968 and is currently being developed by the University of Michigan. It includes data related to employment, income, wealth, expenditures, health, marriage, childbearing, child development, philanthropy, education, and other topics. This document refers specifically to the data collected in 1998.

Link to the document: <https://simba.isr.umich.edu/data/data.aspx>

Sample: The sample of the PSID was originally composed of approximately 18,000 people in 5,000 households, making it a representative sample of the entire country. This sample was complemented with a non-representative oversample of low-income families to facilitate the investigation of issues related to poverty and vulnerability. Over the years, the sample size has changed as the number of members in the families considered originally has grown. There are 7,387 individuals in the total sample and 6,707 individuals in the analysis sample. Section 3 of this document describes the prevalence and pattern of missing data.

Weights: Population weights are computed as the ratio between specific weights for the Panel Study of Income Dynamics (PSID) and the Current Population Survey (CPS). These weights, based on sex, age, and ethnicity, are tailored for different demographic groups. Age groups are formed, and missing values are imputed using historical data. The process ensures a more representative combined dataset for analysis.

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

Circumstances:

- Race 'ethnicity'(white, black, american indian, aleut, eskimo, asian, pacific islander, hispanic and others described in Table 1)
- Sex (female, male)
- Region of upbringing (6 categories described in Table 1)
- Fathers's edu. (years of education, described in Table 2)
- Mother's edu. (years of education, described in Table 2)
- Father's occ. (3 categories, High: includes ISCO 1, 2 and 3, Medium: includes ISCO 4, 5 and 6, and Low: ISCO 7, 8, 9 and 0, described in Table 3)
- Mother's occ. (3 categories, High: includes ISCO 1, 2 and 3, Medium: includes ISCO 4, 5 and 6, and Low: ISCO 7, 8, 9 and 0, described in Table 3)

¹Income variable was equivalized using the square root scale.

2 Descriptive Statistics

Table 1: Respondant's socio-demographics - 1998

	Analysis sample	Total sample
	(N=6,707)	(N=7,387)
Gender		
0 Female	3,672 (54.7%)	4,011 (54.3%)
1 Male	3,035 (45.3%)	3,376 (45.7%)
ethnicity		
1 White	4,762 (71.0%)	5,010 (67.8%)
2 Black	1,399 (20.9%)	1,667 (22.6%)
3 Am. Indian, Aleut, Eskimo	119 (1.8%)	138 (1.9%)
4 Asian, Pac. Islander	130 (1.9%)	173 (2.3%)
5 Hispanic	40 (0.6%)	71 (1.0%)
7 Other	257 (3.8%)	328 (4.4%)
Region of upbringing		
1 Northeast	1,161 (17.3%)	1,204 (16.3%)
2 North central	1,891 (28.2%)	1,989 (26.9%)
3 South	2,170 (32.4%)	2,394 (32.4%)
4 West	818 (12.2%)	860 (11.6%)
5 Alaska, Hawaii	4 (0.1%)	4 (0.1%)
6 Foreign country	663 (9.9%)	806 (10.9%)
Missing	0 (0%)	130 (1.8%)

Table 2: Parental education - 1998

	N	Mean	SD	Median	Min	Max	Missing
Analysis sample - Mother's edu	6,707	3.748	2.254	4	0	8	0
Analysis sample - Father's edu	6,707	3.786	2.045	4	0	8	0
Total sample - Mother's edu	7,387	3.675	2.289	4	0	8	183
Total sample - Father's edu	7,387	3.722	2.090	4	0	8	99

Table 3: Parental occupation - 1998

	Analysis sample	Total sample
	(N=6,707)	(N=7,387)
Father's occupation (ISCO)		
1 Low	2,658 (39.6%)	2,950 (39.9%)
2 Middle	2,239 (33.4%)	2,343 (31.7%)
3 High	1,810 (27.0%)	1,886 (25.5%)
Missing	0 (0%)	208 (2.8%)
Mother's occupation (ISCO)		
1 Low	4,570 (68.1%)	4,857 (65.8%)
2 Middle	503 (7.5%)	535 (7.2%)
3 High	1,634 (24.4%)	1,745 (23.6%)
Missing	0 (0%)	250 (3.4%)

Table 4: Respondant's income - 1998

	N	Mean	SD	Median	Min	Max	Missing
Analysis sample	6,707	42,838	71,721	33,331	14.11	4,312,815	0
Total sample	7,387	41,685	69,478	32,373	14.11	4,312,815	0

3 Missing data analysis

3.1 Missing patterns

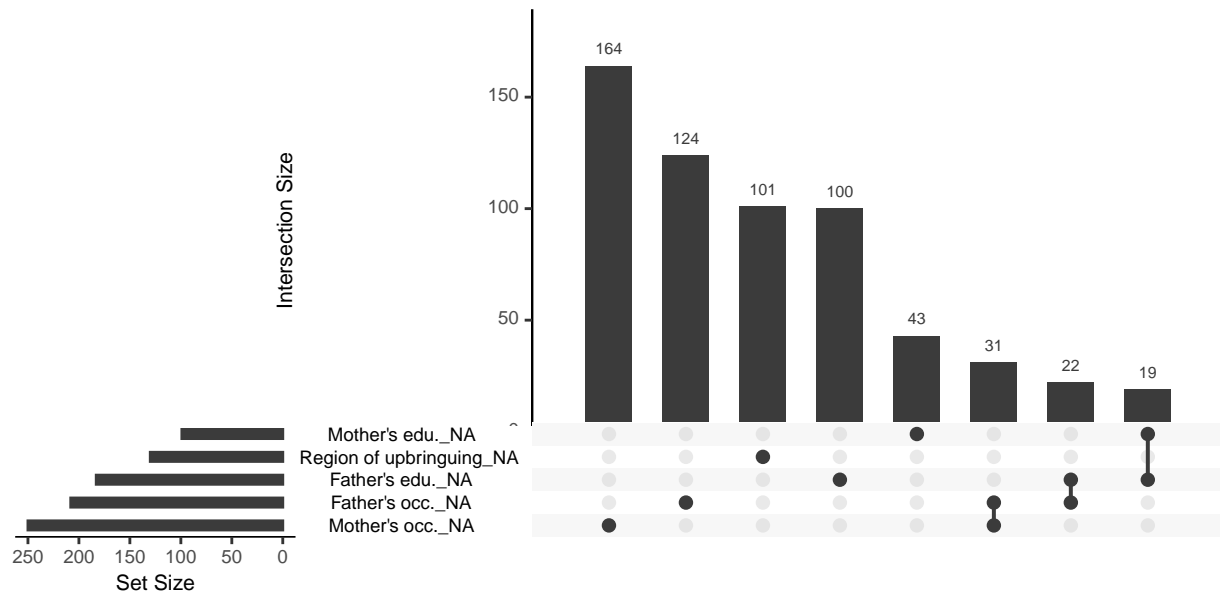


Figure 1: Missing patterns: *Left*: Marginal distribution of missing observations per variable. *Right*: Combination of missingness across cases

3.2 Differences in expected total equivalized household income between samples

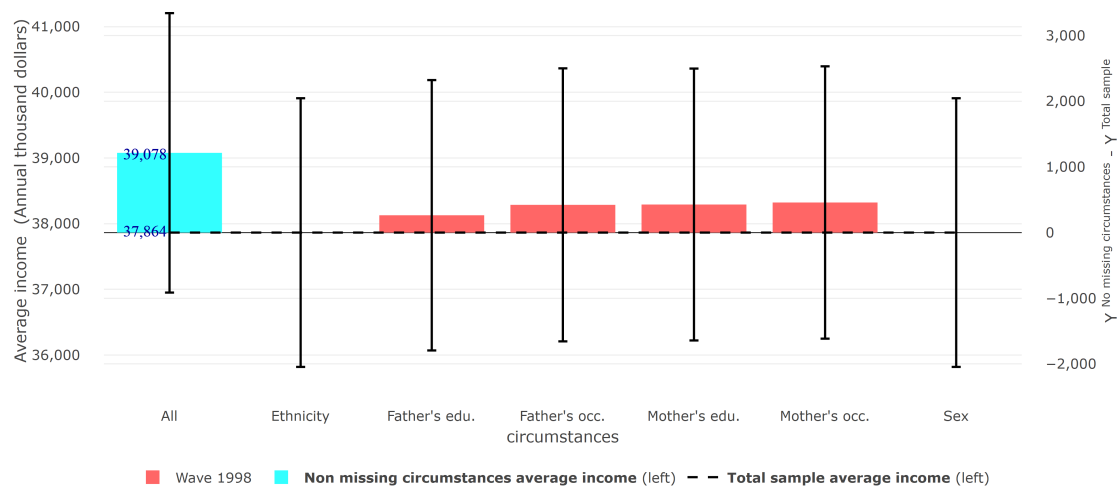


Figure 2: Differences in expected total equivalized household income between the sample with non-missing circumstances and the total sample

3.3 Gini coefficient

Table 5: Gini coefficient in analysis sample and total sample

Wave	Sample	Gini	Lower bound	Upper bound	Average income
Wave 1998	Analysis sample	0.401	0.385	0.432	39,078
Wave 1998	Total sample	0.404	0.389	0.434	37,864

3.4 Differences in Gini coefficient between samples

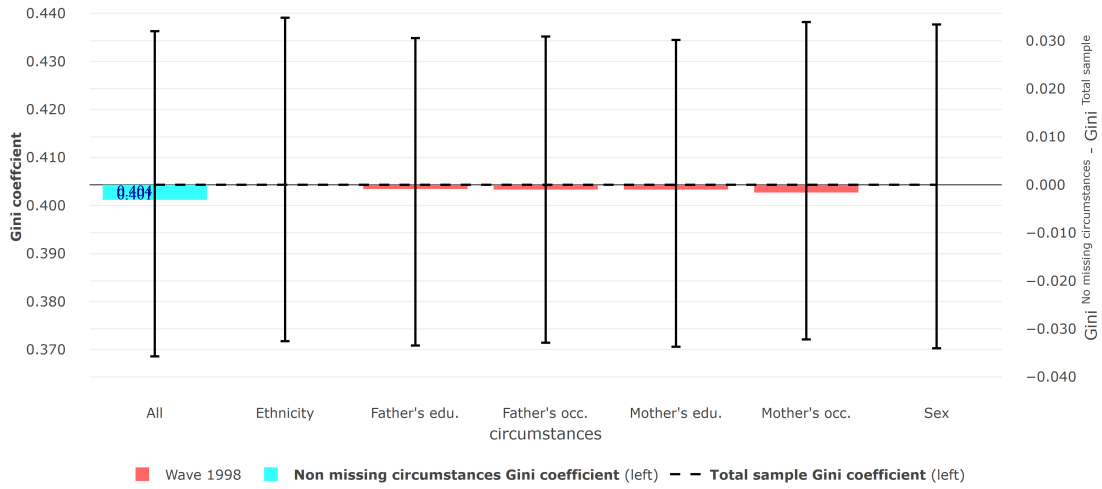


Figure 3: Differences in Gini coefficient between the sample with non-missing circumstances and the total sample