

# United States 2006

## 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 2006.

**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 9,384 individuals in the total sample and 8,289 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.<sup>1</sup>

**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)

---

<sup>1</sup>Income variable was equivalized using the square root scale.

## 2 Descriptive Statistics

Table 1: Respondant's socio-demographics - 2006

	Analysis sample	Total sample
	(N=8,289)	(N=9,384)
<b>Gender</b>		
0 Female	4,702 (56.7%)	5,229 (55.7%)
1 Male	3,587 (43.3%)	4,155 (44.3%)
<b>ethnicity</b>		
1 White	5,103 (61.6%)	5,407 (57.6%)
2 Black	2,518 (30.4%)	3,182 (33.9%)
3 Am. Indian, Aleut, Eskimo	152 (1.8%)	175 (1.9%)
4 Asian, Pac. Islander	120 (1.4%)	144 (1.5%)
5 Hispanic	20 (0.2%)	27 (0.3%)
7 Other	376 (4.5%)	449 (4.8%)
<b>Region of upbringing</b>		
1 Northeast	1,261 (15.2%)	1,346 (14.3%)
2 North central	2,228 (26.9%)	2,447 (26.1%)
3 South	3,071 (37.0%)	3,570 (38.0%)
4 West	1,096 (13.2%)	1,191 (12.7%)
5 Alaska, Hawaii	6 (0.1%)	6 (0.1%)
6 Foreign country	627 (7.6%)	741 (7.9%)
Missing	0 (0%)	83 (0.9%)

Table 2: Parental education - 2006

	N	Mean	SD	Median	Min	Max	Missing
Analysis sample - Mother's edu	8,289	3.816	2.312	4	0	8	0
Analysis sample - Father's edu	8,289	3.861	2.146	4	0	8	0
Total sample - Mother's edu	9,384	3.751	2.341	4	0	8	384
Total sample - Father's edu	9,384	3.807	2.177	4	0	8	144

Table 3: Parental occupation - 2006

	Analysis sample	Total sample
	(N=8,289)	(N=9,384)
<b>Father's occupation (ISCO)</b>		
1 Low	3,415 (41.2%)	3,788 (40.4%)
2 Middle	2,674 (32.3%)	2,819 (30.0%)
3 High	2,200 (26.5%)	2,278 (24.3%)
Missing	0 (0%)	499 (5.3%)
<b>Mother's occupation (ISCO)</b>		
1 Low	5,219 (63.0%)	5,631 (60.0%)
2 Middle	650 (7.8%)	711 (7.6%)
3 High	2,420 (29.2%)	2,622 (27.9%)
Missing	0 (0%)	420 (4.5%)

Table 4: Respondant's income - 2006

	N	Mean	SD	Median	Min	Max	Missing
Analysis sample	8,289	45,494	58,001	33,714	24.53	1,107,570	0
Total sample	9,384	44,132	56,432	32,703	24.53	1,107,570	0

### 3 Missing data analysis

#### 3.1 Missing patterns

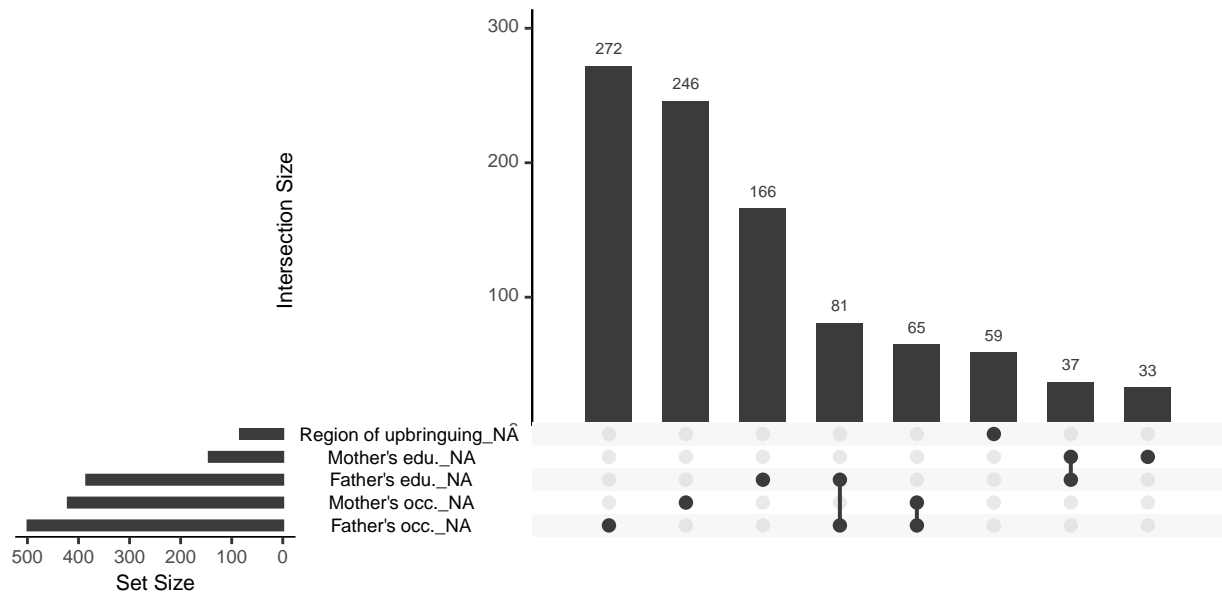


Figure 1: Missing patterns: *Left*: Marginal distribution of missing observations per variable. *Right*: Combination of missingness across cases

#### 3.2 Differences in expected total equivalized household income between samples

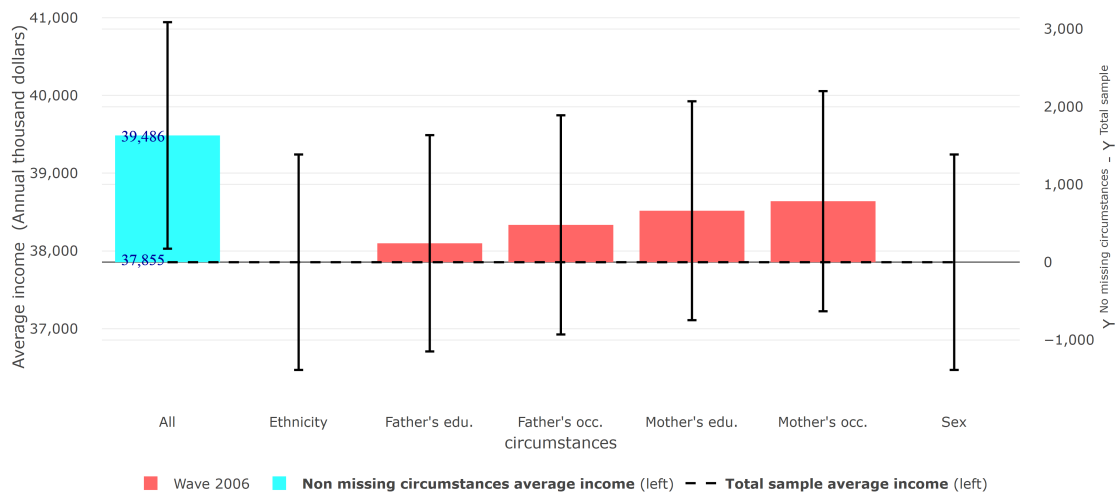


Figure 2: Differences in expected total equivalized household income between the sample with non-missing circumstances and the total sample

### 3.3 Gini coefficient

Table 5: Gini coefficient in analysis sample and total sample

Wave	Sample	Gini	Lower bound	Upper bound	Average income
Wave 2006	Analysis sample	0.442	0.425	0.456	39,486
Wave 2006	Total sample	0.443	0.429	0.458	37,855

### 3.4 Differences in Gini coefficient between samples

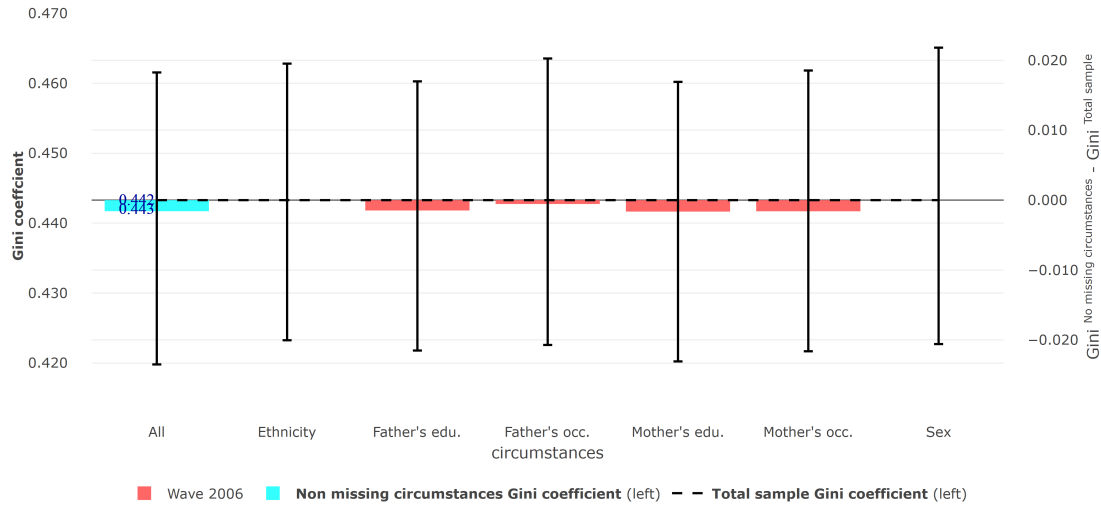


Figure 3: Differences in Gini coefficient between the sample with non-missing circumstances and the total sample