

Information on Annual Social and Economic Supplement to the Current Population Survey data

Aaron Gullickson

3/10/2017

The data we are using come from the 2017 Annual Social and Economic Supplement to the Current Population Survey (CPS-ASEC). The CPS is a survey of the US population that is collected every month by the US Census Bureau and the US Bureau of Labor Statistics. In March of every year, an additional questionnaire is fielded that asks more detailed social and economics questions. This data makes up the Annual Social and Economic Supplement to the Current Population Survey (CPS-ASEC). We will use this data to look at the gender wage gap. I have used the IPUMS-CPS website to create an extract of the CPS-ASEC for 2017 and have then recoded and cleaned the data for our purposes. The dataset that we will use is restricted to individuals aged 18-65 who worked at least 26 weeks in the previous year for at least an average of 10 hours/week.

Here are the variables in the dataset that we will use:

- **wage:** The hourly wage in US dollars of respondents in the previous year (2016). This number is based on the reported income from wages or salary for the previous year and the reported number of weeks worked and hours worked in a typical week for the previous year. Because some values can be extreme, any hourly wages greater than \$300/hour have been re-coded as \$300/hour (this is called top-coding) and wages less than \$1/hour have been re-coded as \$1/hour. Having an hourly wage does not require being paid by the hour. Hourly wages are calculated for salaried individuals as well by dividing income from salary by hours worked in the year.
- **sex:** Self-reported sex of the respondent: male (reference) or female.
- **age:** The age of the respondent.
- **nativity:** Whether the respondent was born in the US (reference) or is foreign-born.
- **degree:** The highest educational degree received by the respondent: No high school diploma (reference), high school diploma, associate's degree, bachelor's degree, or graduate degree.
- **nchild:** Number of own children living in the household with the respondent.
- **occup:** The broad occupational category of the respondent. In the actual CPS data, there are hundreds of different occupations listed. For our purposes, I have simplified this into a broader (and smaller) set of occupational categories that we will use for the analysis. Here are the categories of the occupational variable, along with some examples of specific occupations:
 - *Managers:* Human resources Managers, Operations Managers
 - *Business/Finance Specialist:* Claims Adjusters, Compliance Officers, Accountants, Tax Preparers
 - *STEM:* Computer Programmers, Civil Engineers, Biological Scientists
 - *Doctors:* Dentists, Surgeons, Optometrists
 - *Legal:* Lawyers, Judges, Paralegals
 - *Education:* Preschool and Kindergarten Teachers, Librarians
 - *Arts, Design, and Media:* Artists, Dancers and Choreographers, Writers and Authors
 - *Other Healthcare:* Registered Nurses, Physical Therapists, Dental Hygienists
 - *Social Services:* Clergy, Social Workers
 - *Service:* Waiters and Waitresses, Barbers, Bartenders
 - *Sales:* Cashier, Telemarketer
 - *Administrative Support:* Bank Tellers, Data Entry Keyers, Receptionist
 - *Other Manual:* Carpenters, Logging Workers, Mining Machine Operators, Small Engine Mechanic