

Labor Market Data Documentation

1 DATA SOURCES

1.1 ECONOMIC CONDITIONS

- **Source:** Decennial Census (1980 5% sample, 1990 5% sample, 2000 5% sample); American Community Survey (2010 3-year file, 2011 3-year file, 2019 3-year file, 2020 3-year file)
- **Download link:** All data are downloaded from IPUMS (<https://usa.ipums.org/usa-action/variables/group>)
 - Downloaded on April 14, 2022.
- **Variables used (raw variable names from IPUMS download):**
 - **gq:** Group quarters status
 - **age:** Age
 - **educ:** Educational Attainment
 - **perwt:** Person weight
 - **empstat:** Employment status
 - **wkswork2:** Weeks worked last year
 - **uhrswork:** Usual hours worked per week
 - **incwage:** Wage and salary income
 - **incbus:** Business income
 - **incfarm:** Farm income
 - **incbus00:** Business and farm income
 - **incbusfm:** Business and farm income
 - **race:** Race
 - **hispan:** Hispanic origin
 - **bpl:** Birthplace
 - **citizen:** Citizenship status
 - **sex:** Sex
 - **statefip:** State FIPS code
 - **puma:** Public Use Microdata Area
 - **occ:** Occupation code
 - **ind:** Industry Code

2 DATA PREPARATION

Key datasets

- **Commuting Zone Crosswalk, Industry Crosswalk, Occupational Code Crosswalk:** Crosswalks with 1990 values published by Professor David Dorn ([link](#)).

2.1 ECONOMIC CONDITIONS

- Because county information isn't consistently available across ACS/Census files, we crosswalk on Public Use Microdata Areas ("puma"). Using the state FIPS code and the "puma", we crosswalk to commuting zones using the Commuting Zone Crosswalk.
- We aggregate population, employment, weeks worked, hours worked, wage and salary income, and total earned income to the commuting zone level by sex-age-education-race-foreign born demographic cells. We condense some of the demographic cells as follows:
 - Drop any observations that correspond to the population living in institutions.
 - Drop any observations younger than 16 and older than 64, and group ages: 18-24, 25-39, 40-54, and 55-64.
 - Group education: Less than 4 Years of College ("educ" less than or equal to 100) and 4+ years of college ("educ" in 101-116).
 - Group on race: non-Hispanic white (race = 1 & hispan = 0) and other (all else).
 - Group on foreign-born status: foreign-born (born outside United States or outlying areas/territories, bpl > 120; exclude those born abroad to American parents, citizen > 1) and native born (all else).
 - Create employment using values of "empstat" = 1 (i.e. individuals reporting that they are employed).
 - Create annual weeks worked using the midpoint of the range of weeks worked in "wkswork2".
 - Create annual hours worked by multiplying our weeks worked variable by the usual "uhrswork".
 - Create a total income earned variable by summing wage and salary income with business and farm income.
 - Crosswalk all occupation and industry codes using the Occupational Code Crosswalk and Industry Crosswalk.
- We append the aggregated files for each year (1980, 1990, 2000, 2010, 2011, 2019, 2021) together.
- We then construct the following variables:
 - **emprate:** The employment rate defined as the employed population divided by the population.
 - **nonemprate:** The nonemployment rate defined as 1 less the employment rate.
- Finally, we crosswalk in each state that a commuting zone belongs to for clarity in situations when a commuting zone spans across multiple states.
- The final product contains the below list of constructed variables:
 - **year:** The ACS/Census year that the data corresponds to

- **cz_name:** The name of the commuting zone
- **czone:** The numerical code of the commuting zone (1990 definition)
- **fips_code1:** The FIPS code associated with the first state that the commuting zone is in
- **state1:** The name of the first/primary state that the commuting zone is in
- **fips_code2:** The FIPS code associated with the second state that the commuting zone is in (if applicable)
- **state2:** The name of the second state that the commuting zone is in (if applicable)
- **fips_code3:** The FIPS code associated with the third state that the commuting zone is in (if applicable)
- **state3:** The name of the third state that the commuting zone is in (if applicable)
- **sex:** Sex (male, female)
- **age:** Age category (four categories described above)
- **ed:** Education level (two categories described above)
- **race:** Race group (non-Hispanic white, other)
- **foreign:** Foreign born status (foreign born, native born)
- **population:** Population count
- **employed:** Employed population count
- **emprate:** Employment rate
- **nonemprate:** Non-employment rate
- **weeks:** Total number of weeks worked
- **hours:** Total number of hours worked
- **incwage:** Total wage and salary income earned
- **inccarn:** Total income earned
- **incbusfm:** Total business and farm income earned