This file contains intermediate results for the paper “The Contribution of National Income Inequality to Regional Economic Divergence.” It contains descriptive statistics and measures of sigma divergence using different income definitions and geographies.

“ipums\_dev\_[geography].csv” - Descriptive sigma divergence statistics across geographies (MSAs or Commuting Zones) using IPUMS Census microdata for a variety of income definitions.

Contains the following variables:

* wsd - Population-weighted coefficient of variation
* wsdlog - Population-weighted standard deviation of log income
* wp10 - Population-weighted 10th percentile of metropolitan average income, as a fraction of the national average
* wp25 - Population-weighted 25th percentile of metropolitan average income, as a fraction of the national average
* wp50 - Population-weighted 50th percentile of metropolitan average income, as a fraction of the national average
* wp75 - Population-weighted 75th percentile of metropolitan average income, as a fraction of the national average
* wp90 - Population-weighted 90th percentile of metropolitan average income, as a fraction of the national average
* natstat - National average income in 2015 dollars
* wiqr - Difference between wp75 and wp25
* wd19 - Difference between wp90 and wp10
* wr95 - Ratio of wp90 to wp50
* wr51 - Ratio of wp50 to wp10
* wr15 - Ratio of wp10 to wp50
* wr91 - Ratio of wp90 to wp10
* samp - Income definition used. Values:
  + fam - family income
  + famsqrt - family income / square root of family size
  + famnorm - family income / family size
  + hh - household income
  + hhsqrt - household income / square root of household size
  + hhnorm - household income / household size
  + adult - individual income, ages 18+ only
  + adultmale - individual income, males ages 18+ only
  + adultfemale - individual income, females ages 18+ only
* stat - Statistic used for columns B-O, mean or median

“bea\_dev\_[geography].csv” – Descriptive sigma divergence statistics using BEA Regional Economic Accounts per capita personal income data. Variables are defined as for “ipums\_dev\_[geography].csv.”

“cf\_[geography].csv” – Counterfactual sigma divergence statistics computed in simulations holding either the geographical distribution of individuals at different income levels or the level of income inequality constant at 1980 levels.

Contains the following variables:

* year - Year
* sd\_actual – observed coefficient of variation
* sd\_gXX\_iXX – coefficient of variation if geography and inequality were both as in reality (note this closely matches sd\_actual)
* sd\_g80\_i80 – coefficient of variation if geography and inequality had both remained at 1980 values
* sd\_g80\_iXX – coefficient of variation if geography was held constant at 1980 value but inequality changed as in reality.
* sd\_gXX\_i80 – coefficient of variation if geography changed as in reality but inequality was held constant at 1980 level.
* iqr\_[scenario] – counterfactual scenarios with interquartile range as the statistic of interest
* p19\_[scenario] – counterfactual scenarios with 10-90 range as the statistic of interest
* samp – income definition, defined as above
* psize – size, in percentiles, of buckets into which incomes are grouped prior to conducting simulations (see text)
* popyr – year in which metro population is used for weighting statistics. “cur” = current year, “80” = 1980, “13” = 2013

“dt\_[geography].csv” – Counterfactual sigma divergence statistics computed when dropping the richest 1, 5, and 10 percent of the population.

Contains the following variables:

* year – Year
* sd – Observed coefficient of variation
* sd1 – Coefficient of variation dropping richest 1%
* sd5 – Coefficient of variation dropping richest 5%
* sd10 – Coefficient of variation dropping richest 10%
* iqr[scenario] – Scenarios using interquartile range
* p19d[scenario] – Scenarios using 10-90 range.
* samp – Income definition, defined as above
* stat – Statistic used, median or mean