State Level Income Distributions for net income deciles for the US for historical years (2011-2014) and projections for different SSP scenarios (2015-2100)

Kelly Casper [i], Kanishka Narayan [i] & Brian O'Neill [i] & Stephanie Waldhoff [i]
i – Joint Global Change Research Institute (JGCRI), Pacific Northwest National Lab (PNNL), College Park, Washington DC

Summary

Income distributions are a growing area of interest in the examination of equity impacts brought on by climate change and its responses. We project US state level income distributions using a PCA-based approach, applying a downscaled version of the approach employed by Narayan et al. (2022, in-prep). A state-level dataset had to be synthesized and projected based on existing sources. We apply a PC-based model to our derived state-level dataset, employing projected GINI's from the SSP scenarios. We produce projected income distribution by income decile for three SSPs to year 2100. For the purpose of the projections, we developed a consistent set of tax adjusted net income deciles for all states from 2011 to 2014. This dataset was used for initialization of the projections and for validation.

Summary of outputs

Main output files include*,*

1. *state_decile_shares_gdp_per_capita_2011_2100.csv-* These are the projections of income shares along with the gdp per capita in 2011 USD for the 50 states and Washington D.C. Historical data ranges from 2011to 2015 and projections start in 2020.
2. *US_aggregated_income_distribution_2015_2011.csv-* These are the projections of the decile level GDP per capita along with population and GDP at PPP (all at the decile level) for the US as a whole aggregated from 1. above.

Description of outputs

   A. *state_decile_shares_gdp_per_capita_2011_2100.csv-* This file includes the following columns,

   1. **state**- This is the state code
   2. **state_name**- This is the state name
   3. **year-** This is the year which ranges from 2011-2100. Annual values from 2011 to 2015- and five-year time steps from 2015-2100.
   4. **sce-** This is the scenario. This includes **Historical data**, **SSP2**, **SSP3** & **SSP5**
   5. **shares-** These are the income shares. They are presented as percentages range from 0 to 100 percent
   6. **category-** This is the population decile for each state year combination. Ranges from d1 to d10
   7. **decile_pop-** This is the population in each decile in each state. This is the state level population divided by 10.
   8. **decile_gdp- This** is the decile GDP for each state in 2011 USD at PPP. This is the state level USD multiplied by the shares in point no. 5 above.
   9. **decile_gdp_per_capita-** This is the gdp per capita for each decile in each state in 2011 USD. This is the decile_gdp from point no. 8 above divided by decile_pop from point no 7.

B. *US_aggregated_income_distributiins.csv-* This file includes the following columns,

1. **GCAM_region_ID** – Unique identifier for the US. (Region ID =1)
2. **year-** This is the year which ranges from 2015-2100.
3. **sce-** This is the scenario. This includes **Historical data**, **SSP2**, **SSP3** & **SSP5**
4. **shares-** These are the income shares. They are presented as percentages range from 0 to 100 percent
5. **category-** This is the population decile for each state year combination. Ranges from d1 to d10
6. **tot_pop-** This is the total national population in the US
7. **tot_gdp-** Th**is** is the total national GDP in 2011 USD at PPP.
8. **gdp_per_capita-** This is the average gdp per capita for the US in 2011 PPP.

Reproduction instructions

The above code and data can be reproduced using code available at -
https://github.com/JGCRI/pridr/pull/3/files

Software used

Projections were generated using a PCA based model coded in R, namely the **pridr** software package which is available on GitHub here- https://github.com/JGCRI/pridr.

Other R packages used in the generation of these files are **tidyr**, **ggplot**, **dplyr**. The current code is compatible with latest versions of these software packages.

Recommended version of R is version 4.0 or above. All versions of R are available for download here- https://cran.r-project.org/bin/windows/base/

R studio is recommended for ease of use but not required.

Input data used

Following chart (Figure 1) shows all input data used in generation of these data & projections with appropriate citations

Legend
Input · Output

- National level GDP per capita under the SSPs[3]
- State baseline GDP per capita[3,4]
- State level GDP growth rates[4]
- National level GINI growth rates for the SSPs[2]
- State level GINI in 2015[6]
- State GDP per capita projections
- State SSP population projections[5]
- State GINI projections
- *Deciles generated using PCA algorithm[1]*
- State GDP projections
- State decile projections for shares of income and population
- State decile GDP per capita projections

Data Sources
1. Narayan & O'Neill et al. (2022; in prep)
2. Rao et al. (2019)
3. SSP Database
4. AEO-2020
5. Jiang et al. (2019)
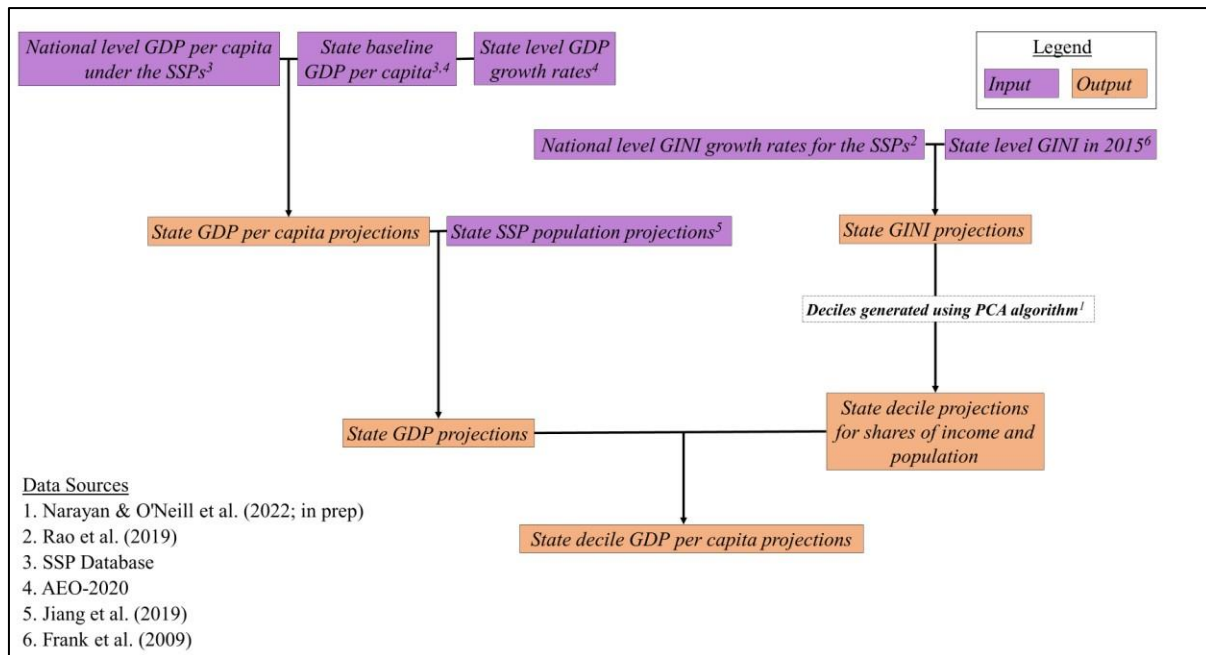6. Frank et al. (2009)

Figure 1. A flow chart of our methodology and data sources. This approach mirrors that of the global model presented in Narayan & O'Neill et al. (2022; in prep).