**Contact**

Madeleine Parker, [madeleine.parker@berkeley.edu](mailto:madeleine.parker@berkeley.edu)

Kamene Dornubari-Ogidi, [k.ogidi@berkeley.edu](mailto:k.ogidi@berkeley.edu)

Matt Nissen, [mnissen1@stanford.edu](mailto:mnissen1@stanford.edu)

**Data Sources**

Decennial, ACS

**Year**

2000, 2013-2017

**Geographies**

All 4 SPARCC sites, Tract-level

**Datasets**

Original datasets: By year, [2000\_hhincome\_race.dta](https://drive.google.com/open?id=1EEnH4gtdHrHXdcm6e4ZsQCONsl97sAGP) and [2017\_hhincome\_race.dta](https://drive.google.com/open?id=1ZGIBe4Fucu3BU3_8nVVmTDJ7A9FTDPJT), in each [year] folder

Processed datasets: Format: [race]\_inc\_[year].dta, found in [Income interpolation by race](https://drive.google.com/open?id=1a_mySKUL9DBkB7CIf_GTB70fZm0LxDh0) folder

**Processing**

Income interpolation by race and by year, using first the [Income\_interpolations\_race\_cleaning.do](https://drive.google.com/open?id=1mK10zGDMOx0jp0PDGvM8vDyM_ZzuauWo) file and then the 2 year do-files: [income\_interpolation\_2017.do](https://drive.google.com/open?id=1MX7LiSsBkQfoY1XgHSMdcJXC3glGvjx-) & [income\_interpolation\_2000.do](https://drive.google.com/open?id=1MyTFZ_N9-TSU3567yrm0qUIhwsajYiNW)

Merge process files (KDO & MN)

*MERGE\_allracesxinc00\_17\_append* [*Link*](https://drive.google.com/open?id=1u4Fhjzsf1KtHdm4jMV-beEvfwjqXNaKw)

* Do file for merge across races for every year. Concludes with code for appending racexincome data for all regions (a process done for each year separately). Also creates CV flag based on ACS and determines population change by race and income group.

*DTA files for each region’s race x income merge - separated by year (in* [*this*](https://drive.google.com/open?id=1kjpj1Hy4u9IMZIrSrBBijxwTjzNqCYgp) *folder)*

* [2000](https://drive.google.com/open?id=1aWB7NSJb4KjSn9Q0k52hkzX-PMCj-x3c)
* [2017](https://drive.google.com/open?id=1l0LyKozGImFTWatxpyZ7kQw6XIdmqy4V)

Final .dta file in main folder, link [here](https://drive.google.com/open?id=1I6kwYEDJe0FVpU1vtYIf5zpMIY2SeSRR).

**Variable Description**

[Used variables in final dataset]. Example for 2000; 2017 uses same format

E.g.:

FIPS Tract ID (2010)

COUNTY County FIPS Code (2010)

hinc\_00 Total median household income (in year)

vli2000\_[race] proportion 0.5 regional median household income

li2000\_[race] proportion 0.5 - 0.8 regional median household income

mi2000\_[race] proportion 0.8 - 1 regional median household income

hmi2000\_[race] proportion 1 - 1.2 regional median household income

hi2000\_[race] proportion 1.2 - 1.5 regional median household income

vhi2000\_[race] proportion 1.5 regional median household income

vli2000\_count\_[race] count 0.5 regional median household income

li2000\_count\_[race] count 0.5 - 0.8 regional median household income

mi2000\_count\_[race] count 0.8 - 1 regional median household income

hmi2000\_count\_[race] count 1 - 1.2 regional median household income

hi2000\_count\_[race] count 1.2 - 1.5 regional median household income

vhi2000\_count\_[race] count 1.5 regional median household income

total\_[race]\_00 Total # of households with [race] householder alone

total\_[race]\_popchange Population change of households with [race] householder

vli\_[race]\_popchange Pop change of hh; 0.5 regional median household income

li\_[race]\_popchange Pop change of hh; 0.5 - 0.8 regional median household income

mi\_[race]\_popchange Pop change of hh; 0.8 - 1 regional median household income

hmi\_[race]\_popchange Pop change of hh; 1 - 1.2 regional median household income

hi\_[race]\_popchange Pop change of hh; 1.2 - 1.5 regional median household income

vhi\_[race]\_popchange Pop change of hh; 1.5 regional median household income

*Only 2000*

vli2000\_se\_[race] standard error 0.5 regional median household income

li2000\_se\_[race] standard error 0.5 - 0.8 regional median household income

mi2000\_se\_[race] standard error 0.8 - 1 regional median household income

hmi2000\_se\_[race] standard error 1 - 1.2 regional median household income

hi2000\_se\_[race] standard error 1.2 - 1.5 regional median household income

vhi2000\_se\_[race] standard error 1.5 regional median household income

**Additional variables in merge files**

stderror\_[race]17\_total standard error total median household income for that race

hinc\_17\_[race] total median household income for that race

*Previously named hinc\_[year]*

hh\_cv17\_race coefficient variation (cv) for total household income

*cv= Total standard error/total household income*

hh\_cv17\_[race]\_15per\_flag coefficient variation is above 15%

hh\_cv17\_[race]\_30per\_flag coefficient variation is above 30%

**Additional Notes**

MN: Developed standard errors for 2000 data based on Census documentation [here](https://www.census.gov/prod/cen2000/doc/sf3.pdf) and [here](https://www.census.gov/prod/cen2000/phc-2-1-pt2.pdf). However, standard errors are not aggregated for AMI groups for 2017 ACS. Might be possible to aggregate these errors in income interpolation .do file, which would allow for individual ttests amongst racexincome groupings. However, original maps only relied on CV flags based on ACS data, decided to do the same process here. Standard errors for 2000 data is left in the data for future reference.