

## **JASA ACS Reproducibility Initiative - Author Contributions Checklist Form**

The purpose of the Author Contributions Checklist (ACC) Form is to document the code and data supporting a manuscript, and describe how to reproduce its main results.

As of Sept. 1, 2016, the ACC Form must be included with all new submissions to JASA ACS.

This document is the initial version of the template that will be provided to authors. The JASA Associate Editors for Reproducibility will update this document with more detailed instructions and information about best practices for many of the listed requirements over time.

## **Data**

### **Abstract (Mandatory)**

1. The file data1.txt contains the dataset we analyzed in Section 4.1. It is about the state-level poverty ratio for the school going related children for the 5- to 17-year old group in 1999 and was originally analyzed by Datta and Mandal (2015).
2. The file data1\_theta0.txt contains the ratio benchmarked census estimation state poverty ratio computed from the 2000 CPS population level poverty ratio. It is used to benchmark the estimates in Section 4.1 and is the same as those used in Datta and Mandal (2015).
3. The file data2.txt is the dataset we used in Section 4.2. It contains 5-year (2007-2011) county-level pooled ACS estimates of overall poverty rates, which was originally analyzed by Chakarborty et. al. (2016).
4. The files in the directory cb\_2013\_us\_county\_20m are the county-level cartographic boundary shapefiles downloaded from [https://www.census.gov/geo/maps-data/data/cbf/cbf\\_counties.html](https://www.census.gov/geo/maps-data/data/cbf/cbf_counties.html). They are used to produce the maps in Section 4.2.

### **Availability (Mandatory)**

Restrictions (if data will not be made publicly available, justify why not)

The datasets we analyzed are publicly available. The state-level data are extracted from poverty estimates of Current Population Survey 1999 and Census 2000. The county-level data can be downloaded from American Fact Finder <http://factfinder.census.gov>.

### **Description (Mandatory if data available)**

Permissions (demonstrate that author has legitimate access to data)

The authors have legitimate access to data since the datasets are publicly available or computed based on publicly available datasets.

Licensing information

Not available.

Link to data

A link to data will be created on the first author's once the manuscript is accepted for publication.

## File format

Both datasets are in plain text format.

## Metadata

1. data1.txt
  - Y: direct estimates of state-level child poverty ratio
  - X1: the number of child exemptions from IRS data
  - X2: IRS non-filer rate
  - X3: residuals from fitting a model for the 1989 census poverty data on X1 and X2
  - D: variances of sampling errors
2. data2.txt
  - Area: county ids
  - y: direct estimates of county level overall poverty rates
  - x: foodstamp participation rates
  - D: variances of sampling errors
  - FIPS: Federal Information Processing Standard county codes

## Code

### **Abstract (Mandatory)**

The file `samplers_and_functions.R` includes the Gibbs samplers and other functions used in the data analysis. The files `state_code.R` and `county_code.R` are R scripts used to produce the results in Section 6.1 and 6.2 respectively.

### **Description (Mandatory)**

How delivered (R package, Shiny app, etc.)

R scripts.

Licensing information (default is MIT License)

MIT License

Link to code/repository

A link to code will be created on the first author's once the manuscript is accepted for publication.

Version information

Not applicable.

### **Optional Information (complete as necessary)**

Supporting software requirements (e.g., libraries and dependencies, including version numbers)

To run the code provided, R and the following R libraries should be installed. The numbers in the parentheses indicate the version of the packages used when performing the analysis presented in the manuscript.

MASS (7.3-45), GIGrv (0.4), maptools (0.8-39), fields (8.4-1), RColorBrewer (1.1-2).

# Instructions for Use

## **Reproducibility (Mandatory)**

What is to be reproduced (e.g., "All tables and figure from paper", "Tables 1-4", etc.)

The code provided can be used to reproduce all the results in Section 6.1 and Section 6.2 including tables and figures.

How to reproduce analyses (e.g., workflow information, makefile, wrapper scripts)

To reproduce the results, put all the files provided in the working directory and execute the following commands in R:

```
> source("state_code.R")  
> source("county_code.R")
```

The analyses presented in the manuscript were performed on Intel Core i5 2.6GHz MacBook with 8GB memory. The execution time is about 5 minutes for `state_code.R` and 2 hours for `county_code.R`.