Figures Replication

This folder contains the Stata code to reproduce figures 1 to 3 in "A Practical Method to Reduce Privacy Loss when Disclosing Statistics Based on Small Samples" by Chetty and Friedman (2019).

To run the code, please set the file paths in in the do-file *all_figures_paper.do* to the files contained in this replication package (*tract_covariates.dta* and *tract_outcomes_early_dta_dp.dta)*.

Implementation of Noise-Infusion Algorithm: Example Code

The sub-folder *example_code_implementation_guide* contains two examples of how to implement the noise-infusion algorithm outlined in the Appendix A of "A Practical Method to Reduce Privacy Loss when Disclosing Statistics Based on Small Samples" by Chetty and Friedman (2019). The three files in the sub-folder contain the following:

Example 1 – Simple regression coefficient as statistic of interest

> *Implementation_guide_simple_reg.do*

> Stata do-file showing a step-by-step example of how to apply the noise-infusion algorithm to publicly release the estimated coefficients of a simple regression estimate of child income rank on parent income rank in each of the cells of a simulated dataset.

Example 2 – Predicted value of Y at a certain value of X as statistic of interest

> *Implementation_guide_p_25_prediction.do*

> Stata do-file showing a step-by-step example of how to apply the noise-infusion algorithm to publicly release the predicted value of child income rank at the 25th percentile of the parental income distribution in each of the cells of a simulated dataset.

Dataset

> *private_data_by_cells.dta*

> Simulated dataset in Stata format containing information on child income rank and parent income rank of 10,000 fictitious individuals grouped in 111 cells. Variable names: parent_rank, kid_rank, cell.