

Data and Code for: Temporal Aggregation for the Synthetic Control Method
Liyang Sun, Eli Ben-Michael, and Avi Feller

Overview

The code in this replication package constructs the analysis file from only one data source (Bell, Stuart and Gemmill (2023)) using R. The same R markdown file runs all of the code to generate the 3 figures in the paper. The replicator should expect the code to run for less than 1 hour.

Data Availability and Provenance Statements

We are secondary data users; we did not generate the data. Rather, we accessed the data from the replication packages of Bell, Stuart and Gemmill (2023), who compiled live birth counts for 2016-2022 from the CDC WONDER (2024).

We certify that we have legitimate access to and permission to use the data used in this manuscript.

We license users to download, copy, and modify this material.

Details on each Data Source

Data.Name	Data.Files	Location	Provided	Citation
“US Monthly Live Births 2016-2022”	compileddata.csv	./	TRUE	Bell, Stuart and Gemmill (2023)

Data on monthly lively births were downloaded from the Natality on CDC WONDER Online Database CDC WONDER (2024). We use 2016-2022 expanded version following Bell, Stuart and Gemmill (2023). Data can be downloaded from <https://wonder.cdc.gov/natality-expanded-current.html>, under “Births” and “Group Results By”, select “State of Residence”, then download. A copy of the data is provided as part of this archive. The data are in the public domain.

Datafile: `compileddata.csv`

Dataset list

Data file	Source	Notes	Provided
<code>compileddata.csv</code>	All listed	Raw data, serves as input for Figure 1-3.	Yes

Computational requirements

Software Requirements

- R 3.4.3
 - tidyverse (2.0.0)
 - augsynth (0.2.0)
 - lubridate (1.9.3)
 - ggrepel (0.9.5)
 - ggpubr (0.6.0)

Controlled Randomness

No simulation is involved in analysis, and thus, there is no need for specifying a random seed.

Memory and Runtime Requirements

Summary Approximate time needed to reproduce the analyses on a standard 2021 desktop machine is 1-5 minutes.

Details The code was last run on a **6-core Intel-based laptop with MacOS version 11.7.3**.

Description of programs/code

- Program `time_aggrigation.rmd` will generate all figures in the manuscript and the online appendix.

Instructions to Replicators

- Download the data file referenced above. It should be stored in the same directory as `time_aggregation.rmd`. Create the `/figure` folder. No further action is needed on the replicator's part.
- Run `time_aggregation.rmd` to generate all figures.

Figure/Table #	Program	Output file
Figure 1	<code>time_aggregation.rmd</code>	<code>/figure/frontier_plot-1.pdf</code>
Figure 2	<code>time_aggregation.rmd</code>	<code>/figure/synth_yearly_monthly_plot-1.pdf</code>
Figure 3	<code>time_aggregation.rmd</code>	<code>/figure/frontier_plot_estimates-1.pdf</code>

References

Bell, Suzanne O., Elizabeth A. Stuart, and Alison Gemmill. 2023. "Texas' 2021 Ban on Abortion in Early Pregnancy and Changes in Live Births." *JAMA*, 330(3): 281–282.

CDC WONDER. 2024. "Births Data Summary." Natality 2016-2022 (Expanded).

Acknowledgements

Some content on this page was adapted from Bell, Stuart and Gemmill (2023), Supplementary data, with the author's permission.