Data and computation of results for "In Utero Exposure to the Great Depression is Reflected in Late-Life Epigenetic Aging Signatures"

Lauren L. Schmitz[1] and Valentina Duque[2]

[1] Robert M. La Follette School of Public Affairs, University of Wisconsin-Madison

[2] School of Public Affairs, American University

August 2022

1. Data access and available data files

Health and Retirement Study (HRS) Data

This study used restricted individual level information on state of birth from the HRS and our contractual agreement does not permit public dissemination of the data before or after it has been merged with publicly available HRS data. Details on how to access restricted data for the HRS can be found at https://hrs.isr.umich.edu/data-products/restricted-data. Details on how to access the publicly available HRS data used in this study can be found at https://hrs.isr.umich.edu/data-products.

Historical State-Level Data

Publicly available, state-level data used in this study are posted on github in the folder `laurenschmitz/great-depression-epigenetic-aging/data`. The code used to clean the data prior to being merged with the HRS data are found in `/project-code`.

Wage index data file: `Salary_Data_w_ETS_FINAL.dta`

- `Pre_Analysis_Adjustments_to_Wage_Index.do` cleans the data so they can be merged with the HRS, transforms the raw wage index data to real terms, and creates event time study variables needed for the analysis.
- *Data source*: U.S. Bureau of Economic Analysis, SAINC7H Wages and Salaries by Industry (Historical) 1929-1957

Car sale index data file: `Car_Sale_Index_Data_w_ETS_FINAL.dta`

- `Pre_Analysis_Adjustments_to_Car_Sales_Index.do` cleans the data so they can be merged with the HRS, transforms the raw car sale data into an index, and creates event time study variables needed for the analysis.
- *Data source:* J. K. Hausman, Fiscal Policy and Economic Recovery: The Case of the 1936 Veterans' Bonus. *Am. Econ. Rev.* 106, 1100–1143 (2016).

Employment index data file: `Employment_Index_Data_w_ETS_FINAL.dta`

- `Pre_Analysis_Adjustments_to_Employment_Index.do` cleans the data so they can be merged with the HRS and creates event time study variables needed for the analysis.
- *Data source:* J. J. Wallis, Employment in the Great Depression: New data and hypotheses. *Explor. Econ. Hist.* 26, 45–72 (1989).

Maternal mortality rate data file: `Maternal_Mortality_Rate_Data_FINAL.dta`

- `Pre_Analysis_Adjustments_to_Maternal_Mortality_Rate.do` cleans the 1929 data so they can be merged with the HRS.
- *Data source:* S. Jayachandran, A. Lleras-Muney, K. V. Smith, Replication data for: Modern Medicine and the Twentieth Century Decline in Mortality: Evidence on the Impact of Sulfa Drugs. Nashville, TN: Am. Econ. Assoc. [publisher], 2010. Ann Arbor, MI: Inter-university Consort. for Pol. Soc. Rese [distributor], 2019 https:/doi.org/10.3886/E113743V1.

Manufacturing employment share data file: `Census_Manf_Share_Data_FINAL.dta`

- *Note*: Share of gainfully employed persons in the manufacturing sector was created by dividing the total number of wage earners in manufacturing by state in 1929 from the Census of Manufactures by the total number of gainfully occupied persons in that state in 1930 from the 1930 Census Occupation bulletin. After calculating the employment share, we generated an indicator equal to one if a state's share of employment in manufacturing was in the 75th percentile nationally in 1929 and zero otherwise for analysis.
- `Pre_Analysis_Adjustments_to_Manufacturing_Share.do` cleans the data and creates the manufacturing share variable.
- *Data sources*:
    - *Manufacturing share numerator*: U.S. Census Bureau, 1930 Census: Manufacturers, 1929 Volume 3. Reports by States. Statistics for Industrial Areas, Counties, and Cities. (1933). *Note:* Data were taken from the "Wage earners (average for the year)" column for 1929 for each state.
    - *Manufacturing share denominator*: U.S. Census Bureau, 1930 Census: Volume 4. Occupations, by States. Reports by States, Giving Statistics for Cities of 25,000 or More. United States Summary. (1933). *Note:* Data for each state were taken from Table 5, p. 18, Column 2, "Total gainfully occupied persons".

Infant mortality data file: `NBER_Infant_Mortality_Data_FINAL.dta`

- `Pre_Analysis_Adjustments_to_Infant_Mortality.do` cleans the data so they can be merged with the HRS and creates the infant mortality rate per 1000 births in 1928 for analysis.
- *Data source:* D. Norton, Data on Infant Mortality and Births, 1920-1945. Natl. Bur. Econ. Res. (2007).

Proportion farmland data file: `Proportion_Farmland_Census_of_Agriculture_FINAL.dta`

- *Data source*: U.S. Census Bureau, 1930 Census: Agriculture Volume 2. Reports by States, with Statistics for Counties and a Summary for the United States (1932). *Note:* Data on the proportion of land area in farms for each state was taken from "Tables for Geographic Divisions and States", Table 3, p. 24, Column 4, "Per ct. of approx. land area". Note that for Alabama, Arkansas, Nevada, North Dakota, and Oklahoma we used data on the percentage of farmland from the 1925 Census of Agriculture due to differences in how the data were originally sourced for this project. Numbers for these states are comparable and, in some cases, identical to data from the 1930 Census.

New Deal spending data file: `New_Deal_Spending_Data_FINAL.dta`

- `Pre_Analysis_Adjustments_to_New_Deal_Spending.do` cleans the data so they can be merged with the HRS and creates a total per capita New Deal relief spending variable.
- *Note:* To generate per capita state estimates, we aggregated data from the county level to the state level and divided by the state population in 1930. Following Fishback et al., total spending included spending on 1) *relief* from Federal Emergency Relief Administration grants, Civil Works Administration grants, Works Progress Administration grants, and Public Assistance grants (Social Security Act); 2) *public works* including Public Works

Administration federal and nonfederal grants and nonfederal loans, public roads administration grants, and public buildings administration grants; 3) *farm programs* including Agricultural Adjustment Administration grants, Farm Credit Administration loans, Farm Security Administration Rural Rehab grants, Farm Security Administration Rural Rehab loans**,** and Rural Electrification Administration loans; and 4) the *housing market* including Reconstruction Finance Corporation loans, Home Owners Loan Corporation loans, U.S. Housing Administration loan contracts, and U.S. Housing Administration grants. After calculating per capita spending, we generated an indicator for whether the share of New Deal spending was in the 75th percentile nationally for analysis.

- *Data source:* Fishback, Price, and Kantor, Shawn. New Deal Studies: New Deal Spending. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2018-11-18. https://doi.org/10.3886/E101199V1-24102

Rural electrification loans and climate data file:

`Fishback_Electrification_Climate_Data_FINAL.dta`

- `Pre_Analysis_Adjustments_to_Electrification_and_Climate_Data.do` cleans the data so they can be merged with the HRS and creates state-year estimates from county-year data.
- *Data source:* Fishback, Price, and Kantor, Shawn. New Deal Studies: Weather, Demography, and the New Deal. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2018-11-18. https://doi.org/10.3886/E101199V1-124306

World War II mobilization rate data file: `WWII_Mobilization_Rates_FINAL.dta`

- `Pre_Analysis_Adjustments_to_WWII_Mobilization_Rates.do` cleans the data so they can be merged with the HRS.
- *Note:* for analysis, we generated an indicator for high mobility equal to one if a state had a mobilization rate above the national average and zero otherwise.
- *Data source:* D. Acemoglu, D. H. Autor, D. Lyle. Women, War, and Wages: The Effect of Female Labor Supply on the Wage Structure at Midcentury. J. Polit. Econ. 112, 497–551 (2004). Link to data online: https://economics.mit.edu/faculty/dautor/data/autacemly06

2. Computation of results

The following Stata do files merge all needed data and produce our main results:

- `1_HRS_Data_Merge.do` merges the public and restricted use Health and Retirement Study (HRS) data.
- `2_HRS_Panel_Reshape.do` reshapes the data from wide to panel format.
- `3_Merge_Historical_State_Data_with_HRS_data.do` merges the historical state-level data with the HRS panel. The following .do files detailed above in Section 1 clean the historical data prior to this step so they can be merged with the HRS:

```
Pre_Analysis_Adjustments_to_Wage_Index.do
Pre_Analysis_Adjustments_to_Car_Sale_Index.do
Pre_Analysis_Adjustments_to_Employment_Index.do
```

`Pre_Analysis_Adjustments_to_Maternal_Mortality_Rate.do`
`Pre_Analysis_Adjustments_to_Manufacturing_Share.do`
`Pre_Analysis_Adjustments_to_Infant_Mortality.do`
`Pre_Analysis_Adjustments_to_New_Deal_Spending.do`
`Pre_Analysis_Adjustments_to_Electrification_and_Climate_Data.do`
`Pre_Analysis_Adjustments_to_WWII_Mobilization_Rates.do`

- `4_Clean_Data_and_Code_Variables_for_Analysis.do` cleans and codes the HRS and state-level data for analysis.

- `5_Primary_Analysis.do` computes the results for Tables 1-2, Figure 2, Tables S2-S15, and Tables S18-S20.

  Note: Coefficients in Tables 1-2, Figure 2, Tables S3-S5, S7-S8, and S18-S20 are reported as standardized effect sizes (i.e., Pearson's r). Coefficients were converted post analysis using the following formulas:

$$r = \hat{\beta} \times \frac{SD(X)}{SD(Y)}$$

  Where Y denotes the outcome, X is the exposure variable, and $\hat{\beta}$ is the effect coefficient obtained from a linear model regressing Y on X. The same scaling factor was applied to the standard error estimates and 95% confidence intervals:

$$SE(r) = SE(\hat{\beta}) \times \frac{SD(X)}{SD(Y)}$$

$$CI_{0.95} = [r - 1.96 \times SE(r), r + 1.96 \times SE(r)]$$

- `Figure_2_Pearson_r_Coefficient_Plot.do` plots standardized Pearson r coefficients for the preconception, in utero, and childhood periods estimated for Figure 2 in `5_Primary_Analysis.do`. Standardized effect sizes were calculated with respect to the SD of the wage index in utero and the SD of either GrimAge EAA or DunedinPoAm.

- `Census_1940_Fertility_and_Sex_Ratio_Analysis.do` computes results for Tables S16-S17 using the 1% representative sample of the 1940 Census.