# README for Beach and Hanlon “Culture and the Historical Fertility Transition” Replication

## Overview

## This document outlines the various components of the Culture and the Historical Fertility Transition replication package. We begin with a list of data sets and their sources. Next, we outline the software requirements. The programs to generate the results are split between “processing” programs, which prepare the data for analysis and “analysis” programs, which conduct the analysis. The instructions to the replicators outline the steps necessary to conduct the analysis. We end with a table that maps the various analysis programs to the corresponding manuscript figure/table.

## Data Availability and Provenance Statements

All of the data needed to replicate “Culture and the Historical Fertility Transition” are saved in the “Raw Data” directory. Below is a table listing each dataset and its source. The code to combine the data and generate the analysis is described later.

|  |  |
| --- | --- |
| **Data file** | **Notes** |
| mitchell\_data.dta | Authors’ transcription of Mitchell (2003). |
| newspaper\_hits.xlsx | This file lists results for two queries: “Monday” and “Bradlaugh OR Besant” conducted on August 24, 2020 for three newspaper database websites: newspapers.com, chroniclingamerica.loc.gov, and gale.com. The websites are either freely available or with a subscription, which anyone can obtain for a fee.  There is no specific citation, other than the websites themselves. |
| **Canada Data (contained in “Canada” subdirectory)** | |
| Canada\_1881\_Raw\_IPUMS\_00001.dta | This is an extract of the 1881 Canadian census from IPUMS.org, which we use to generate our microdata results. Data are free and publicly available. They can be downloaded via the IPUMS upon registration at: https://uma.pop.umn.edu/ipumsi/user/new?return\_url=https://international.ipums.org.  Data are cited as: Minnesota Population Center (2019). |
| County\_controls.dta | Series digitized by the authors from scanned copies of original *Census of Canada* manuscripts obtained from HathiTrust, Google Books and other online sources |
| CountyXCensus\_pop\_cats.dta | Series digitized by the authors from scanned copies of original *Census of Canada* manuscripts obtained from HathiTrust, Google Books and other online sources |
| Henripin\_Canada\_birth\_rates.xlsx | Authors’ transcription of Henripin (1968). |
| province\_county\_lat\_lon.dta | Latitude and longitude data for counties obtained from Google Maps.   There is no citation other than the website itself. |
| **England and Wales Data (contained in “England\_and\_Wales” subdirectory)** | |
| 1895\_Directory\_with\_BNA.dta | Authors’ transcription of Mitchell (1895). |
| BNA\_Index.xlsx | This index categorizes every article mentioning “Bradlaugh” or “Besant” in 1877, as identified by a query of britishnewspaperarchive.co.uk in December of 2019. The categorization is our own, as informed by reading each article.  Beach and Hanlon (Forthcoming) is the relevant citation. |
| BNA\_Transcribed\_Sample | This folder contains a sample of newspaper articles mentioning “Bradlaugh” or “Besant” in 1877, as identified by a query of britishnewspaperarchive.co.uk in December of 2019. The transcriptions were conducted by authors’ research team.  Beach and Hanlon (Forthcoming) is the relevant citation. |
| district\_births\_data\_standardized.dta | Data transcribed from annual reports of the Registrar General, aggregated into time-consistent registration districts by authors’ research team. Annual Registrar Reports are available for download from histpop.org.  Due to the aggregation, Beach and Hanlon (Forthcoming) is the relevant citation. |
| district\_pop\_by\_age\_combined.dta | District-level population tabulations from each census year based on published census reports and aggregated into time consistent districts by authors’ research team. These census reports are available for download from histpop.org.  Due to the aggregation, Beach and Hanlon (Forthcoming) is the relevant citation. |
| district\_shapefiles | This folder contains the district shapefiles for England and Wales from Satchell et al. (2018) |
| EW\_District\_Chars.dta | This dataset provides district-level data from the following sources:   * District latitude and longitude for main town or district center (if no main town) (authors calculations from Google Maps). * “articles” data based on keyword queries using British Newspaper Archive, as mentioned above. * “new\_papers” series and “established papers” are based on the 1895 Newspaper Press Directory (cited above and below as Mitchell, 1895) * The population series are from published census reports (as made available by histpop.org) * Births, marriages, and deaths from Registrar General’s reports (as made available by histpop.org) * Local economic structure data from the Census reports (as made available by histpop.org)   Each of the above sources were aggregated into time consistent districts by authors’ research team.  Due to the aggregation, Beach and Hanlon (Forthcoming) is the relevant citation. |
| marriage\_data\_district\_standardized.dta | From annual reports of the Registrar General (available from histpop.org), aggregated into time-consistent registration districts.  Due to the aggregation, Beach and Hanlon (Forthcoming) is the relevant citation. |
| Newspaper\_Monthly\_Queries.dta | Dataset lists the results for various keyword queries using britishnewspaperarchive.co.uk, as conducted in December of 2019. Data are available to anyone that purchases a subscription.  There is no citation other than the website itself. |
| Newspaper\_Yearly\_Queries.dta | Dataset lists the results for various keyword queries using britishnewspaperarchive.co.uk, as conducted in December of 2019. Data are available to anyone that purchases a subscription.  There is no citation other than the website itself. |
| WS\_a3\_1\_birth\_rate\_series.dta | Transcribed from Wrigley and Schofield (1981). |
| **South Africa Data (contained in “South\_Africa” subdirectory)** | |
| South\_Africa\_1875\_1891.dta | Data digitized by the authors from scanned copies of the original *Cape Colony Census* manuscripts (1875 and 1891 versions) obtained from HathiTrust, Google Books, and other similar online sources. |
| **U.S. Data (contained in “United\_States” subdirectory)** | |
| 1880\_US\_IPUMS\_extract.dta | This is an extract of the 1880 U.S. census microdata from IPUMS.org, which we use to generate our microdata results. Data are free and publicly available. They can be downloaded via the IPUMS upon registration at: https://uma.pop.umn.edu/ipumsi/user/new?return\_url=https://international.ipums.org.  Data are cited as: Ruggles et al. (2020). |
| Newspaper\_Mentions.xlsx | This file lists results for two queries: “Malthusian” and “Bradlaugh OR Besant” conducted on August 24, 2020 for three newspaper database websites: newspapers.com, chroniclingamerica.loc.gov, and gale.com. The websites are either freely available or with a subscription, which anyone can obtain for a fee.  There is no specific citation, other than the websites themselves. |

### Software Requirements

Stata (code was last run with version 16.1) and the following user-generated commands, which are currently stored in the root ado folder of the replication package.

* + unique
  + shp2dta
  + eststo, estadd, estout, esttab (all from st0085\_2 package)
  + ftools
  + reghdfe
  + psacalc
  + ivreg2
  + ranktest
  + ivreghdfe
  + sutex
  + texsave
  + geodist
  + spmap
  + ols\_spatial\_hac

## Description of programs

* Programs in the Processing folder prepare data for analysis.
* Programs in the Analysis folder generate all tables and figures.
* The program Analysis/Analysis\_Master\_Replication.do will run everything in the correct order. Each program called from master.do identifies the table or figure it creates (e.g., T1\_Canada\_agg\_main.do). Output files are called appropriate names (T1\_Canada\_agg\_main.tex) and should be easy to correlate with the manuscript.

## Instructions to Replicators

* Analysis/Analysis\_Master\_Replication.do needs to be modified to include the user’s “root” directory (where the replication folder is stored).
* Run Analysis/Analysis\_Master\_Replication.do to generate all of our results and save them in the Results folder.

### Details

* Each table and figure is saved and run as a separate program.
* If the user wants to run a program individually it is crucial to run the processing files from “Analysis\_Master\_Replication.do”. This is where the control variables are defined as a global that can be called on in subsequent do files.
* The programs were last run top to bottom on April 30, 2022.

## List of tables and programs

The below table lists the figures and tables in the order in which they appear in the manuscript and the appendix. The table includes the program name/location and the output files. The starting directory for all program locations is the “Analysis” folder. Unless otherwise stated, results are saved in the “Results” folder.

|  |  |  |  |
| --- | --- | --- | --- |
| Figure/Table | Program | Output file(s) | Note |
| Figure 1 | F1\_ew\_national\_time\_series.do | F1\_ew\_national\_time\_series\_a.pdf F1\_ew\_national\_time\_series\_b.pdf F1\_ew\_national\_time\_series\_c.pdf | Figure 1 consists of 3 panels and each panel is exported as a separate PDF. |
| Figure 2 | F2\_Canada\_time\_series.do | F2\_canada\_time\_series.pdf |  |
| Table 1 | T1\_Canada\_main\_agg.do | T1\_Canada\_agg.tex |  |
| Table 2 | T2\_Canada\_microdata.do | T2\_Canada\_mico.tex |  |
| Figure 3 | F3\_keyword\_time\_series.do | F3\_BNA\_article\_time\_series.pdf |  |
| Figure 4 | F4\_US\_keyword\_time\_series.do | F4\_US\_article\_time\_series.pdf |  |
| Table 3 | T3\_EW\_Balance.do | T3\_EW\_Balance.tex |  |
| Figure 5 | F5\_EW\_DD\_Directories.do | F5\_EW\_DD\_Newspapers.pdf |  |
| Table 4 | T4\_EW\_DD.do | T4\_EW\_Base.tex |  |
| Figure 6 | F6\_EW\_Falsification\_DD.do | F6\_EW\_Falsification.pdf |  |
| Table 5 | T5\_EW\_IV.do | T5\_EW\_IV.tex |  |
| Figure 7 | F7\_EW\_births\_long.do | F7\_ew\_cbr\_long.pdf |  |
| Figure 8 | F8\_EW\_Scotland\_graph.do | F8\_EW\_Scotland\_graph.pdf |  |
| Figure 9 | F9\_England\_France\_Germany.do | F9\_England\_France\_Germany.pdf |  |
| Figure 10 | F10\_Other\_Euro\_Fertility.do | F10\_Other\_Euro\_Fertility.pdf |  |
| Table 6 | T6\_Canada\_agg\_sum\_stats.do | T6\_Canada\_agg\_sum\_stats.tex |  |
| Table 7 | T7\_Canada\_micro\_sum\_stats.do | T7\_Canada\_micro\_sum\_stats.tex |  |
| Table 8 | T8\_Canada\_origin\_comparisons.do | T8\_Canada\_origin\_comparisons.tex |  |
| Table 9 | T9\_Canada\_agg\_robust.do | T9\_Canada\_Robust.tex |  |
| Table 10 | T10\_Canada\_spatial\_SE.do | T10\_Canada\_spatial\_SE.tex |  |
| Table 11 | T11\_Canada\_micro\_robust.do | T11\_Canada\_Micro\_Robust.tex |  |
| Table 12 | T12\_T13\_US\_Sum\_Stats.do | T12\_US\_Sum\_Stats.tex |  |
| Table 13 | T12\_T13\_US\_Sum\_Stats | N/A | Tabulations appear in results window of Stata |
| Table 14 | T14\_US\_Microdata.do | T14\_US\_Microdata.tex |  |
| Table 15 | T15\_US\_microdata\_windows.do | T15\_US\_microdata\_windows.tex |  |
| Table 16 | T16\_SA\_Summary\_Stats.do | T16\_SA\_Summary\_Stats.tex |  |
| Table 17 | T17\_SA\_main\_results.do | T17\_SA\_main\_results.tex |  |
| Table 18 | T18\_SA\_Spatial\_SE.do | T18\_SA\_Spatial\_SE.tex |  |
| Figure 11 | N/A |  | No replication code. Figure is a screenshot of raw data. |
| Figure 12 | N/A |  | No replication code. Figure is a screenshot of raw data. |
| Figure 13 | N/A |  | No replication code. Figure is a screenshot of raw data. |
| Figure 14 | F14\_newspaper\_databases.do | F14\_newspaper\_databases.pdf |  |
| Figure 15 | F15\_EW\_news\_maps.do | F15\_EW\_news\_maps\_a.pdf F15\_EW\_news\_maps\_b.pdf | Figure consists of two panels. Each panel is exported as separate PDF. |
| Table 19 | T19\_20\_BNA\_Selection.do | T19\_BNA\_entry.tex |  |
| Table 20 | T19\_20\_BNA\_Selection.do | T20\_BB\_Article\_Ratio.tex |  |
| Figure 16 | N/A |  | Figure 16 is a screenshot of an actual article. No code necessary. |
| Figure 17 | F17\_BNA\_map.do | F17\_BNA\_map.pdf |  |
| Table 21 | T21\_BNA\_classification.do | T21\_BNA\_classification.tex |  |
| Table 22 | N/A |  | These word stems were prepared by an outside vendor using newspaper article transcriptions provided by the authors (transcriptions available in Raw\_Data/England\_and\_Wales/BNA\_Transcribed\_Sample |
| Table 23 | T23\_T24\_EW\_Sum\_Stats.do | T23\_EW\_Sum\_Stats\_Full.tex |  |
| Table 24 | T23\_T24\_EW\_Sum\_Stats.do | T24\_EW\_Sum\_Stats\_DD.tex |  |
| Table 25 | T25\_EW\_Alt\_Windows.do | T25\_EW\_Alt\_Windows.tex |  |
| Table 26 | T26\_EW\_Spatial\_SE.do | T26\_EW\_Spatial\_SE.tex |  |
| Table 27 | T27\_EW\_Robust.do | T27\_EW\_Robust.tex |  |
| Figure 18 | F18\_EW\_Event\_Study.do | F18\_EW\_Event\_Study.pdf |  |
| Figure 19 | F19\_EW\_Marriages\_Births\_AR.do | F19\_EW\_Marriages\_Births\_AR.pdf |  |
| Table 28 | T28\_EW\_Marriages.do | T28\_EW\_Marriages.tex |  |
| Figure 20 | F20\_EW\_Mar\_Counterfactual.do | F20\_EW\_Mar\_Counterfactual.pdf |  |
| Table 29 | T29\_EW\_Mar\_Counterfactual.do | T29\_EW\_Mar\_Counterfactual.tex |  |

## References

Beach, Brian and W. Walker Hanlon. Forthcoming. “Culture and the Historical Fertility Transition.” *The Review of Economic Studies*.

Henripin, Jacques. 1968. Tendances et Facteurs de la Fecondite au Canada. Ottawa: Bureau Federal de la Statistique

Minnesota Population Center. Integrated Public Use Microdata Series, International: Version 7.2 [dataset]. Minneapolis, MN: IPUMS, 2019. https://doi.org/10.18128/D020.V7.2

Mitchell, Brian. 2003. International Historical Statistics, Europe 1750-2000. Hampshire, England: Palgrave Macmillan

Mitchell, Charles. 1895. *Newspaper Press Directory*. United Kingdom: C. Mitchell.

Satchell, M., Shaw-Taylor, L., Wrigley, E., Kitson, P., Newton, G. (2018). 1851 England and Wales census registration districts. [data collection]. UK Data Service. SN: 852948, DOI: 10.5255/UKDA-SN-852948

Steven Ruggles, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas and Matthew Sobek. IPUMS USA: Version 10.0 [dataset]. Minneapolis, MN: IPUMS, 2020. <https://doi.org/10.18128/D010.V10.0>

Wrigley, Edward and Roger Schofield. 1981. *The Population History of England 1541-1871*. Cambridge: Cambridge University Press.