This README file contains instructions for replicating the results in "When Did Growth Begin?" by Paul Bouscasse, Emi Nakamura, and Jón Steinsson. All results referenced in the paper can be replicated by running two files in R: *main_estimation.R* and *main_post.R*. The first one launches the estimation of the various specifications used in the paper; the second one produces the figures and tables.

**Data and Code Availability Statement**

All data used is available in ~\*bns_replication\data\raw*. The data was either (i) downloaded from publicly available websites, (ii) manually digitized from published or working papers referenced in our article, (iii) obtained from Gregory Clark in private correspondence, or (iv) assembled by Jón Steinsson for his lecture notes. Details for each raw data file are in appendix A below.

All code used is available in ~\*bns_replication\code*.

**Computation Requirements**

The replication requires the installation of R (version used: 4.3.1), RStudio (version used: 2024.09.0+375), Rtools (version used: Rtools43, GNU Make 4.4.1), and RStan (version used: 2.26.23). Instructions for the installation of RStan are available here:

https://mc-stan.org/users/interfaces/rstan

Other required R packages are: dplyr (version used: 1.1.3), ggpubr (0.6.0), reshape2 (1.4.4), broom (1.0.5), tidyverse (2.0.0), xtable (1.8.4), nleqslv (3.3.4), pracma (2.4.2), invgamma (1.1), ramify (0.3.3), stringr (1.5.0), patchwork (1.1.3), readxl (1.4.3), EnvStats (2.8.1), openxlsx (4.2.5.2), latex2exp (0.9.6), scales (1.2.1). Those are automatically installed if they are missing.

The estimation results were obtained on a Station Dell PRECISION 7920 bought in 2023 with 2 Intel(R) Xeon(R) Gold 5218R CPU @ 2.10GHz processors (20 cores/40 threads each) and 64Go of RAM. The computer's operating system is Windows 10 Enterprise, version 22H2, 64 bits.

We have used 24 chains of 1,000,000 draws for each specification. Each specification takes between 12 and 36 hours, provided chains can run in parallel.[1] The total running time on that computer was about 2 weeks if we ran specifications in turn. (Computing time can be cut if specifications themselves are run in parallel.) For replication with less computing power, we recommend cutting the number of draws (*iters)* and chains (*chain*) on lines 33 and 34 of *main_estimation.R*. In our experience, 200,000 draws and 8 chains is enough for reliable results.

**Instructions for Data Preparation and Analysis**

The estimation can be replicated by running *main_estimation.R*. The figures and tables can be generated by running *main_post.R.* Both files can be run independently since results from *main_estimation.R* have been saved to ~\*bns_replication\output\RData*. In both cases, the user needs to set *path0*, the path of the

---

[1]  Specifications with a structural estimation of the slope are the slowest.

*bns_replication* folder. Charts and tables are saved in ~\*bns_replication\output\charts* and ~\*bns_replication\output\tables*.

Appendix B below contains detailed explanations for each code file.

## Appendix A: List of Raw Data Files

file: allen05.xlsx

data: agricultural productivity

reference: Allen (2005), "English and Welsh agriculture"

note: digization of table 16


file: allen07.xlsx

data: wage and price

reference: Allen (2007), "Pessimism Preserved"

note: digization of appendices I and II


file: Allreturn2005.xls

data: interest rates

reference: Clark (2010), "Macroeconomic Aggregates for England"

note: shared by Gregory Clark in private correspondance on 06/09/2020


file: a-millennium-of-macroeconomic-data-for-the-uk.xlsx

data: Broadberry et al.'s population estimate

reference: Broadberry et al. (2015), "British Economic Growth, 1270–1870"

https://www.bankofengland.co.uk/statistics/research-datasets

downloaded: 09/13/2024


file: clark07.xlsx

data: population

reference: Clark (2007), "The Long March of History"

note: digization of table 9

file: clark10.xlsx

data: miscellaneous

reference: Clark (2010), "Macroeconomic Aggregates for England"

note: digization of tables 1, 13, and 33


file: England NNI - Clark - 2015.xlsx

data: wage and price

reference: Clark (2010), "Macroeconomic Aggregates for England"

url: https://faculty.econ.ucdavis.edu/faculty/gclark/data.html

downloaded: 09/12/2024


file: feinstein88.xlsx

data: capital stock

reference: Feinstein (1988), "National statistics, 1760-1920"

note: digitization of table VIII, columns (4) and (7)


file: GDP-Efficiency 2023.xlsx

data: efficiency index

reference: Clark (2016), "Microbes and Markets"

note: shared by Gregory Clark in private correspondance on 02/18/2023


file: HowGrowthBeganFiguresTables.xlsx

data: labor productivity

reference: Steinsson (2024), "How Did Growth Begin? The Industrial Revolution and its Antecedents"

version obtained: 09/13/2024

file: hw19.xlsx

data: annual income and hours

reference: Humphries & Weisdorf (2019), "Unreal Wages"

note: digitization of tables A2 and A3


file: MalthusFigures.xlsx

data: wage and population

reference: Steinsson (2021), "Malthus and Pre-Industrial Stagnation"

url: https://eml.berkeley.edu/~jsteinsson/teaching.html

downloaded: 09/12/2024


file: Wages 2014.xlsx

data: wage

reference: Clark (2010), "Macroeconomic Aggregates for England"

url: https://faculty.econ.ucdavis.edu/faculty/gclark/data.html

downloaded: 09/12/2024

## Appendix B: List and Explanation of Code Files

Launchers:

- *main_estimation.R*: launches the estimation of every specification in the paper

    o user needs to set *path0*, the path of *bns_replication* folder

    o saves draws to *~\bns_replication\output\RData*. Draws are already save, user does not need to run *main_estimation.R* again

    o if user wants to rerun particular specifications, they can pick a subset of the possible values for *nt* in the *for* loop. Possible values for *nt* are listed after line 48 with appropriate explanations

- *main_post.R*:

    o user needs to set *path0*, the path of *bns_replication* folder

    o produces the tables and charts contained in the paper and saves them in *~\bns_replication\output\tables* and *~\bns_replication\output\charts*

Estimation files:

- *estimation.R*: contains the code for performing the estimation. Called by *main_estimation.R*

- *malthus_model.stan*: Stan model file. Called by *estimation.R*

Figure- or table-generating programs:

These files generate figures and tables. They require the appropriate results generated by *estimation.R* (called *~\bns_replication\output\RData\.RDataZZZ*):

- *raw_data_charts.R:* figures based on raw data. Only requires transformed data file

    o figures 3, 4, 8, 9, 12, A1

- *charts_simple.R*: figures based on a single specification

    o figures 1, 2, 5, 6, 7, 11, 15, 16, 17, A4, A9, A10

- *charts_comparison.R:* figures that compare results across specifications

    o figures 10, 13, A2, A5, A6, A7, A8

- *tables.R*: all dynamic tables

    o tables 1, 4, 5, A1 (tables 2, 3 are hard coded)

- *irf_figures.R*: figures based on impulse response function

    o figures 14, H1 to H4

- *clark_experiments.R*: figures for appendix G

    o figures G1 to G6

- *densities.R*: density figure A3

- *to_excel.R*: Excel file with main posterior times series

- *text_numbers.R*: computes numbers cited in text


Support files:

Called by *main_estimation.R* and/or *main_post.R*:

- *libraries.R*: package loading (packages are automatically installed if missing)

- *chart_theme.R*: templates for charts

- *chart_functions.R*: defines functions which are used to produce charts

- *dataset.R*: assembles dataset used in the estimation

    o uses raw data files contained in ~\bns_replication\data\raw

    o saves dataset in ~\bns_replication\data\transformed\malthus_data.csv