Revolutionary Transition: Inheritance Change and Fertility Decline

Victor Gay* Paula E. Gobbi† Marc Goñi‡

Version 1.0 (June 2025)

Replication Package Readme File

1 Overview **1**

2 Data Availability **1**
2.1 Henry Database . . . 1
2.2 Geni Database . . . 2
2.3 Customs and Inheritance Data . . . 2
2.4 Other data . . . 2

3 Instructions to Replicators **4**
3.1 Restricted Data Access . . . 4
3.2 Reproduction Workflow . . . 5
3.3 Maps Reproduction Files . . . 9

4 Raw Data Files Sources and Preparation **12**
4.1 Henri Database (`1_1_henri`) . . . 12
4.2 Geni Database (`1_2_geni`) . . . 12
4.3 France (`1_3_france`) . . . 14
4.4 Customs (`1_4_customs`) . . . 15
4.5 Bailliages (`1_5_bailliages`) . . . 16
4.6 Wheat Prices (`1_6_wheat_prices`) . . . 17

*Department of Social and Behavioral Sciences, Toulouse School of Economics, University of Toulouse Capitole. Email: victor.gay@tse-fr.eu.
†Université Libre de Bruxelles (ECARES) and CEPR. Email: paula.eugenia.gobbi@ulb.be.
‡University of Bergen and CEPR. Email: marc.goni@uib.no.

4.7 Population (1_7_population) . . . 18
4.8 Administrative Centers (1_8_admin_centers) . . . 18
4.9 Political Societies (1_9_political_societies) . . . 19
4.10 Rebellions in 1779–89 (1_10_rebellions) . . . 19
4.11 Cassini roads (1_11_roads) . . . 20
4.12 Horse-Post Network (1_12_horse_post) . . . 20
4.13 Clerical Peril (1_13_clerical_peril) . . . 21
4.14 Soil texture (1_14_soil_texture) . . . 21
4.15 Terrain Ruggedness (1_15_ruggedness) . . . 22
4.16 Caloric Suitability (1_16_caloric) . . . 22
4.17 Easter Dates (1_17_easter_dates) . . . 22
4.18 Fertility Trends (1_18_fertility_trends) . . . 23
4.19 Familinx (1_19_familinx) . . . 23
4.20 Cadastre (1_20_cadastre) . . . 23
4.21 Agricultural census of 1852 (1_21_agricensus) . . . 24

5 List of Tables and Figures **24**
5.1 List of Main Text Tables and Figures . . . 25
5.2 List of Appendix Tables and Figures . . . 25

6 Technical Requirements **28**
6.1 Software . . . 28
6.2 Packages and Computing Time . . . 28

7 Licenses **30**

1 Overview

This replication package contains all the codes and information needed to reproduce the results in Gay et al. (2025) and its online appendix. It also contains the code and instructions to assemble the final datasets for the analysis starting from their original sources. We provide instructions on how to access the Henry (Séguy 2001) and the Geni database, as well as all the `Stata` do files, `R` codes, and `GIS` files needed to combine them with self-collected data on customs and inheritance (Gay et al. 2024b; Gay et al. 2024a)—which we provide in the replication package—and data from various other secondary datasets. Scripts in `2_scripts/2_1_data` run all the necessary code to generate the final datasets analyzed in the paper. Scripts in `2_scripts/2_2_analysis` replicate the figures and tables in the paper and the online appendix. We provide the shapefiles of the maps in the paper and in the online appendix, as well as instructions on how to reproduce them. The replicator should expect the code to run for about 5 hours in total.

The remaining of the `readme` file is organized as follows. Section 2 describes the provenance of each dataset used in the analysis. Section 3 provides guidance for replicators: how to access the restricted data used in the article (Section 3.1); the workflow for running the replication scripts (Section 3.2); and the files required to generate the article's maps (Section 3.3.2). Section 4 then outlines the sources of the raw data used to build the analytical datasets. Section 5 lists all tables and figures in the article, along with the corresponding datasets and scripts needed to reproduce them. Finally, Section 6 specifies the technical requirements for running the replication materials (software, packages, computing time, and operating system), and Section 7 details the licensing terms for the code and replication material.

2 Data Availability

This replication package uses self-collected data on inheritance customs and secondary data from both publicly available and restricted-access sources. Below, we provide a brief overview of the availability of the datasets used for the analysis. Detailed access instructions are available in Section 3.1. The location of each dataset in this replication package is available in Section 4.

2.1 Henry Database

The dataset referred to as the *Henry database* in this article corresponds to the nominative component of the *Enquête Louis Henry* (Fleury and Henry 1958; Séguy et al. 1999). This survey data was distributed by INED on CD-ROM, bundled with a dedicated publication (Séguy 2001). It is therefore proprietary data and not publicly available.

2.2 Geni Database

The dataset referred to as the *Geni database* in this article corresponds to online genealogies in geni.com, a MyHeritage Company. The dataset was obtained through a non-exclusive, non-transferrable, recoverable license between MyHeritage, Ltd. and Paula Gobbi, to access and use data from public profiles on geni.com.

2.3 Customs and Inheritance Data

The customs and inheritance data used in this article are based on shapefiles produced by the authors and made publicly available through the Harvard Dataverse (Gay et al. 2023a; Gay et al. 2023b). For detailed information on the construction of these datasets, see Gay et al. (2024b) and Gay et al. (2024a).

2.4 Other data

The analyses in this article uses other publicly available datasets and data sources:

- **Shapefile of contemporary French municipalities** The GEOFLA 2011 dataset (IGN 2011) is publicly available and can be downloaded from https://geoservices.ign.fr/geofla.
- **Wheat prices in Ancien Régime France** Wheat prices data are based on Ridolfi (2019) and are available upon request from Leonardo Ridolfi. We gratefully acknowledge his generosity in sharing the raw price series with us.
- **Population in 1793** Population data come from the dataset titled *État géolocalisé des paroisses et autres lieux habités de la Carte Générale de la France (1756–1789) devenus chefs-lieux de communes entre 1789 et 1999, accompagné du nombre d'habitants de ces communes relevé lors des 33 recensements nationaux réalisés entre l'an III et 1999 (based on cassini.ehess.fr)* (Cristofoli et al. 2021). This dataset is publicly available and can be downloaded from the Didomena repository at https://didomena.ehess.fr/concern/data_sets/6395wb092.
- **Administrative centers** The dataset of administrative centers is based on the list of municipalities that served as administrative centers prior to the French Revolution, as documented in Nordman et al. (1989, pp. 74–80).
- **Political societies** The dataset of political societies is based on the list of municipalities that hosted a political society in 1789–93, as documented in Boutier et al. (1992, pp. 77–101).

- **Rebellions** The dataset on rebellions against the state from 1779 to 1789 is based on the Jean Nicolas Survey (Nicolas 2002). The corresponding database is described and publicly distributed by Gay (2025).

- **Cassini roads** The shapefile of Cassini roads is based on Perret et al. (2015a). This dataset is publicly available and can be downloaded from the Harvard Dataverse at https://doi.org/10.7910/DVN/28674 (Perret et al. 2015b).

- **Horse-post network** The dataset of horse-post relays in 1790 is described and publicly distributed by Albertus and Gay (2025).

- **Clerical peril** The dataset of the share of clergy members who refused the oath of loyalty to the state in 1791 is from the shapefile of Squicciarini (2020b) and based on Tackett (1986). This data is publicly available and can be downloaded on the ICPSR repository at https://doi.org/10.3886/E119862V1 (Squicciarini 2020a).

- **Soil texture** The soil texture shapefile is from the soil geographical database of France (INRA 1998). This dataset is publicly available and can be downloaded from the Harvard Dataverse at https://doi.org/10.15454/BPN57S.

- **Terrain ruggedness** The terrain ruggedness index data is from Nunn and Puga (2012) based on EROS (2018)'s GTOPO30 dataset. This dataset is publicly available and can be downloaded from Diego Pugga's webpage at https://diegopuga.org/data/rugged (accessed in June 2025).

- **Caloric suitability** The post-1500 average caloric suitability index data is from Galor and Özak (2016) based on FAO (2012)'s GAEZ dataset. This dataset is publicly available and can be downloaded from Ömer ÖZark's webpage at https://ozak.github.io/Caloric-Suitability-Index (accessed in June 2025).

- **Easter dates** Easter dates data are publicly available and can be accessed at https://tlarsen2.tripod.com/thomaslarsen/easterdates.html (accessed in June 2025).

- **Fertlity trends** Fertility at the end of the first phase of the demographic transition is computed using publicly available data from SGF (1861)'s vital statistics (*Mouvement de la population*) for 1855, that can be accessed at https://www.insee.fr/fr/statistiques/2659665?sommaire=2591397#consulter (accessed in June 2025). Fertility at the end of the demographic transition is computed using publicly available data from from the 1954 *Enquête famille* (INSEE 1954), that can be requested at https://data.progedo.fr/studies/doi/10.13144/lil-0283?tab=access. Country-level fertility trends based on the crude birth rate and I_g index are from Chesnais (1992) and Weir (1994), respectively.

- **Familinx** The familinx data is a sub-sample of the Geni database scrapped by Kaplanis et al. (2018). At the time our project started, it could be downloaded from `https://familinx.org` (accessed in February 2017).

- **Cadastre** The cadastre data are publicly available from various départemental archives. See Appendix F for more details.

- **Arrondissements of 1852** The nomenclatures and shapefiles of arrondissements of 1852 are based on Gay (2020)'s dataset, which is publicly available and can be accessed at `https://doi.org/10.7910/DVN/98ERPB`.

- **France's waterways** The shapefile of France's waterways is from SANDRE (2017). It is publicly available and can be accessed at `https://www.sandre.eaufrance.fr/atlas/atlas/api/records/7381de46-42f7-42df-9abe-0ecd4b946034` (accessed in June 2025).

- **Agricultural census of 1852** The data of the 1852 agricultural census was produced by the French Ministry of agriculture (1858, 1860) and made available by Marin and Marraud (2011). This dataset is publicly available and can be accessed at `https://doi.org/10.4000/acrh.3696`.

3 Instructions to Replicators

3.1 Restricted Data Access

This section describes how to access the restricted data used in this article, which cannot be provided as part of this replication package. However, to preserve its folder structure on the JPE Dataverse, these files are substituted with empty `blank.txt` files when the deletion would leave the folder empty.

3.1.1 Henry Database

Accessing the Henry database requires first purchasing Séguy (2001).[1] Then, extract the following SAS datasets from the CD-ROM: `enfants.sas7bdat` and `parents.sas7bdat`. Finally, place these two files into the folder `1_raw_data/1_1_henri`.

3.1.2 Geni Database

Accessing the Geni database requires to request the following files to MyHeritage, Ltd:

[1] This book can be purchased via INED's website at `https://www.ined.fr/fr/publications/editions/etudes-enquetes-historiques/la-population-de-la-france-de-1670-a-1829/` (accessed in June 2025).

- `geni_profiles.csv` (20.2 GB, created on April 4, 2022, at 11:39:31),
- `geni_unions.csv` (7.2 GB, created on April 4, 2022, at 11:55:59),
- `geni_union_details.csv` (2.6 GB, created on April 4, 2022, at 11:54:34).

The data request should include a description of the research project. Paula Gobbi sent her ERC (B1) project together with some preliminary results using Familinx data and details of which work packages required the usage of the Geni database. Once the data is obtained, place the three files listed above into the folder `1_raw_data/1_2_geni`.

Once the access to Geni is approved, the following files should be asked from the authors:

- `fr-clean-geoloc.dta`
- `20km-from-henry.dta`
- `GENI_SAMPLE` (shapefile)

These files are described in detail in Section 4. Place `fr-clean-geoloc.dta` into the folder `1_raw_data/1_2_geni/auxiliary_files`, `20km-from-henry.dta` into the folder `1_raw_data/1_18_fertility_trends`, and the `GENI_SAMPLE` folder into the `1_raw_data/1_2_geni` folder.

In addition, since geni.com comprises crowdsourced genealogies which are continuously updated, the data accessed by the replicator may not be identical to or may have a different structure than the data obtained by the authors at the time of this article. For perfect reproduceability, the authors will provide, together with the codes used to extract the data from these raw files, the file `fr-clean`, which contains the geni sub-sample used in this article. Please place `fr-clean` into the folder `3_outputs/3_1_datasets`.

3.2 Reproduction Workflow

This section outlines the steps required to run the scripts and reproduce the results in this article. All scripts are located in the `2_scripts` folder. Replicators should first execute the scripts in the `2_1_data` folder, which generate analyses datasets, and then proceed to the scripts in the `2_2_analyses` folder, which produce the analytical outputs. Scripts should be run sequentially. For setting file paths, replicators just need to open `do` files from the directory where they are placed. Instructions regarding required software packages are provided in Section 6. If ran from a Dropbox or a OneDrive folder, we recommend pausing syncronization to avoid `r(608)` errors (i.e., directories marked as read-only). We provide additional script-specific details below.

3.2.1 Data Scripts

These scripts are located in the `2_1_data` folder and generate analyses datasets.

Henry database (`00_henry-data-prep.do`) This Stata `do` file combines the Henry and inheritance data to generate the final analysis dataset, `final-henry.dta`. Before running this script, gain access to the Henry nominative database (see Section 3.1). The `do` file structure is documented in the preamble.

Geni sample (`01_geni_data_to_sample`) This folder contains `R` scripts that extract the sub-sample used in this article—`fr-clean.csv`—from the full database provided by `geni.com`. We are deeply grateful to Matthew Curtis for writing these scripts and for his invaluable assistance with the Geni database. Before running the scripts in the order listed below, users must first obtain access to the raw Geni data files (see Section 3.1). Then, to ensure the `R` scripts run without read errors, execute the following terminal commands to convert `geni_profiles.csv` and `geni_union_details.csv` into `geni_profiles_patched.csv` and `geni_union_details_patched.csv`, respectively.

```
awk 'c>=70␣&&␣NR>1␣&&␣/^[0-9]{1,},/␣{printf␣\"\\n\";}␣c>=70␣&&␣NR>1␣&&␣/^[0-9]{1,},/␣{
    c=0}{␣c=␣c+gsub(\",\",␣\",\")}{printf␣\"%s\",$0}' ./Data/Raw/geni_profiles.csv| tr
     '\\r' '␣'> ./Data/Patched/geni_profiles_patched.csv
```

```
awk 'c>=28␣&&␣NR>1␣&&␣/^[0-9]{1,},/␣{printf␣\"\\n\";}␣c>=28␣&&␣NR>1␣&&␣/^[0-9]{1,},/␣{
    c=0}{␣c=␣c+gsub(\",\",␣\",\")}{printf␣\"%s\",$0}' ./Data/Raw/geni_union_details.
    csv| tr '\\r' '␣'> ./Data/Patched/geni_union_details_patched.csv
```

- `fr-cities-manual-fix.R` This `R` script retrieves the geolocations of French cities in 1700, as listed in Buringh (2020), using the GeoNames geographical database. To run the script, users must first create a GeoNames username, available at `https://www.geonames.org`, and enter it at the beginning of the script. The output—the `fr-cities-manual-fix.csv` file—is saved in the `auxiliary_files` folder and will be used by `7-xwalk-locations.R`.
- `0-chunk-profiles.R` This `R` script cuts the `geni_profiles_patched.csv` data into 16 `PARQUET` type files for efficiency purposes.
- `0-chunk-union.R` This `R` script cuts the `geni_unions.csv` data into 17 `PARQUET` type files.
- `0-chunk-union-details.R` This `R` script cuts the `geni_union_details_patched.csv` data into 11 `PARQUET` type files.
- `1-extract-profile-bd.R` This `R` script takes the raw Geni profiles file `geni_profiles.csv` and creates the `fr-pids-bd.csv` file, which contains a list of `profile_ids` who were born or died in France.

- `2-extract-profile-ids-m.R` This `R` script takes the raw Geni `geni_union_details.csv` and `geni_unions.csv` files and and creates the `fr-pids-m.csv` file, a list of `profile_ids` who married in France.

- `3-merge-in-unions.R` This `R` script takes the `fr-pids-m.csv` file file and iteratively merges in union identifiers and profile identifiers to build up the family tree of each `profile_id` either born, married, or deceased in France. It iterates six times and creates a final `fr-union6.csv` file.

- `4-family-ids.R` This `R` script provides the links between `profile_id` and marriage-type unions (up to 3), together with the location and date of the marriage. The output is the `fr-family-ids.csv` file.

- `5-relink-profiles.R` This `R` script takes the `fr-family-ids.csv` file and merges in the information from the `geni_profiles.csv` file, and uses baptisms and burials to infer births and deaths when necessary. The output is the `fr-raw.csv` file.

- `6-georef.R` This `R` script assigns latitude and longitude to each location of birth, marriage, and death using Geonames data. The output files are `fr-geocoded-v3.csv` and `fr-locations-original.csv`.

- `7-xwalk-locations.R` This `R` script takes latitudes and longitudes and assigns inheritance rules and distance to city. The output is the `fr-locations-customs.csv` file.

- `8-finish-fert.R` This `R` script assigns geographical information to individuals, creates variables, and restricts to the final sample we use. The final output file is the `fr-clean.csv` file.

Note: since geni.com comprises crowdsourced genealogies which are continuously updated, the data accessed by the replicator may not be identical to or may have a different structure than the data obtained by the authors at the time of this article. This may cause the previous `R` scripts to work imperfectly.

Geni controls (`02_geni-inheritance-and-controls.do`) This Stata `do` file combines the Geni sample and data on inheritance customs and controls to generate the `geni_controls.dta` file. After gaining access to the Geni data, this `do` file can be ran after the `R` scripts in `01_geni_data_to_sample`, or by placing the author-provided `fr-clean` file into the folder `3_outputs/3_1_datasets`. The `do` file structure is documented in the preamble.

Geni database (`03_geni-final-data-prep.do`) This Stata `do` file combines the Geni sample, data on inheritance customs and controls, and generates variables for analysis, the `final-geni.dta` file. The `do` file structure is documented in the preamble.

Familinx database (`04_familinx-data-prep.do`) This Stata `do` file combines the Familinx data files and generates the familinx annual fertility series used in Appendix Figure B2. The `do` file structure is documented in the preamble. Before running it, use "Text File Splitter"[2] to break `profiles-anon.txt` into 151 parts; name them `part1.txt`, `part2.txt`, ..., and `part151.txt`; and place them in the folder `2_scripts/2_0_tempfiles`.

Cadastre database (`05_cadastre-data-prep.do`) This Stata `do` file creates the cadastre database used in Appendix F. The `do` file structure is documented in the preamble.

Agricultural census database (`06_agricensus-data-prep.do`) This Stata `do` file creates the 1852 agricultural census database used in Appendix F. The `do` file structure is documented in the preamble.

3.2.2 Analyses Scripts

These scripts are located in the `2_2_analyses` folder and produce analytical outputs of this article.

Main outputs based on the Henry database (`01_results-paper-henry.do`) This Stata `do` files contains the script to reproduce the figures and tables in the main text using the Henry database. Specifically, it produces Figures 1 and 4, and Tables 1, 2, 3, and 5.

Main outputs based on the Geni database (`02_results-paper-geni.do`) This Stata `do` files contains the script to reproduce the figures and tables in the main text using the Geni database. Specifically, it produces Figure 5 and Table 4.

Appendix outputs based on the Henry database (`03_results-appendix-henry.do`) This Stata `do` files contains the script to reproduce the figures and tables in the appendix using the Henry database. Specifically, it produces Tables A1, A3, A4, A5, A7, A8, A10, C1, D1, D2, D3, D4, D5, D6, D7, D8, D9, H1, and H2, and Figures B5, B6, C1, D1, D2, D3, D4, and D5.

Appendix outputs based on the Geni database (`04_results-appendix-geni.do`) This Stata `do` files contains the script to reproduce the figures and tables in the appendix

[2] https://apps.apple.com/us/app/text-file-splitter/id831715956?mt=12&ign-mpt=uo%3D5

using the Geni database. Specifically, it produces Tables A9, E1, and E2, and Figures E1, E2, E3, and E4.

Appendix outputs based on the cadastre database (`05_results-appendix-cadastre.do`) This Stata `do` files contains the script to reproduce the figures and tables in Appendix F using the cadastre database. Specifically, it produces Tables F2 and F3, and Figures F3 and F4.

Appendix outputs based on the agricultural census database (`06_results-appendix-agricensus.do`) This Stata `do` files contains the script to reproduce the figures and tables in Appendix F using the 1852 agricultural census database. Specifically, it produces Tables F4 and F5, and Figure F9.

Appendix model output (`07_model-appendix.nb`) This Mathematica `nb` file contains the script to reproduce the model simulation in Appendix G. Specifically, it produces the content of Figure G1—a `png` file that is then called by a `tex` file.

3.3 Maps Reproduction Files

We detail the files required to generate the maps presented in the main text and appendix using QGIS. The corresponding PNG map files are located in the `3_outputs` folder. All maps use a common base shapefile representing the territory of France as of 1789, available as `FRANCE_1789` in the `1_raw_data/1_3_france` folder. This shapefile is based on the data from Gay et al. (2023a).

3.3.1 Main Text Maps

Figure 2a. Customs in Ancien Régime France. `figure2a.png`. Import the `CUSTOMS` layer located in the `1_raw_data/1_4_customs` folder to display customary region boundaries, then the `WRITTEN_LAW` layer located in the same folder to overlay the types of customary regions between customary law, written law, and mixed regions.

Figure 2b. Inheritance in Ancien Régime France. `figure2b.png`. Import and overlay the following inheritance layers located in the `1_raw_data/1_4_customs` folder: `INHERITANCE_TREATED`, `INHERITANCE_WOMEN`, `INHERITANCE_PARTIBLE`, `INHERITANCE_ IMPARTIBLE`, and `INHERITANCE`.

Figure 3a. Henry locations. DD setting. `figure3a.png`. Import the `INHERITANCE_TREATED` layer located in the `1_raw_data/1_4_customs` folder, then overlay the `HENRY_SAMPLE` layer located in the `1_raw_data/1_1_henri` folder with the locations of Henry municipalities and their sample sizes.

Figure 3b. Geni locations. RD-DD setting. `figure3b.png`. Import the `INHERITANCE_TREATED` layer located in the `1_raw_data/1_4_customs` folder, then overlay the `GENI_SAMPLE` and `GENI_SAMPLE_30` layers located in the `1_raw_data/1_2_geni` folder with the locations of Geni individuals along with their position relative to the inheritance treatment frontier.

3.3.2 Appendix Maps

Figure B1. Judicial districts in Ancien Régime France. `figureB1.png`. Import the `BAILLIAGES` layer located in the `1_raw_data/1_5_bailliages` folder.

Figure B3.a. Written-law and customary-law areas (Klimrath). `figureB3a.png`. Import the `WRITTEN_LAW_KLIMRATH` layer located in the `1_raw_data/1_4_customs` folder.

Figure B3.b. Written-law and customary-law areas (Gay, Gobbi, Goñi). `figureB3b.png`. Import the `WRITTEN_LAW` layer located in the `1_raw_data/1_4_customs` folder.

Figure B7. Inheritance systems in four administrative centers. `figureB7.png`. Import the `INHERITANCE` layer located in the `1_raw_data/1_4_customs` folder and zoom-in on the appropriate locations.

Figure F5.a. Terrain ruggedness. `figureF5a.png`. Import the `tri_france` raster layer located in the `1_raw_data/1_15_ruggedness` folder.

Figure F5.b. Soil texture. `figureF5b.png`. Import the `SOIL_TEXTURE` layer located in the `1_raw_data/1_14_soil_texture` folder.

Figure F7.a. Terrain ruggedness. `figureF7a.png`. Import the `RUGGEDNESS_ ARRONDISSEMENTS` layer located in the `1_raw_data/1_15_ruggedness` folder.

Figure F7.b. Share of sandy soils. `figureF7b.png`. Import the `SOIL_TEXTURE_ ARRONDISSEMENTS` layer located in the `1_raw_data/1_14_soil_texture` folder.

Figure F8.a. Arable hectares per landowner. `figureF8a.png`. Import the `FARM_SIZE` layer located in the `1_raw_data/1_21_agricensus` folder and use the `size1` variable.

Figure F8.b. Cultivated hectares per landowner. `figureF8b.png`. Import the `FARM_SIZE` layer located in the `1_raw_data/1_21_agricensus` folder and use the `size2` variable.

Figure H3. Raster map of wheat prices in the 1780s. `figureH3.png`. Import the `WHEAT_PRICES_IDW_1780` raster layer located in the `1_raw_data/1_6_wheat_prices` folder and clip it to the extant of 1789 France using the `FRANCE_1789` layer located in the `1_raw_data/1_3_france` folder through the `Clip raster by mask layer` command, then overlay the `WHEAT_PRICE_LOCATIONS_1780` layer located in the `1_raw_data/1_6_wheat_prices` folder.

Figure H4.a. Évêchés. `figureH4a.png`. Import the `EVECHES` layer located in the `1_raw_data/1_8_admin_centers` folder.

Figure H4.b. Bailliages. `figureH4b.png`. Import the `BAILLIAGES` layer located in the `1_raw_data/1_8_admin_centers` folder.

Figure H4.c. Recettes des finances. `figureH4c.png`. Import the `RECETTES` layer located in the `1_raw_data/1_8_admin_centers` folder.

Figure H4.d. Subdélégations. `figureH4d.png`. Import the `SUBDELEGATIONS` layer located in the `1_raw_data/1_8_admin_centers` folder.

Figure H5.a. Political societies (1789–93). `figureH5a.png`. Import the `POLITICAL_SOCIETIES` layer located in the `1_raw_data/1_9_political_societies` folder.

Figure H5.b. Rebellions (1779–89). `figureH5b.png`. Import the `REBELLIONS` layer located in the `1_raw_data/1_10_rebellions` folder.

Figure H6.a. Cassini roads (late 1700s). `figureH6a.png`. Import the `france_cassini_roads` layer located in the `1_raw_data/1_11_roads` folder.

Figure H6.b. Horse-post network (1790). `figureH6b.png`. Import the `RELAYS_1790` and `SECTIONS_1790` layers located in the `1_raw_data/1_12_horse_post` folder.

Figure H7.a. Caloric suitability. `figureH7a.png`. Import the `post1500AverageCalories_france` raster layer located in the `1_raw_data/1_16_caloric` folder.

Figure H7.b. Terrain ruggedness. `figureH7b.png`. Import the `tri_france` raster layer located in the `1_raw_data/1_15_ruggedness` folder.

4 Raw Data Files Sources and Preparation

This section outlines the sources of the raw data files used to construct the analytical datasets for this article. With the exception of the Henri and Geni data files, most raw data files processed by the scripts are first manually geolocated in QGIS at the level of municipalities' *chef-lieux* using the `COMMUNES` shapefile from IGN (2011), which we convert into the `CHEF_LIEUX` point layer (see Section 4.3). The scripts then retain only those locations that fall within the boundaries of France as of 1789, as defined by the `FRANCE_1789` shapefile from Gay et al. (2023a). Detailed procedures for processing each raw data file are provided below.[3]

4.1 Henri Database (`1_1_henri`)

4.1.1 Shapefiles

`HENRY_SAMPLE` This folder contains the `HENRY_SAMPLE` shapefile, which provides the distribution of Henry municipalities along with their sample sizes.

4.1.2 SAS and CSV files

`enfants.sas7bdat` This file contains the children's data of the nominative part of the Henry database.

`parents.sas7bdat` This file contains the parents' data of the nominative part of the Henry database.

`henri_villages.csv` This file contains the list of municipalities in the Henry database along with their INSEE identifiers.

`henry_controls.csv` This file provides the control and treatment variables for each municipality in the Henry sample. These variables are constructed using sources and methods similar to those employed for the Geni database, but are for the 40 henry villages and are specific to the DD design that we use in this analysis.

4.2 Geni Database (`1_2_geni`)

4.2.1 Geni chunked

`1_2_geni_chunked` This folder contains 3 subfolders, `geni_profiles`, `geni_union_details`, and `geni_union` with chunks of the Geni raw data in PARQUET type files ensuring effi-

[3]The QGIS project file `raw_data.qgz`, located at the root of the `1_raw_data` folder, loads the main layers of this project.

ciency in the construction of the Geni data sample.

4.2.2 Shapefiles

`GENI_SAMPLE` This folder contains the `GENI_SAMPLE` shapefile, which provides the distribution of individuals in the Geni sample. The `GENI_SAMPLE_30` shapefile restricts the Geni sample to individuals that are within 30 kilometers of the inheritance treatment frontier.

4.2.3 Auxiliary files (`auxiliary_files`)

This folder contains auxiliary files used for the preparation of the Geni database.

`brette_bailliages.xlsx` This file contains the links between bailliages and custom's names. This file was produced by the authors. For details see Gay et al. (2024a).

`buringh.xlsx` This file contains the Buringh (2020)'s database. After downloading it, open the `European urban population, 700 - 2000.txt` file in excel, select the data, convert text to columns using "@" as the character separating each field, and save it in the folder `auxiliary_files` as `buring.xlsx`.

`FR.txt` This file contains place names in France from GeoNames. It was extracted from the `FR.zip` downloaded from `https://download.geonames.org/export/dump/` on 12.12.2022.

`fr-cities-manual-fix.csv` This file is produced in the script `fr-cities-manual-fix.R` in `01_geni_data_to_sample` and contains the 324 cities available in Buringh (2020) in France in 1700, geocoded using the geonames database.

`fr-fix-bailliages.txt` This file provides manual corrections to non-numeric bailliages identifiers in `brette_bailliages.xlsx` to allow for successful merging with the shapefile in `z-French-bailliages` that uses numeric identifiers.

`list-customs_FR.txt` This file was produced by the authors and provides detailed information on inheritance laws for each custom.

`z-French-bailliages` This folder contains the shapefile for the bailliages in France in 1789, excluding Corsica, and whether each of these bailliages was under written law. The files in the folder were produced by the authors based on the `BAILLIAGES_1789_BRETTE` shapefile of Gay et al. (2023a), which draws on Brette (1904). For details see Gay et al. (2024a) and Gay et al. (2024b).

`fr-clean-geoloc.dta` This file contains the set of final Geni profiles along with their manually-corrected geolocations, i.e., their INSEE municipality identifier. More specifically, we start with the output of the `01_geni_data_to_sample` scripts (`fr-clean.csv`). We then extract Geni profiles born between 1700 and 1803 and located in France, and keep profile identifiers, their original location strings, and their corresponding automatically georeferenced locations obtained through Geonames. Next, we project these georeferenced points onto the contemporary shapefile of France's municipalities (`COMMUNES`), attribute the corresponding municipality information (INSEE identifier and name) through the `Join attributes by location` command on QGIS and compare it manually to the original location string. This procedure reveals that the automatic georeferencing has an accuracy rate close to 70 percent. We correct the remaining inaccuracies manually—the essence of this procedure is explained in the main text (Footnote 23). Finally, we save this file in `dta` format and use it at the end of the `02_geni-inheritance-and-controls.do` file.

4.3 France (`1_3_france`)

4.3.1 Shapefiles

`FRANCE_1789` This folder contains the `FRANCE_1789` shapefile of the territory of France as of 1789, excluding Corsica. It is based on the `France_1789_BRETTE` shapefile of Gay et al. (2023a), which draws on Brette (1904). See Gay et al. (2024b) for more details.

`GEOFLA_2011` This folder contains the `COMMUNES` shapefile from IGN (2011).[4] From this file, we construct the `CHEF_LIEUX` point layer. This is done by multiplying the `X_CHF_LIEU` and `Y_CHF_LIEU` fields in the attribute table by 100, exporting the table as a CSV file (`chef_lieux.csv`), and then importing it into QGIS as a delimited text layer using the RGF93 CRS.

4.3.2 CSV files

`coordinates_rfg93.csv` This file contains the latitude and longitude of municipalities in the `CHEF_LIEUX` point layer obtained by exporting the `CHEF_LIEUX` shapefile using the RGF 93 CRS and requesting `AS_XY` as layer options. It is also provided in `dta` format.

`coordinates_wgs84.csv` This file contains the latitude and longitude of municipalities in the `CHEF_LIEUX` point layer obtained by exporting the `CHEF_LIEUX` shapefile using the WGS 84 CRS and requesting `AS_XY` as layer options.

[4] Specifically, the path to the `COMMUNES` shapefile within the publicly available dataset is: `GEOFLA_1-1_COMMUNES_SHP_LAMB93_FXX_2011-11-14/GEOFLA/1_DONNEES_LIVRAISON_2021-03-00087/ GEOFLA_1-1_SHP_LAMB93_FR-ED111/COMMUNES`.

4.4 Customs (1_4_customs)

4.4.1 Shapefiles

CUSTOMS This folder contains the CUSTOMS shapefile of customary regions.

WRITTEN_LAW This folder contains the WRITTEN_LAW shapefile of types of customary regions, between customary law, written law, and mixed areas.

WRITTEN_LAW_KLIMRATH This folder contains the WRITTEN_LAW_KLIMRATH shapefile of types of customary regions, between customary law and written law, based on Klimrath (1837).

INHERITANCE This folder contains the INHERITANCE shapefile of inheritance customs. It details the distribution of inheritance customs between partible (option, strict) and impartible (primogeniture, ultimogeniture, unigeniture) areas, and between areas where women were excluded versus included in inheritance.

INHERITANCE_IMPARTIBLE This folder contains the INHERITANCE_IMPARTIBLE shapefile that splits France into primogeniture, ultimogeniture, and unigeniture impartible inheritance areas. It is based on the INHERITANCE shapefile of inheritance customs.

INHERITANCE_PARTIBLE This folder contains the INHERITANCE_PARTIBLE shapefile that splits France into strict and option partible inheritance areas. It is based on the INHERITANCE shapefile of inheritance customs.

INHERITANCE_TREATED This folder contains the INHERITANCE_TREATED shapefile that splits France into treated (impartible and/or women excluded) and untreated (partible and women included) areas by the 1793 inheritance reforms. It is based on the INHERITANCE shapefile of inheritance customs.

INHERITANCE_TREATED_BORDER This folder contains the INHERITANCE_TREATED_BORDER_50 and INHERITANCE_TREATED_BORDER_100 shapefiles. They represent the interior border between treated and untreated areas, divided into 50- and 100-kilometer segments, respectively. The borders are represented as point layers, which allows for the calculation of the shortest distance from each municipality to the treatment border. These shapefiles are derived from the INHERITANCE_TREATED shapefile of treated areas.

More precisely, we perform the following operations: we transform the polygons in the INHERITANCE_TREATED shapefile into lines using the `Polygons to lines` command; split these lines by 50 and 100-kilometer segments using the `Split lines by maximum length` command; create a segment identifier through the field calculator using the `@id`

command; calculate each segment length (in kilometers) through the field calculator using the `$length/1000` command; transform lines into points distributed every 100 meters using the `Points along geometry` command; and finally manually erase all the points located on the territorial frontiers of France.

`INHERITANCE_WOMEN` This folder contains the `INHERITANCE_WOMEN` shapefile that splits France into areas where women were excluded versus included in inheritance. It is based on the `INHERITANCE` shapefile of inheritance customs.

4.4.2 CSV files

`affected.csv`, `fem_inherit.csv`, `partimpart.csv` These file map municipalities in the `CHEF_LIEUX` point layer to the `INHERITANCE` shapefile of inheritance customs through the `Join attributes by location` command. Specifically, we join to features in `CHEF_LIEUX` they `are within` by comparing to `INHERITANCE` (note: we create a spatial index for both layers before running this command). We do so for whether women are included in inheritance (`fem_inherit.csv`), for partible and impartible areas (`partimpart.csv`), and for treated and untreated areas (`affected.csv`).

`affected_distance.csv`, `affected_segment_50.csv`, `affected_segment_100.csv` These files provide the minimum distance from municipalities in the `CHEF_LIEUX` point layer to the treatment frontier in the `INHERITANCE_TREATED_BORDER` shapefiles, along with the identifier of the 50 or 100-kilometer segment associated with this nearest point. Specifically, we calculate this distance through the `Distance to nearest hub (points)` command from the `CHEF_LIEUX` source points layer to the `INHERITANCE_TREATED_BORDER_50` destination hubs layer in kilometers, keeping the `SEGMENT` identifier as the hub layer name attribute. We do the same for the 100-kilometer segments.

`affected_segment_50_length.csv`, `affected_segment_100_length.csv` These files provide the length of the segments in the `affected_segment_50.csv` and `affected_segment_100.csv` files.

4.5 Bailliages (`1_5_bailliages`)

4.5.1 Shapefiles

`BAILLIAGES` This folder contains the `BAILLIAGES` shapefile of bailliages in 1789, excluding Corsica. It is based on the `BAILLIAGES_1789_BRETTE` shapefile of Gay et al. (2023a), which draws on Brette (1904). See Gay et al. (2024b) for more details.

4.5.2 CSV files

`bailliages.csv` This file maps municipalities in the `CHEF_LIEUX` point layer to the `BAILLIAGES` shapefile of bailliages through the `Join attributes by location` command. Specifically, we join to features in `CHEF_LIEUX` they `are within` by comparing to `BAILLIAGES` (note: we create a spatial index for both layers before running this command).

4.6 Wheat Prices (`1_6_wheat_prices`)

4.6.1 Shapefiles

`WHEAT_PRICE_LOCATIONS` This folder contains shapefiles for each decade of the locations of decade-average wheat prices based on the `wheat_prices.csv` file.

`WHEAT_PRICES` This folder contains shapefiles for each decade of the spatially interpolated decade-average wheat prices by municipality. Specifically, we attribute an mean interpolated price to each municipality by using the `Zonal statistics` command with the `COMMUNE` layer as the input layer and the `WHEAT_PRICES_IDW_`*`YEAR`* layer as the interpolated raster layer.

`WHEAT_PRICES_IDW` This folder contains raster layers for each decade of the spatially interpolated decade-average wheat prices. Specifically, we interpolate prices using the `IDW interpolation` command with the `WHEAT_PRICE_LOCATIONS_`*`YEAR`* as the input layer, `price`*`year`* as the interpolation attribute, on the extent of the `FRANCE_1789` layer and with an output raster size of 300 rows and 304 columns such that each pixel is about the size of a municipality.

4.6.2 CSV files

`wheat_prices_ridolfi.csv` This file contains the raw wheat prices data from Ridolfi (2019).

`wheat_prices.csv` This file files contain the average wheat price per decade and municipality based on the `wheat_prices_ridolfi.csv`, along with municipalities' spatial coordinates. The files `wheat_prices_1700.csv` to `wheat_prices_1790.csv` report the corresponding information for each decade.

4.7 Population (1_7_population)

4.7.1 CSV files

`arrondissements_cl_1837.csv` The `arrondissements_cl_1837.csv` file contains the list of municipalities that were *arrondissement* chef-lieu per the *Statistique de la France* 1837 (pp. 7–10).

`cassini_base.dta` This file corresponds to the Cassini database based on the `lieux_cassini_devenus_communes.csv` file (Cristofoli et al. 2021).

`geofla_area.csv`, `ID_GEOFLA.csv` These files are based on the GEOFLA database (IGN 2011). They contain the area of each municipality polygon computed by dividing the `SUPERFICIE` variable by 100 and the correspondance between each polygon and its INSEE and IGN identifier.

`table_passage_annuelle_2021.csv` This file provides the correspondance between municipalities in 2021 geography to municipalities in past geographies, including 2011. It is publicly available from the INSEE website at https://www.insee.fr/fr/information/7671867 (accessed in June 2025). From there the `missing_insee_com1_cassini.dta` is generated, containing the municipalities that merged between 2011 and 2021.

4.8 Administrative Centers (1_8_admin_centers)

4.8.1 Shapefiles

`ADMIN_CENTERS` This folder contains shapefiles of the municipalities that hosted an administrative center before the Revolution based on the join of the `admin_centers.csv` with the `CHEF_LIEUX` shapefile using the `INSEE_COM` municipality identifier. Specifically, the folder contains the following shapefiles: `EVECHES` for Church administration (*évêché* capitals), `RECETTES` for tax collection (*recettes des finances* capitals), `BAILLIAGES` for judicial district seat (*bailliage* capitals), and `SUBDELEGATION` for territorial administration (*subdélégation* capitals). Note that we only keep administrative centers in France as of 1789 and exclude Corsica by clipping these shapefiles to the `FRANCE_1789` shapefile using the `Clip` command.

4.8.2 CSV files

`admin_centers.csv` The `admin_centers.csv` file provides the list of municipalities that hosted an administrative center before the Revolution based on Nordman et al. (1989, pp. 74–80).

*admin_centers*_distance.csv The *admin_centers*_distance.csv files provides the minimum distance between municipalities in the CHEF_LIEUX point layer and administrative centers in the ADMIN_CENTER shapefiles. Specifically, we calculate this distance through the Distance to nearest hub (points) command from the CHEF_LIEUX source points layer to the ADMIN_CENTER destination hubs layer in kilometers. We do so for each administrative center and generate center-specific csv files: eveches_distance.csv, recettes_distance.csv, bailliages_distance.csv, and subdelegations_distance.csv.

4.9 Political Societies (1_9_political_societies)

4.9.1 Shapefiles

POLITICAL_SOCIETIES This folder contains the POLITICAL_SOCIETIES shapefile of municipalities that hosted a political society in 1789–93 based on the join of the political_societies.csv with the CHEF_LIEUX shapefile using the INSEE_COM municipality identifier. Note that we only keep political societies in France as of 1789 and exclude Corsica by clipping this shapefile to the FRANCE_1789 shapefile using the Clip command.

4.9.2 CSV files

political_societies.csv The political_societies.csv file provides the list of municipalities that hosted a political society in 1789–93 based on Boutier et al. (1992, pp. 77–101).

political_societies_distance.csv The political_societies_distance.csv file provides the minimum distance between municipalities in the CHEF_LIEUX point layer and political societies in the POLITICAL_SOCIETIES shapefile. Specifically, we calculate this distance through the Distance to nearest hub (points) command from the CHEF_LIEUX source points layer to the POLITICAL_SOCIETIES destination hubs layer in kilometers.

4.10 Rebellions in 1779–89 (1_10_rebellions)

4.10.1 Shapefiles

REBELLIONS This folder contains the REBELLIONS shapefile of municipalities that experienced rebellions against the state in 1779–89. To create this file, we import the rebellions.csv file as a delimited text layer.

4.10.2 CSV and DTA files

`nicolas_events.dta` The `nicolas_events.dta` file is the raw dataset of the Jean Nicolas database from Gay (2025) based on the Jean Nicolas Survey (Nicolas 2002).

`rebellions.csv` The `rebellions.csv` file contains the number of rebellions against the state in 1779–89 by municipality along with spatial coordinates. More specifically, it is based on rebellions in the `nicolas_events.dta` file which primary type is `01 Rejection of state reform initiatives`, `02 Resistance to state taxation`, or `03 Resistance to the state judiciary, military, or police`.

`rebellions_distance.csv` The `rebellions_distance.csv` file provides the minimum distance between municipalities in the `CHEF_LIEUX` point layer and rebellions in the `REBELLIONS` shapefile. Specifically, we calculate this distance through the `Distance to nearest hub (points)` command from the `CHEF_LIEUX` source points layer to the `REBELLIONS` destination hubs layer in kilometers.

4.11 Cassini roads (`1_11_roads`)

4.11.1 Shapefiles

`CASSINI_ROADS` This folder contains the `france_cassini_roads` shapefile of Cassini roads based on Perret et al. (2015a) and available in Perret et al. (2015b).

4.11.2 CSV files

`roads_distance.csv` The `roads_distance.csv` file provides the minimum distance between municipalities in the `CHEF_LIEUX` point layer and Cassini roads in the `france_cassini_roads` shapefile. Specifically, we first transform the Cassini roads vectors into points distributed every 100 meters using the `Points along geometry` command, then calculate the distance through the `Distance to nearest hub (points)` command from the `CHEF_LIEUX` source points layer to the `france_cassini_roads_points` destination hubs layer in kilometers.

4.12 Horse-Post Network (`1_12_horse_post`)

4.12.1 Shapefiles

`HORSE_POST` This folder contains the `RELAYS_1790` shapefile of horse-post relays in 1790 and the `SECTIONS_1790` shapefile of horse-post roads in 1790 based on data from Albertus and Gay (2025).

4.12.2 CSV files

`horse_post_distance.csv` The `horse_post_distance.csv` file provides the minimum distance between municipalities in the `CHEF_LIEUX` point layer and horse-post relays in the `RELAYS_1790` shapefile. Specifically, we calculate this distance through the `Distance to nearest hub (points)` command from the `CHEF_LIEUX` source points layer to the `RELAYS_1790` destination hubs layer in kilometers.

4.13 Clerical Peril (`1_13_clerical_peril`)

4.13.1 Shapefiles

`CLERICAL_PERIL` This folder contains the shapefile of districts with the share of refractory clergy, who refused the oath of loyalty to the state in 1791. Specifically, we use Squicciarini (2020b) shapefile of 552 districts in 1791 and input data of the number of clergy members who took an oath in Spring 1791 directly from Tackett (1986). We input relevant département-level values when a specific district is missing.

4.13.2 CSV files

`clerical_peril.csv` This file maps municipalities in the `CHEF_LIEUX` point layer to the `CLERICAL_PERIL` shapefile of districts through the `Join attributes by location` command. Specifically, we join to features in `CHEF_LIEUX` they `are within` by comparing to `CLERICAL_PERIL` (note: we create a spatial index for both layers before running this command).

4.14 Soil texture (`1_14_soil_texture`)

4.14.1 Shapefiles

`SOIL_TEXTURE` This folder contains the `SOIL_TEXTURE` shapefile based on data from INRA (1998). It provides the dominant surface textural class across 318 soil mapping units along 5 categories. To construct this shapefile, we rely on the `text1` variable in the `stu` dataset provided by INRA (1998). In details, we match the `stu` dataset, which describes the dominant surface textural class at the level of 917 soil typological units (`stu`), to the `stuorg` dataset, which maps soil typological units into 318 soil mapping units (`smu`) together with the proportion of the area of soil mapping units covered by each soil typological unit (`pcarea`). We then calculate a relative surface textural class for each soil mapping unit using the `pcarea` variable as weights. Finally, we match these data to the `30169_L93` shapefile of the 318 soil mapping units. We then intersect this shapefile with the `COMMUNES` layer to create the `COMMUNE_TEXTURE_INTERSECTED` shapefile, then calculate area-weighted averages for each municipality polygon.

We use the same method with the ARRONDISSEMENTS_1852 layer to create the SOIL_TEXTURE_ARRONDISSEMENTS_INTERSECTED shapefile. The SOIL_TEXTURE_ARRONDISSEMENTS layer is the created based on the soil_texture_arrondissements.csv dataset (see below).

4.14.2 CSV files

soil_texture_arrondissements.csv This file contains soil texture data for at the level of 1852 arrondissements. It is based on the soil texture information contained in the SOIL_TEXTURE_ARRONDISSEMENTS_INTERSECTED shapefile.

4.15 Terrain Ruggedness (1_15_ruggedness)

4.15.1 Shapefiles

RUGGEDNESS This folder contains the tri_france raster layer, which corresponds to tri raster layer of Nunn and Puga (2012) clipped to the extent of France using the Clip raster by mask layer command. We then polygonize this raster using the Polygonize command, weigh each cell using the cell_area_france raster, rescale it in hundreds of meters, and intersect it with the COMMUNES layer to create the RUGGEDNESS_COMMUNE_INTERSECTED shapefile.

We use the same method with the ARRONDISSEMENTS_1852 layer to create the RUGGEDNESS_ARRONDISSEMENTS_INTERSECTED shapefile. The RUGGEDNESS_ARRONDISSEMENTS layer is the created based on the soil_texture_arrondissements.csv dataset (see below).

4.15.2 CSV files

ruggedness_arrondissements.csv This file contains terrain ruggedness data for at the level of 1852 arrondissements. It is based on the ruggedness information contained in the RUGGEDNESS_ARRONDISSEMENTS_INTERSECTED shapefile.

4.16 Caloric Suitability (1_16_caloric)

CALORIC_SUITABILITY This folder contains the post1500AverageCalories_france raster layer, which corresponds to the post1500AverageCalories raster layer of Galor and Özak (2016) clipped to the extent of France using the Clip raster by mask layer command. We then polygonize this raster using the Polygonize command and intersect it with the COMMUNES layer to create the post1500AverageCalories_COMMUNE shapefile.

4.17 Easter Dates (1_17_easter_dates)

Easter-dates-1700-1819.txt This file contains easter dates from 1700 to 1819.

4.18 Fertility Trends (1_18_fertility_trends)

`mouvement_population.dta` This file contains département-level data based on SGF (1861)'s vital statistics (*Mouvement de la population*) for 1855.

`fam1954.dta` This file corresponds to the 1954 *Enquête famille* (INSEE 1954).

`dep_treat_weight_1851.dta` This file contains the population weights of each département, equal to its population share relative to the entire population of its relevant pre-reform inheritance area per the 1851 census.

`reg_treat_weight_1954.dta` This file contains the population weights of each region available in the 1954 *Enquête famille*, equal to its population share relative to the entire population of its relevant pre-reform inheritance area per the 1954 census.

`ferttrends.csv` This file contains country-level fertility trends from 1740 to 1911 based on the crude birth rate from Chesnais (1992), the I_g index from Weir (1994), and the completed fertility of mothers from the Henry database.

`20km-from-henry.dta` This file contains the individuals in the geni data `fr-clean.csv` who were born within 20 kilometers of any location from the villages in the Louis Henry dataset. The locations of the Louis Henry dataset were manually georeferenced by the authors.

4.19 Familinx (1_19_familinx)

`profiles-anon.txt` This file contains the subsample of Geni profiles in the familinx database. Note that this file is too large to be uploaded as-is on the JPE Dataverse—it is close to 15 GB. Hence, we provide it in a `zip` folder. Users should unzip this file before running the scripts.

`relations-anon.txt` This file contains the father-child links in the familinx database.

`LICENSE` This file contains the familinx licence.

4.20 Cadastre (1_20_cadastre)

`cadastre.xslx` This file contains the raw cadastre data for each municipality section together with inheritance rules, soil characteristics, area, and population of each municipality with cadastre information.

`cadastre_distribution.xlsx` This file contains the absolute and relative land distribution of treated and untreated municipalities based on cadastre data. It is generated by the `05_results-appendix-cadastre.do` file.

4.21 Agricultural census of 1852 (`1_21_agricensus`)

4.21.1 Shapefiles

`ARRONDISSEMENTS_1852` This folder contains the `ARRONDISSEMENTS_1852` shapefile of 1852 arrondissements based on Gay (2020)'s shapefile of 1870 arrondissements. It also contains the `ARRONDISSEMENTS_1852_RIVERS` shapefile which is produced by mapping the `COURS_D_EAU` shapefile of waterways to the `ARRONDISSEMENTS_1852` shapefile of arrondissements through the `Join attributes by location` command.

`BD CARTHAGE 2017 (SANDRE)` This folder contains the `COURS_D_EAU` shapefile of waterways of France produced by SANDRE (2017).

`FARM_SIZE` This folder contains the `FARM_SIZE` shapefile of arrondissement-level average farm size based on the 1852 agricultural census.

4.21.2 Other files

`arrondissements.xlsx` This file contains the nomenclature of 1852 *arrondissements* based on an adaptation of Gay (2021)'s 1870 arrondissement nomenclature available in Gay (2020).

`cultures_diverses.xls` This file contains a subset of the 1852 agricultural census produced by the French Ministry of agriculture (1858, 1860). It is based on the dataset produced by Marin and Marraud (2011).

`economie_rurale.xls` This file contains a subset of the 1852 agricultural census produced by the French Ministry of agriculture (1858, 1860). It is based on the dataset produced by Marin and Marraud (2011).

`av_farm_size.csv` This file contains arrondissement-level average farm size based on the 1852 agricultural census.

5 List of Tables and Figures

The tables below list all tables and figures from the main text and appendix, along with the datasets and scripts required to reproduce them. Datasets are located in the

3_outputs/3_1_datasets folder, and scripts, in the 2_scripts/2_2_analyses folders. Tables and figures are located in the3_outputs/3_2_main and 3_outputs/3_3_appendix folders.

5.1 List of Main Text Tables and Figures

Output	Dataset	Script
Table 1: Balancedness of pre-reform characteristics	final-henry.dta	01_results-paper-henry.do
Table 2: Difference-in-differences estimates, Henry database	final-henry.dta	01_results-paper-henry.do
Table 3: Flexible-trend difference-in-differences estimates, Henry database	final-henry.dta	01_results-paper-henry.do
Table 4: Spatial regression-discontinuity estimates, Geni database	final-geni.dta	02_results-paper-geni.do
Table 5: Heterogeneous effects by soil conditions for small versus large farms	final-henry.dta	01_results-paper-henry.do
Figure 1: Fertility decline in France, 1700–1850	final-henry.dta ferttrends.xlsx	01_results-paper-henry.do
Figure 2: Customs and inheritance systems in Ancien Régime France	See Section 3.3.2 (map)	
Figure 3: Observations across Egalitarian and Inegalitarian Inheritance	See Section 3.3.2 (map)	
Figure 4: Completed fertility by cohort across inheritance systems		
Figure 4.a: Fertility trends	final-henry.dta	01_results-paper-henry.do
Figure 4.b: Event-study estimates	final-henry.dta	01_results-paper-henry.do
Figure 5: Fertility and distance to the inheritance border		
Figure 5.a: Cohorts fertile before 1793	final-geni.dta	02_results-paper-geni.do
Figure 5.b: Cohorts fertile after 1793	final-geni.dta	02_results-paper-geni.do

5.2 List of Appendix Tables and Figures

Output	Dataset	Script
Table A1: Summary statistics (women born in 1700–1803), Henry database	`final-henry.dta`	`03_results-appendix-henry.do`
Table A2: Accuracy of Henry forms	Produced manually based on Séguy et al. (1999, p. 22)	
Table A3: Fertility control mechanisms, Henry database	`final-henry.dta`	`03_results-appendix-henry.do`
Table A4: Inheritance reforms and child mortality, Henry database	`final-henry.dta` `enfants.sas7bdat`	`03_results-appendix-henry.do`
Table A5: Inheritance reforms and childlessness across samples, Henry database	`final-henry.dta`	`03_results-appendix-henry.do`
Table A6: List of sources describing local effects of the inheritance reforms for land fragmentation	Produced manually	
Table A7: Heterogeneous effects by categories of soil textures, Henry database	`final-henry.dta`	`03_results-appendix-henry.do`
Table A8: Heterogeneous effects by soil conditions for small versus large farms, additional fertility outcomes, Henry database	`final-henry.dta`	`03_results-appendix-henry.do`
Table A9: Heterogeneous effects by soil conditions for small versus large farms, RD-DD, Geni database	`final-geni.dta`	`04_results-appendix-geni.do`
Table A10: Relationship between inheritance reforms and distribution factors determining women's bargaining power in the household, Henry database	`final-henry.dta`	`03_results-appendix-henry.do`
Figure B1: Judicial districts in Ancien Régime France	See Section 3.3.2 (map)	
Figure B2: Customary boundaries based on Klimrath (1837)	Produced manually from Klimrath (1837)	
Figure B3: Written-law and customary-law areas	See Section 3.3.2 (map)	
Figure B4: Example of an entry in the nominative part of the Henry database	Produced manually from Séguy (2001)	
Figure B5: Crude and net completed fertility for women born in 1650–1800, Henry database	`final-henry.dta`	`03_results-appendix-henry.do`

Figure B6: Trends in completed fertility of mothers, Geni versus Henry	`final-henry.dta`
	`fr-clean.csv`
	`familinx-year-series.csv`
	`20km-from-henry.dta`
		`03_results-appendix-henry.do`

Figure B7: Inheritance systems in four administrative centers	See Section 3.3.2 (map)	
Table C1: Fertility convergence after the inheritance reforms reforms	`mouvement_population.dta.dta` `fam1954.dta`	`03_results-appendix-henry.do`
Figure C1: Overall implications of the inheritance reform for the demographic transition	`final-henry.dta`	`03_results-appendix-henry.do`
Table D1: Placebo test	`final-henry.dta`	`03_results-appendix-henry.do`
Table D2: Alternative definitions of sample, exposure, and treatment group	`final-henry.dta`	`03_results-appendix-henry.do`

Continued on next page

Continued from previous page

Output	Dataset	Script
Table D3: Effects of abolishing impartible inheritance	final-henry.dta	03_results-appendix-henry.do
Table D4: Effects of including women in inheritances	final-henry.dta	03_results-appendix-henry.do
Table D5: Robustness to migration and changes in mortality	final-henry.dta	03_results-appendix-henry.do
Table D6: Adjusted fertility using the first-name repetition technique	final-henry.dta	03_results-appendix-henry.do
Table D7: Robustness controlling for soil, climate, and terrain characteristics	final-henry.dta	03_results-appendix-henry.do
Table D8: Robustness to excluding three outliers in Alps and Pyrenees with extreme ruggedness	final-henry.dta	03_results-appendix-henry.do
Table D9: Alternative inference methods using the wild bootstrap procedure	final-henry.dta	03_results-appendix-henry.do
Figure D1: Permutation tests	final-henry.dta	03_results-appendix-henry.do
Figure D2: Sensitivity of main estimates to outliers	final-henry.dta	03_results-appendix-henry.do
Figure D3: Baseline versus placebo exercise	final-henry.dta	03_results-appendix-henry.do
Figure D4: Alternative definitions of treatment	final-henry.dta	03_results-appendix-henry.do
Figure D5: Conley adjusted standard errors with different distance cutoffs	final-henry.dta	03_results-appendix-henry.do
Table E1: Covariate balance test for spatial regression-discontinuity analysis	final-geni.dta	04_results-appendix-geni.do
Table E2: Sensitivity to additional RD-DD specifications	final-geni.dta	04_results-appendix-geni.do
Figure E1: Balance RD plots	final-geni.dta	04_results-appendix-geni.do
Figure E2: Trends in completed fertility under reformed and not reformed inheritance	final-geni.dta	04_results-appendix-geni.do
Figure E3: Sensitivity to bandwidth choice for spatial RD-DD	final-geni.dta	04_results-appendix-geni.do
Figure E4: Conley adjusted standard errors with different distance cutoffs	final-geni.dta	04_results-appendix-geni.do
Table F1: Sample of parcel-level cadastral information	Produced manually	
Table F2: Average land size	cadastre.dta	05_results-appendix-cadastre.do
Table F3: Inequality indices in the distribution of land	cadastre.dta	05_results-appendix-cadastre.do
Table F4: Summary statistics for soil and land characteristics	agricensus.dta	06_results-appendix-agricensus.do
Table F5: Soil characteristics and farm size in 1852	agricensus.dta	06_results-appendix-agricensus.do
Figure F1: Cadastral map of section D of Echevronne	Produced manually	
Figure F2: Parcel-level table of section D of Echevronne	Produced manually	
Figure F3: Distribution of parcels across size bins	cadastre.dta	05_results-appendix-cadastre.do
Figure F4: Cumulative distributions of parcel sizes	cadastre.dta	05_results-appendix-cadastre.do

Continued on next page

Continued from previous page

Output	Dataset	Script
Figure F5: Terrain and soil characteristics	See Section 3.3.2 (map)	
Figure F6: Soil texture diagram	Produced manually from USDA	
Figure F7: 1852 arrondissement-level terrain and soil characteristics	See Section 3.3.2 (map)	
Figure F8: 1852 arrondissement-level farming land structure	See Section 3.3.2 (map)	
Figure F9: Farming land structure and soil characteristics	`agricensus.dta`	`06_results-appendix-agricensus.do`
Figure G1: Relationship between fertility and land under egalitarian and inegalitarian inheritance	Simulation	`07_model-appendix.nb`
Figure H1: Lent and advent marriages between 1700 and 1815	`final-henry.dta`	`03_results-appendix-henry.do`
Figure H2: Balancedness on the religiosity index	`final-henry.dta`	`03_results-appendix-henry.do`
Figure H3: Raster map of wheat prices in the 1780s	See Section 3.3.2 (map)	
Figure H4: Spatial distribution of administrative centers in 1789	See Section 3.3.2 (map)	
Figure H5: Spatial distribution of political societies and rebellions	See Section 3.3.2 (map)	
Figure H6: Spatial distribution of paved roads and horse-post relays	See Section 3.3.2 (map)	
Figure H7: Land characteristics	See Section 3.3.2 (map)	

6 Technical Requirements

6.1 Software

Most scripts are executed using Stata (SE version 18.5). Geni processing scripts are executed in RStudio (version 2024.12.1+563). The model figure G1 is produced using Mathematica (version 14.1.0). Maps and mapping manual operations are carried out using QGIS (version 3.40.5–Bratislava).

6.2 Packages and Computing Time

Required packages are installed automatically at the beginning of each script.[5] The overall computing time is about 5 hours using the following operating system:

[5]One exception is the `georeference` R package, which requires first installing `devtools` as follows: `install.packages('devtools')`, `library(devtools)`, and `install_github('editio/georeference')`.

- **Processor** Intel(R) Core(TM) i9-14900K @ 3.20 GHz
- **RAM** 64.0 GB
- **System** TYPE 64-bit

Script	Required packages	Computing time
2_1_data		2 hours, 3 minutes
00_henry-data-prep.do	tabmiss	1 second
01_geni_data_to_sample	arrow 20.0.0.2 (Richardson et al. 2025) data.table 1.17.0 (Barrett et al. 2025) dplyr 1.1.4 (Wickham et al. 2023) geonames 0.999 (Rowlingson 2019) progress 1.2.3 (Csárdi and FitzJohn 2023) readxl 1.4.5 (Wickham and Bryan 2025) RecordLinkage 0.4.12.4 (Sariyar and Borg 2022) sf 1.0.20 (Pebesma 2018) stringr 1.5.1 (Wickham 2023) tictoc 1.2.1 (Izrailev 2024) tidytable 0.11.2 (Fairbanks 2024) tidyverse 2.0.0 (Wickham et al. 2019)	1 hour, 50 minutes
02_geni-inheritance-and-controls.do	fre (Jann 2007)	10 seconds
03_geni-final-data-prep.do	None	2 minutes
04_familinx-data-prep.do	None	10 minutes
05_cadastre-data-prep.do	None	10 seconds
06_agricensus-data-prep.do	None	2 seconds
2_1_analyses		3 hours, 1 minute
01_results-paper-henry.do	tabmiss reghdfe (Correia 2014) ftools (Correia 2016) coefplot (Jann 2013) outreg2 (Wada 2005)	14 seconds
02_results-paper-geni.do	rdrobust (Calonico et al. 2018) reghdfe (Correia 2014) ftools (Correia 2016) outreg2 (Wada 2005)	2 minutes
03_results-appendix-henry.do	reghdfe (Correia 2014) ftools (Correia 2016) estout (Jann 2004) outreg2 (Wada 2005) grc1leg2 (Over 2024) ritest (Hess 2019) coefplot (Jann 2013) acreg (Colella et al. 2020) ranktest (Kleibergen et al. 2007) hdfe (Correia 2015)	14 minutes
04_results-appendix-geni	rdrobust (Calonico et al. 2018) reghdfe (Correia 2014)	2 hours, 44 minutes

Continued on next page

Continued from previous page

Script	Required packages	Computing time
	`ftools` (Correia 2016) `outreg2` (Wada 2005) `grc1leg2` (Over 2024) `coefplot` (Jann 2013) `acreg` (Colella et al. 2020) `ranktest` (Kleibergen et al. 2007) `hdfe` (Correia 2015)	
`05_results-appendix-cadastre`	`fre` (Jann 2007) `ineqdeco` (Jenkins 1999)	3 seconds
`06_results-appendix-agricensus`	`reghdfe` (Correia 2014) `ftools` (Correia 2016) `binscatter` (Stepner 2013)	3 seconds
`07_model-appendix.nb`		30 seconds

7 Licenses

This work is under the Attribution 4.0 International license (CC BY 4.0). https://creativecommons.org/licenses/by/4.0. You are free to:

- Share: copy and redistribute the material in any medium or format
- Adapt: remix, transform, and build upon the material for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms under the following terms:

- Attribution: you must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
- No additional restrictions: you may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

Scripts The scripts in this replication package are under the MIT License. https://spdx.org/licenses/MIT.html#licenseText. Copyright (c) 2025 Victor Gay, Paula E. Gobbi, and Marc Goñi. Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software,

and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

References

Albertus, M., and V. Gay. 2025. "The Road to Rebellion: Rural Uprisings and State-Building in the Run-Up to the French Revolution." TSE Working Paper 24–1557.

Barrett, T., M. Dowle, A. Srinivasan, J. Gorecki, M. Chirico, T. Hocking, B. Schwendinger, and I. Krylov. 2025. *`data.table`: Extension of `data.frame`*. R package version 1.17.0. https://CRAN.R-project.org/package=data.table.

Boutier, J., P. Boutry, and S. Bonin. 1992. *Atlas de la Révolution française, Tome 6. Les sociétés politiques*. Paris: EHESS.

Brette, A. 1904. *Atlas des bailliages ou juridictions assimilées ayant formé unité électorale en 1789 dressé d'après les actes de la convocation conservés aux Archives nationales*. Paris: Imprimerie nationale.

Buringh, E. 2020. *European urban population, 700 - 2000 [database]*. DANS Data Station Social Sciences and Humanities.

Calonico, S., M. D. Cattaneo, M. H. Farrell, and R. Titiunik. 2022 [2018]. *`RDROBUST`: Stata Module to Provide Robust Data-Driven Inference in the Regression-Discontinuity Design*. Boston, MA: Statistical Software Components, Boston College Department of Economics. https://ideas.repec.org/c/boc/bocode/s458483.html.

Chesnais, J.-C. 1992. *The Demographic Transition*. Oxford: Oxford University Press.

Colella, F., R. Lavive, S. O. Sakalli, and M. Thoenig. 2020. *`ACREG`: Stata Module to Perform Arbitrary Correlation Regression*. Boston, MA: Statistical Software Components, Boston College Department of Economics. https://ideas.repec.org/c/boc/bocode/ss458889.html.

Correia, S. 2023 [2014]. *`REGHDFE`: Stata Module to Perform Linear or Instrumental-Variable Regression Absorbing Any Number of High-Dimensional Fixed Effects*. Boston, MA: Statistical Software Components, Boston College Department of Economics. https://ideas.repec.org/c/boc/bocode/s457874.html.

———. 2016 [2015]. *`HDFE`: Stata Module to Partial Out Variables with Respect to a Set of Fixed Effects*. Boston, MA: Statistical Software Components, Boston College Department of Economics. https://ideas.repec.org/c/boc/bocode/s457985.html.

———. 2023 [2016]. *FTOOLS: Stata Module to Provide Alternatives to Common Stata Commands Optimized for Large Datasets*. Boston, MA: Statistical Software Components, Boston College Department of Economics. https://ideas.repec.org/c/boc/bocode/s458213.html.

Cristofoli, P., M.-C. Vouloir, B. Duménieu, and C. Motte. 2021. *Des chefs-lieux de Cassini aux communes de France (1756–1999) [database]*. Paris: Laboratoire de démographie et d'histoire sociale [producer]; Paris: EHESS [distributor]. https://doi.org/10.5281/zenodo.4515223.

Csárdi, G., and R. FitzJohn. 2023. *progress: Terminal Progress Bars*. R package version 1.2.3. https://CRAN.R-project.org/package=progress.

EROS. 2018. *Digital Elevation – Global 30 Arc-Second Elevation (GTOPO30) [database]*. Earth Resources Observation and Science (EROS) Center [distributor]. https://doi.org/10.7910/DVN/HJISNR.

Fairbanks, M. 2024. *tidytable: Tidy Interface to data.table*. R package version 0.11.2. https://CRAN.R-project.org/package=tidytable.

FAO. 2012. *Global Agro-Ecological Zones (GAEZ v3.0) [database]*. Food and Agriculture Organization of the United Nations [distributor]. https://www.gaez.iiasa.ac.at/.

Fleury, M., and L. Henry. 1958. "Pour connaître la population de la France depuis Louis XIV. Plan de travaux par sondage." *Population* 13 (4): 663–86.

Galor, O., and Ö. Özak. 2016. "The Agricultural Origins of Time Preference." *American Economic Review* 106 (10): 3064–103.

Gay, V. 2020. *TRF-GIS Arrondissements (1870–1940) [database]*. Harvard Dataverse [distributor]. https://doi.org/10.7910/DVN/98ERPB.

———. 2021. "Mapping the Third Republic. A Geographic Information System of France (1870–1940)." *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 54 (4): 189–207.

———. 2025. "The Jean Nicolas Database. The French Rebellion, 1661–1789."

Gay, V., P. E. Gobbi, and M. Goñi. 2023a. *Bailliages in 1789 France [database]*. Harvard Dataverse [distributor]. https://doi.org/10.7910/DVN/T8UXHK.

———. 2023b. *Customs in Ancien Régime France [database]*. Harvard Dataverse [distributor]. https://doi.org/10.7910/DVN/HYE2O9.

———. 2024a. "The Customary Atlas of Ancien Régime France." *Explorations in Economic History* 93:101588.

———. 2024b. "The Atlas of Local Jurisdictions of Ancien Régime France." *Journal of Historical Geography* 84:49–60.

———. 2025. "Revolutionary Transition: Inheritance Change and Fertility Decline." *Journal of Political Economy*.

Hess, S. 2020 [2019]. *RITEST: Stata Module to Perform Randomization Inference and Permuation Tests*. Boston, MA: Statistical Software Components, Boston College Department of Economics. https://ideas.repec.org/c/boc/bocode/s458599.html.

IGN. 2011. *Geofla Communes Édition 2011 France métropolitaine [database].* Paris: IGN. https://geoservices.ign.fr/geofla.

INRA. 2018 [1998]. *Base de données géographique des sols de France à 1/1 000 000 version 3.2.8.0 [database].* Harvard Dataverse [distributor]. https://doi.org/10.15454/BPN57S.

INSEE. 2021 [1954]. *Enquête famille 1954 [database].* Progedo-Adisp [distributor]. https://doi.org/10.13144/lil-0283.

Izrailev, S. 2024. *tictoc: Functions for Timing R Scripts, as Well as Implementations of "Stack" and "StackList" Structures.* R package version 1.2.1.

Jann, B. 2023 [2004]. *ESTOUT: Stata Module to Make Regression Tables.* Boston, MA: Statistical Software Components, Boston College Department of Economics. https://ideas.repec.org/c/boc/bocode/s439301.html.

———. 2021 [2007]. *FRE: Stata Module to Display One-Way Frequency Table.* Boston, MA: Statistical Software Components, Boston College Department of Economics. https://ideas.repec.org/c/boc/bocode/s456835.html.

———. 2023 [2013]. *COEFPLOT: Stata Module to Plot Regression Coefficients and Other Results.* Boston, MA: Statistical Software Components, Boston College Department of Economics. https://ideas.repec.org/c/boc/bocode/s457686.html.

Jenkins, S. P. 2024 [1999]. *INEQDECO: Stata Module to Calculate Inequality Indices with Decomposition by Subgroup.* Boston, MA: Statistical Software Components, Boston College Department of Economics. https://ideas.repec.org/c/boc/bocode/s366002.html.

Kaplanis, J., A. Gordon, T. Shor, O. Weissbrod, D. Geiger, M. Wahl, M. Gershovits, B. Markus, M. Sheikh, M. Gymrek, G. Bhatia, D. G. MacArthur, A. L. Price, and Y. Erlich. 2018. "Quantitative Analysis of Population-Scale Family Trees with Millions of Relatives." *Science* 360 (6385): 171–5.

Kleibergen, F., M. E. Schaffer, and F. Windmeijer. 2020 [2007]. *RANKTEST: Stata Module to Test the Rank of a Matrix.* Boston, MA: Statistical Software Components, Boston College Department of Economics. https://ideas.repec.org/c/boc/bocode/s456865.html.

Klimrath, H. 1837. *Études sur les coutumes.* Paris: Levraud.

Marin, B., and M. Marraud. 2011. "L'enquête agricole de 1852." *L'Atelier du Centre de recherches historiques*, vol. Les Enquêtes Collectives du CRH. https://doi.org/10.4000/acrh.3696.

Ministère des travaux publics, de l'agriculture et du commerce. 1858. *Statistique de la France: statistique agricole. Première partie.* Paris: Imprimerie impériale.

———. 1860. *Statistique de la France: statistique agricole. Seconde partie.* Paris: Imprimerie impériale.

Nicolas, J. 2002. *La Rébellion française. Mouvements populaires et conscience sociale, 1661–1789.* Paris: Gallimard.

Nordman, D., M.-V. Ozouf-Marignier, and A. Laclau. 1989. *Atlas de la Révolution française, Tome 5. Le territoire (2).* Paris: EHESS.

Nunn, N., and D. Puga. 2012. "Ruggedness: The Blessing of Bad Geography in Africa." *Review of Economics and Statistics* 94 (1): 20–36.

Over, M. 2024. *GRC1LEG2: Stata Module to Combine Multiple Graphs with a Single Common Legend*. Boston, MA: Statistical Software Components, Boston College Department of Economics. https://ideas.repec.org/c/boc/bocode/s459360.html.

Pebesma, E. 2018. "Simple Features for R: Standardized Support for Spatial Vector Data." *The R Journal* 1 (10): 439–46.

Perret, J., M. Gribaudi, and M. Barthelemy. 2015a. "Roads and Cities of 18th Century France." *Scientific Data* 2:150048.

Perret, J., M. Gribaudi, M. Barthelemy, N. Abadie, S. Baciocchi, C. Bertelli, O. Bonin, P. Bordin, B. Costes, P. Cristofoli, B. Dumenieu, J. Gravier, J.-P. Hubert, P.-A. Le Ny, E. Mermet, C. Motte, M. Pardoen, A.-M. Raimond, S. Robert, and M.-C. Vouloir. 2015b. *The 18th century Cassini roads and cities dataset [database]*. Harvard Dataverse [distributor]. https://doi.org/10.7910/DVN/28674.

Richardson, N., I. Cook, N. Crane, D. Dunnington, R. François, J. Keane, D. Moldovan-Grünfeld, J. Ooms, J. Wujciak-Jens, and Apache Arrow. 2025. *arrow: Integration to 'Apache' 'Arrow'*. R package version 20.0.0.2.

Ridolfi, L. 2019. "Six Centuries of Real Wages in France from Louis IX to Napoleon III: 1250–1860." *The Journal of Economic History* 79 (3): 589–627.

Rowlingson, B. 2019. *geonames: Interface to the Geonames Spatial Query Web Service*. R package version 0.999? https://CRAN.R-project.org/package=geonames.

SANDRE. 2017. *Cours d'eau – Métropole 2017 – BD Carthage [database]*. Service d'administration nationale des données et référentiels sur l'eau [distributor]. https://www.sandre.eaufrance.fr/atlas/atlas/api/records/7381de46-42f7-42df-9abe-0ecd4b946034.

Sariyar, M., and A. Borg. 2022. *RecordLinkage: Record Linkage Functions for Linking and Deduplicating Data Sets*. R package version 0.4-12.4. https://CRAN.R-project.org/package=RecordLinkage.

Séguy, I. 2001. *La population de la France de 1670 à 1829. L'enquête Louis Henry et ses données*. Paris: INED.

Séguy, I., H. Colençon, and C. Méric. 1999. "Enquête Louis Henry. Notice descriptive de la partie nominative." INED documents de travail 76.

Squicciarini, M. P. 2020a. *Data and Code for: Devotion and Development: Religiosity, Education, and Economic Progress in 19th-Century France [database]*. American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E119862V1.

———. 2020b. "Devotion and Development: Religiosity, Education, and Economic Progress in Nineteenth-Century France." *American Economic Review* 110 (11): 3454–91.

Statistique générale de la France. 1861. *Mouvement de la population pendant les années 1855, 1856 et 1857*. Strasbourg: Imprimerie administrative de Veuve Berger-Levrault.

Stepner, M. 2013. *BINSCATTER: Stata Module to Generate Binned Scatterplots*. Boston, MA: Statistical Software Components, Boston College Department of Economics. https://ideas.repec.org/c/boc/bocode/s457709.html.

Tackett, T. 1986. *Religion, Revolution, and Regional Culture in Eighteenth-Century France: The Ecclesiastical Oath of 1791*. Princeton: Princeton University Press.

Wada, R. 2014 [2005]. *OUTREG2: Stata Module to Arrange Regression Outputs into an Illustrative Table*. Boston, MA: Statistical Software Components, Boston College Department of Economics. https://ideas.repec.org/c/boc/bocode/s456416.html.

Weir, D. R. 1994. "New Estimates of Nuptiality and Marital Fertility in France, 1740–1911." *Population Studies* 48 (2): 307–31.

Wickham, H. 2023. *stringr: Simple, Consistent Wrappers for Common String Operations*. R package version 1.5.1.

Wickham, H., M. Averick, J. Bryan, W. Chang, L. D. McGowan, R. François, G. Grolemund, A. Hayes, L. Henry, J. Hester, M. Kuhn, T. L. Pedersen, E. Miller, S. M. Bache, K. Müller, J. Ooms, D. Robinson, D. P. Seidel, V. Spinu, K. Takahashi, D. Vaughan, C. Wilke, K. Woo, and H. Yutani. 2019. "Welcome to the `tidyverse`." *Journal of Open Source Software* 4 (43): 1686.

Wickham, H., and J. Bryan. 2025. *readxl: Read Excel Files*. R package version 1.4.5. https://CRAN.R-project.org/package=readxl.

Wickham, H., R. François, L. Henry, K. Müller, and D. Vaughan. 2023. *dplyr: A Grammar of Data Manipulation*. R package version 1.1.4.