

README – Choi and Joubert (2026)

Contents

1. [Overview](#)
2. [Data Availability](#)
3. [Instructions for Replicators](#)
4. [List of Exhibits](#)
5. [Requirements](#)
6. [Folder Structure](#)

Overview

This package contains the code required to construct the analytical datasets and reproduce the tables, figures, and simulations in the paper. The main Stata do-file runs four tasks in sequence:

1. Construct the EPS individual-year panel.
2. Harmonize and append the CASEN surveys.
3. Prepare the Vital Statistics data for analysis.
4. Produce the final tables and figures.

Before running the code, replicators must update the user-specific paths at the beginning of the main do-file. Tasks should be run in order because later steps depend on outputs created earlier. Most of the analysis is written in Stata, while a small set of simulations is produced in MATLAB.

Data Availability

Data Sources

Data Source 1: Encuesta de Proteccion Social (EPS) with linked Administrative Pension Records

- **Description:** Longitudinal survey (EPS), rounds 2002 to 2015 individually linked to administrative pension contribution histories for the main survey respondent.
- **Source:** The restricted version of the EPS, which contains administrative records on pension contributions was made available by the Subsecretaría de Previsión Social.
- **Access year:** 2021.

- **Data availability:** The public version of the EPS survey can be downloaded from <https://previsionsocial.gob.cl/datos-estadisticos/descargar-bases-de-datos-eps/> but the individual link to the administrative pension records is not publicly available. The files are provided for the purpose of replication only in the folder .\OneDrive – WBG\RawDataDivorcio\raw_data\ but should not be published.

Data Source 2: Administrative Records on Marriages, Births, Divorces

- **Description:** Administrative records on marriages; administrative records on births; administrative records on divorces; administrative records on marriage annulments.
- **Source:** Chilean Civil Registry and Identification Service (Servicio de Registro Civil e Identificación, SRCel). The data were provided by the Transparency Unit of the Civil Registry in response to a request submitted under Chile's Transparency Law.
- **Access year:** 2021.
- **Data availability:** No data can be made publicly available. The data were requested through a formal public information request under Law No. 20,285 on Access to Public Information. The request was submitted to the Transparency Unit of the Civil Registry. Access was granted subject to the exclusions and limitations established by the institution, mainly to protect personal data, statistical confidentiality, and information outside the Civil Registry's institutional competence. Because of the size of the files, the data were made available for in-person retrieval in CD format. The files are provided for the purpose of replication only in the folder .\OneDrive – WBG\RawDataDivorcio\raw_data\ but should not be published.

Data Source 3: Encuesta de Caracterización Socioeconómica Nacional

- **Description:** Repeated cross section of the CASEN survey, rounds 1990 to 2017
- **Source:** Ministerio de Desarrollo Social y Familia – Encuesta de Caracterización Socioeconómica Nacional (CASEN).
- **Access year:** 2021.
- **Data availability:** All data are publicly available. The files are provided for the purpose of replication in the folder .\OneDrive – WBG\RawDataDivorcio\raw_data\ but are also available from <https://observatorio.ministeriodesarrollosocial.gob.cl/encuesta-casen#objetivos>

Additional Public Inputs

The replication package also uses publicly available auxiliary files to harmonize geographic identifiers and administrative divisions:

- DPA2018.xls, from the Departamento de Estadísticas e Información de Salud (DEIS), available at <https://repositoriodeis.minsal.cl/ContenidoSitioWeb2020/uploads/2019/11/DPA2018.xls>;
- CUT_2018_v04.xls, from the Subsecretaría de Desarrollo Regional y Administrativo (SUBDERE), available at https://www.subdere.gov.cl/sites/default/files/documentos/CUT_2018_v04.xls. These files are used only for geographic coding and are publicly available.

The package also uses publicly available Google Trends data for the search term “divorcio,” downloaded from Google Trends for January 2004–December 2009 and used only to construct the descriptive figure on search interest around the implementation of the divorce law.

Statement about Rights

- The authors do not have permission to publicly redistribute all data used in this replication package. Parts of the EPS and CASEN data can be downloaded from the given links, but the administrative records from the Chilean Pension Superintendence and the Chilean Civil Registry are subject to confidentiality and data-protection restrictions. We are therefore providing the raw data for the purpose of replication review but they should not be publicly posted, redistributed, or made available beyond the verification process.

Instructions for Replicators

New users should follow these steps to run the package successfully:

1. Raw data files will be provided to the replicators in a separate folder. The data can be copied into the data/raw_data folder, or the path to the raw data folder can be updated in the macro data_rawdata in the 01_master_project.do
2. Ensure all required software and dependencies are installed as listed in the [Requirements](#) section.
3. Run the 01_master_project.do file.

List of Exhibits

The code provided reproduces the following exhibits. All tables and figures are generated by 02_tables_figures.do:

Exhibit name	Output filename	Note
Figure 1	figure_01.png	Found in output/figures
Figure 2	figure_02.png	Found in output/figures
Figure 3	figure_03.png	Found in output/figures
Figure 4	figure_04.png	Found in output/figures
Figure 5	figure_05.png	Found in output/figures
Figure 6	figure_06.png	Found in output/figures
Figure 7	figure_07.png	Found in output/figures
Figure 8	figure_08.png	Found in output/figures
Figure A.1	figure_a01.png	Found in output/figures
Figure A.2	figure_a02.png	Found in output/figures
Figure A.3	figure_a03.png	Found in output/figures
Table 1	table_01.xls	Found in output/tables
Table 2	table_02.xls	Found in output/tables
Table 3	table_03.xls	Found in output/tables
Table 4	table_04.xls	Found in output/tables
Table 5	table_05.xls	Found in output/tables
Table A.1	table_a01.xls	Found in output/tables
Table A.2	table_a02.xls	Found in output/tables
Table A.3	table_a03.xls	Found in output/tables

Requirements

Computational Requirements

The replication package is written primarily in Stata and was designed to be run from a single main do-file after updating the user-specific paths at the beginning of the code. The code assumes the folder structure described in the replication package.

Software Requirements

- **Stata version 16 or later**
 - texdoc
 - estout

- coefplot
 - csdid
 - reghdfe
 - ftools
- **MATLAB R2019b or later**
 - Statistics and Machine Learning Toolbox

Folder Structure

The replication package is organized into three main folders: code, data, and output.

- code
 - 00_master
 - 01_code_casen
 - 02_code_eps
 - 03_code_vital
 - 04_code_matlab
- data
 - raw_data
 - casen
 - eps
 - vital
 - working_data
 - casen
 - eps
 - vital
- output
 - figures
 - tables