

README for the Reproducibility Package for Labor Market Effects of Public Employment Guarantee

Overview

The code in this replication package uses six data sources using Stata 19 to produce 1 table and 11 figures in the main manuscript and 23 figures in the Appendix. The replicator should expect the code to run for approximately 20 hours given the large sample size.

Data Availability Statement

1. 1986, 1996, and 2006 Egyptian Population Censuses (10-percent Individual-Level Samples from IPUMS-International)

Source: Integrated Public Use Microdata Series, International

Year: 2024

URL: https://international.ipums.org/international-action/sample_details/country/eg#tab_eg1986a

Access Date (Month-Year): October, 2024

Note: : Raw IPUMS extract (1986, 1996, 2006 Egyptian Population, Housing and Establishments Census); also used for administrative unit harmonization; located at: 1_Population census/1_Raw IPUMS data/ and 0_Administrative units/1_Raw data/2_IPUMS/

Access Type: Accessible

License: IPUMS Conditions of Use

License URL: <https://international.ipums.org/international/cite.shtml>

Citation: Integrated Public Use Microdata Series, International. 2024. "1986, 1996, and 2006 Egyptian Population, Housing and Establishments Census" [dataset]. https://international.ipums.org/international-action/sample_details/country/eg#tab_eg1986a. Accessed October, 2024.

2. Annual Education Statistics 1959/1960 and 1969/1970 الإحصاء السنوي للتعليم (Egypt)

Source: Imprimerie Nationale

Year: 1960 (for school year 1959/1960) and 1970 (for school year 1969/1970)

URL: -Not available online.

Access Date (Month-Year): July 2012, December 2024

Note: data/2_Schools statistics/1_Digitized school statistics

Access Type: Restricted (copyright). Mohamed Saleh obtained the 1959/1960 report in December 2024 as a PDF through an inter-library loan from Princeton University to London School of Economics. Mohamed Saleh scanned (photographed) and digitized the 1969/1970 report in July 2012 from an original hard copy of the published report in Arabic at the Library of the Ministry of Education in Cairo, Egypt.

License: Custom License

License URL: -

Citations: Imprimerie Nationale. 1960. "Annual Education Statistics (Egypt)" [dataset]. Accessed July, 2012.

Imprimerie Nationale. 1970. "Annual Education Statistics (Egypt)" [dataset]. Accessed July, 2012.

3. Annuaire Statistique 1949-1950 et 1950-1951 (Egypt)

Source: Imprimerie Nationale

Year: 1953 (includes education statistics for 1951-1952)

URL: - Not available online.

Access Date (Month-Year): July 2014

Note: data/2_Schools statistics/1_Digitized school statistics

Access Type: Restricted (copyright). Downloaded by Mohamed Saleh as PDF from the website Central Agency for Public Mobilization and Statistics (CAPMAS) (Egypt) in July 2014. However, it is no longer available online. It may be available though as hard copy in libraries in Western universities.

License: -

License URL: -

Citation: Imprimerie Nationale. 1953. "Annuaire Statistique (Egypt)" [dataset]. Accessed October, 2024.

4. 1947, 1960, 1986 Egyptian Population Censuses (Sub-district Level Data)

Source: Centre d'Etudes et de Documentation Economiques, Juridiques, et Sociales (CEDEJ). A Century of Censuses CD-ROM: Egypt 1882- 1996. Cairo: CEDEJ, 2003.

Year: - 2003

URL: - Not available online

Access Date (Month-Year): April 2008. CD-ROM is available at various libraries. Mohamed Saleh downloaded the data for 1882-1996 from the CD-ROM in April 2008.

Note: This is the source for all three sub-district-level population census data for 1947, 1960, 1986:

data/1_Population census/3_Raw census data/1947.xls.

data/1_Population census/3_Raw census data/1960.xls.

data/1_Population census/3_Raw census data/1986.xls.

This is also the source for the Egyptian administrative units (1947, 1960, 1986) used for harmonization of administrative boundaries; located at 0_Administrative units/1_Raw data/0_References M. Saleh/1947_admin_units.xls, 1960_admin_units.xlsx, 1986_admin_units.xls.

Access Type: Restricted

License: -

License URL: -

Citation: -

5. Land Reform in Egypt (1961)

Source: Moshrif, R., 2025. Long-term Land Inequality and Post-Colonial Land Reform in Egypt (1896-2020)

Year: 2025

URL: Not available online.

Access Date (Month-Year): October, 2025. Data obtained upon request from Rowaida Moshrif upon the condition that authors do not share with third parties without Rowaida's consent. Rowaida digitized this dataset from the Agricultural Census for 1961. Details are in Moshrif (2025). Contact information : rowaida.moshrif@psemail.eu

Note: data/3_Land reform/1_Raw data

Access Type: Restricted

License: -

License URL: -

Citation: Moshrif, R., 2025. Long-term Land Inequality and Post-Colonial Land Reform in Egypt (1896-2020). 2025. "Land Reform in Egypt" [dataset]. Accessed October, 2025.

6. Egypt - Administrative Boundaries

Source: Humanitarian Data Exchange

Year: 2006

URL: <https://data.humdata.org/dataset/egypt-administrative-boundaries>

Access Date (Month-Year): August 2025

Note: data/6_Mapping data/Egypt_Admin_shape

Access Type: Open

License: Creative Commons Attribution 4.0 International (CC-BY-4.0)

License URL: <https://creativecommons.org/licenses/by/4.0/>

Citation: Humanitarian Data Exchange. 2006. "Egypt - Administrative Boundaries" [dataset].

<https://data.humdata.org/dataset/egypt-administrative-boundaries>. Accessed August 2025.

Statement about Rights

I certify that the author(s) of the manuscript have legitimate access to and permission to use the data used in this manuscript.

I certify that the author(s) of the manuscript have documented permission to redistribute/publish the data contained within this replication package. Appropriate permissions are documented in the LICENSE.txt file.

Instructions for Replicators

To produce the figures and tables available in the manuscript, the replicator should:

1. Use Stata version 19
2. Install the following Stata packages: estwide, spmap, spmatrix, xtile, shp2dta, coefplot, palettes, estout, reghdfe, ftools, require
3. Run the script "main.do" available in the "codes" folder, which executes the full workflow
4. The script will save all tables in "outputs/1_Tables" and all figures in "outputs/2_Figures"

Computational Requirements and Runtime

OS: macOS Tahoe version 26.5

Processor: Apple M2
Memory: 8 GB

Software Requirements

Stata version 19

Code Description

The full workflow can be executed by running the "main.do" script which in turn runs the following do-files :

- 1_cleaning ipums census 1986_1996_2006:** Cleans the 1986, 1996, and 2026 Egyptian Population Census (recode missing, clean age variable, rename variables, relabel variables, and order variables)
- 2_extracting IPUMS administrative units:** Extract birthplace and current residence administrative units from the 1986, 1996, and 2026 Egyptian Population Census to harmonize them with the 1947 administrative boundaries
- 3_harmonizing administrative units:** Harmonizes the birthplace and current residence administrative units from the 1986, 1996, and 2026 Egyptian Population Census with the 1947 administrative boundaries
- 4_merging harmonized administrative units:** Merges the 1947 administrative units with the 1951, 1959, and 1969 school statistics data
- 5_cleaning 1960 census:** Cleans the 1960 Egyptian Population Census and computes the population aged 15 to 19 years old at the district and province level to use it to compute our treatment measure
- 6_calculating treatment intensity variable:** Computes the treatment intensity variable at the district and province level
- 7_harmonize land reform districts:** Merges the 1947 administrative units with the land reform share data
- 8_merging data:** Merges all data sources together in one. dta
- 9_cleaning merged sample:** Cleans the merged data set and generates outcome variables needed for the analysis
- 10_secondary schools map:** Generate a map for Egypt with the share of secondary schools at the district level
- 11_neighbors treatment assignment:** Assigns the treatment value of neighbouring districts within 100km to use in the robustness checks
- 12_descriptive statistics:** Generate descriptive tables and figures
- 13_results:** Generate all regressions results

Folder Structure

Folders in the replication folder are organized as follows:

codes/: Contains all Stata script files. Includes a subfolder ado/ where Stata stores the required user-written commands, alongside the 14 do-files that constitute the analysis workflow.

data/: Contains all datasets used and generated throughout the analysis. It is organized into the following subfolders:

0_Administrative units/: Crosswalks and harmonized administrative unit codes used to standardize geographic identifiers across data sources.

1_Raw data/

0_References M. Saleh/: 1947, 1960, and 1986 administrative unit codes from M. Saleh, used as reference

1_Schools/: 1951 and 1959 school statistics administrative units, extracted for harmonization with 1947 units

2_IPUMS/: 1986, 1996, and 2006 Egyptian birth and residence administrative unit codes extracted from population censuses, for harmonization with 1947 units

2_Harmonized data/

0_References M. Saleh/: 1947 district unit codes saved as .dta

1_Schools statistics/: 1951 and 1959 school statistics administrative units harmonized with 1947 codes; saved as .xlsx and .dta

2_IPUMS: 1986, 1996, and 2006 birth and residence administrative unit codes harmonized with 1947 codes; saved as .xlsx and .dta

3_Census/: 1960 administrative unit codes harmonized with 1947 codes; saved as .xlsx and .dta

1_Population census/: Population census data at various stages of processing.

1_Raw IPUMS data/: Raw IPUMS microdata as downloaded

2_Raw 1960 census/: Raw 1960 Egyptian census data

3_Clean data/: Cleaned 1960, 1986, 1996, and 2006 population census data prepared for analysis

2_Schools statistics/: School statistics data at various stages of processing

1_Digitized school statistics: Digitized 1951, 1959, and 1969 school statistics

2_Clean dta: 1951, 1959, and 1969 school statistics merged with 1947 administrative unit codes; district- and province-level treatment variables generated

3_Land reform/: Land reform data at various stages of processing

1_Raw data/: Raw land reform data

2_Clean data/: Cleaned land reform data harmonized with 1947 administrative units codes

4_Merged_data/: Census and school statistics datasets merged into a single file

5_Analysis data: Final dataset used for analysis, after construction of all outcome variables.

6_Mapping data/: Geographic data used for mapping and spatial robustness checks

Egypt_Admin_shape/: Egypt boundary shapefiles

Egypt_Admin_shape_dta/: Egypt geographical data used to generate maps

Geographic_location_dta/: Egypt geographical data used for the spatial robustness check

outputs/: Contains all outputs generated by Stata. Includes two subfolders:

1_Tables/

2_Figures/