Read-Me

Gregory Casey, Stephie Fried, and Ethan Goode

Overview

This replication package provides all data inputs and programs required to reproduce the paper's results. The general structure of the replication package divides programs into the /programs/ subdirectory, data inputs into the /data/raw/ or /data/climProj folders, and results into /results/tables/ or /results/figures/ based on the generated output. Programs are written in R, Stata, and Matlab and take approximately x hours to run from start to finish. This *"read-me"* provides documentation of all data inputs, computational requirements, and program functionality. Explicit instructions describing the sequence required to replicate all results are included in the section titled *"Instructions for replicators"*.

Data Availability and Provenance Statements

Statement about Rights

- We certify that the authors of the manuscript have legitimate access to, and permission to, use the data used in the manuscript.

Summary of Availability

- All data **are** publicly available.

Details on each data source

The following list provides details on where all raw inputs into the project are sourced from.

- Climate model projections are downloaded directly from the World Meteorological Organization's KNMI Climate Change Atlas (KNMI and (WMO), n.d.), hosted at "http://climexp.knmi.nl/plot_atlas_form.py?id=" as of writing. Country specific climate model output is stored in "/data/climProj/". The following options were used to download climate model output for each country:

  - Type: countries
  - Season: First month "Jan", length "12" months
  - Dataset: GCM: CMIP5 (full set)
  - Variable: "near-surface temperature", "absolute"
  - Output: "time series"
  - Scenario(s): "RCP8.5"
  - Plot period: "1900-2100"

- Historical country-level weather data is sourced from the replication data for "Global nonlinear effect of temperature on economic production" (Miguel, Burke, and Hsiang 2016). We download the data directly from "https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/LSDCAY". The data is stored in "/data/raw/GrowthClimateDataset.csv".
- Country-level economic data comes from the Penn World table version 10.0 (Feenstra, Inklaar, and Timmer 2015). The data can be downloaded directly from "https://www.rug.nl/ggdc/docs/pwt100.xlsx". This data can be found in "data/raw/pwt100.xlsx".
- Data on value-added from agriculture as percent of total GDP is collected from two sources:
  - The bulk of the data is collected from the World Bank (Bank 2012), housed at "https://data.worldbank.org/indicator/NV.AGR.TOTL.ZS", but available for direct download from "https://api.worldbank.org/v2/en/indicator/NV.AGR.TOTL.ZS?downloadformat=csv". This data is stored in "data/raw/API_NV.AGR.TOTL.ZS_DS2_en_csv_v2_4258562.csv".
  - We supplement data for the country of Vietnam using a series sourced from the Economic Transformation Database (Vries and Woltjer 2021). Data can be directly downloaded from "https://www.rug.nl/ggdc/docs/etd-release2021.xlsx".
- Classifications of countries into different regions are sourced from the appendix for "Temperature Shocks and Economic Growth: Evidence from the Last Half Century" (Dell, Jones, and Olken 2012), which can be downloaded from https://www.aeaweb.org/aej/mac/app/2010-0092_app.pdf

Dataset list

Raw data (all files are provided in /data/raw/ and /data/climProj/ are publicly available)

Data file	Source	Notes
API_NV.AGR.TOTL.ZS_DS2_en_csv_v2_4258562.csv	World Bank	Value-added as percent of total GDP by country.
CountryTempChange_RCP85_NA.csv	(Miguel, Burke, and Hsiang 2016)	Temperature change by country in 2100 under the RCP 8.5 climate scenario.
etd-release2021.xlsx	Groningen Growth and Development Centre	Value-added as percent of total GDP for developing countries, notably Vietnam.
GrowthClimateDataset.csv	(Miguel, Burke, and Hsiang 2016)	Historical temperature, precipitation.
pwt100.xlsx	Groningen Growth and Development Centre	Historical GDP, Capital, Investment, Labor, and population.
regional_classification_tblDJO.xlsx	(Dell, Jones, and Olken 2012)	Classification of countries into regions.
/data/climProj	(KNMI and (WMO), n.d.)	Temperature projections out to 2100 under the RCP 8.5 scenario.

Cleaning (cleaning files are provided in /data/raw/ and are created manually by the authors)

Data file	Notes
projections_crosswalk.csv	Crosswalk for differing country naming conventions.
stata_cid_crosswalk.xlsx	Crosswalk for numeric country indicators to country names.

Derived data (all files in /data/output/ are produced by following the "Instructions for replicators" notes below)

Data file	Program	Notes
df_tfpv2.dta	create_data.R	Cleaned data for regressions.
regSampleTfp.csv	regression_tables_baseline.do	Defines countries included in the sample used for regressions.
regtable1.csv	regression_tables_baseline.do	Output from the regression presented in table 1.
regtable1_1990_dummy.csv	regression_tables_het.do	Output from the regression presented in table B6.
regtable1_linear_trends.csv	regression_tables_baseline.do	Output from the regression presented in table B3.
regtable1_quadratic_trends.csv	regression_tables_baseline.do	Output from the regression presented in table B4.

Data file	Program	Notes
regtable1_region _fe.csv	regression_tables _baseline.do	Output from the regression presented in table B5.
regtable2.csv	regression_tables _baseline.do	Output from the regression presented in table 2.
Results_bootstrap _tfp_le.csv	run_bootstrap_tfp.R	Model output using estimates from each bootstrap sample.
rich_ind.dta	regression_tables _het.do	Classification of countries into above or below median global income levels.
tfp8_irf_lp.csv	regression_tables _LP.do	Output from the regression presented in figure C6.
tfp13_irf_lp.csv	regression_tables _LP.do	Output from the regression presented in figure C6.
tfp18_irf_lp.csv	regression_tables _LP.do	Output from the regression presented in figure C6.

Computational requirements

Software requirements

- Stata (code was last run with version 15)
  - ivreg2 (Baum, Schaffer, and Stillman 2002)
  - estout Janna (2007)
- R (code was last run with R version 4.2.0 (2022-04-22 ucrt) – "Vigorous Calisthenics")
  - tidyverse 1.3.2
  - readxl 1.4.0
  - here 1.0.1
  - parallel 4.2.0
  - RColorBrewer 1.1-3
  - stargazer 5.2.3
  - viridis 0.6.2
  - scales 1.2.0
  - ggmap 3.0.0

  - foreign 0.8-82
  - maps 3.4.0
  - renv 0.15.5
- Matlab (code was last run with version R2020a)

Memory and runtime requirements

Approximate time needed to process the raw data on a standard (2019) laptop computer:

~9 seconds

Approximate time needed to process the bootstrap sample results on a standard (2019) laptop computer:

TBD

Approximate time needed to process the remaining results on a standard (2019) laptop computer:

~30 minutes

Description of programs

All programs required to replicate the paper can be found in /programs/. Programs can generally be categorized as cleaning data, iperforming analysis, or producing figures and tables. The remaining uncategorized programs play a supporting role by serving as containers for functions required to perform analysis for the project. These supporting programs are labeled below as "Set-up" programs.

Data cleaning programs

- create_data.R cleans and processes all data for the paper into a single dataset to be used for analysis.

Analysis programs

- regression_tables_baseline.do[1] runs the baseline specification and its derivatives.
- regression_tables_het.do runs specifications which allow for differential outcomes with respect to economic characteristics (e.g. agricultural vs. non-agricultural economies).
- regression_tables_lag.do runs specifications which experiment with the lag structure of temperature shocks.
- regression_tables_LP.do runs specifications which use the local projection estimator.
- regression_tables_clustering.do runs specifications which use clustered standard errors.
- regression_tables_mg.do runs specifications which use the mean group estimator.
- regression_tables_pwttfp.do runs specifications which use an alternative measure of TFP, pulled from the Penn World Tables.
- run_analysis_gdppc.R uses a reduced-form approach to simulate forward the time-path of GDP per capita.
- run_analysis_tfp.R simulates growth of GDP per capita and its components via the model described in the paper.
- run_bootstrap.do performs the baseline regression on 1000 bootstrapped samples.
- run_bootstrap_tfp.R simulates growth of GDP per capita and its components for each set of estimates obtained in run_bootstrap.do.

Figures and tables

- create_figures.R creates all non-map figures (excluding figure 1) included in the paper.
- create_maps.R creates all maps included in the paper.
- create_tables.R creates the table shown in Appendix A.
- create_fig1.m creates the first figure in the paper.

[1] While the main purpose of the programs with the prefix *"regression_tables"* is to obtain results, these programs also produce regression tables featured in the paper. For more details on mapping programs to specific tables see the section titled *"List of tables and figures created"*.

Set-up

- initialize_env.R loads in all required R libraries and custom functions.
- regSim.R houses the function used to simulate GDP per capita using a reduced-form approach.
- solowv2.R houses the Solow growth model used to simulate GDP per capita, and its components.
- bootSimv2.R houses the function used to simulate forward the time-path of TFP implied by estimates obtained on each of 1000 bootstrap samples.

Instructions for replicators

Replicators need only to run programs in the /programs/ directory to obtain all results included in the paper. Tables produced for this project can be found in /results/tables/ and figures can be found in /results/figures/. Intermediate datasets created by the programs are not included by default, but are saved during the replication process in /data/output/. To replicate the full paper, follow the instructions below.

1. Load required R libraries and custom functions: run /programs/initialize_env.R.
2. Clean all data: run /programs/create_data.R.
3. Perform regression analysis and produce results tables. These programs can be run in any order:
    - Run baseline specification including robustness checks that involve manipulating dependent and independent variables used for analysis: run regression_tables_baseline.do
    - Run specifications which allow for heterogeneous effects of temperature on outcome: run regression_tables_het.do
    - Run specifications which test the robustness of the baseline specification to different lag structures: regression_tables_lag.do
    - Run specifications which test the robustness of the baseline specification to the clustering of standard errors: regression_tables_clustering.do
    - Run specifications which test the robustness of the baseline specification to the local projections estimator: regression_tables_LP.do
    - Run specifications which test the robustness of the baseline specification to the mean group estimator: regression_ tables_mg.do
    - Run specifications which test the robustness of the baseline specification to different measures of TFP: regression_tables_pwttfp.do
4. Simulate time-path of the economy:

- Simulation via model for estimates obtained from regressing TFP on variables of interest: run run_analysis_tfp.R
- Simulation via reduced-form approach for estimates obtained from regression GDPPC on variables of interest: run_analysis_gdppc.R

5. Perform bootstrapping procedure and simulate time-path of the economy for each set of coefficients produced:
   - Estimate baseline specification on 1000 bootstrap samples: run run_bootstrap.do
   - Perform simulation for each set of estimates: run run_bootstrap_tfp.R
6. Create all figures and remaining tables. Programs create_maps.R and create_tables.R do **not** depend on output from step five, but are grouped here for the purpose of producing the remaining figures and tables in a single step:
   - Create all non-map figures: Run create_figures.R
   - Create all map figures: Run create_maps.R
   - Create Appendix A summary table: Run create_tables.R

The replicator can use /data/output/Results_bootstrap_tfp_le.csv provided in replication package to skip step five, which is the most time intensive step required to replicate the full project.

List of tables and figures created

Figure/table number	Program	Line Number	Output file
Figure 1	create_fig1.m	16-216	TFP2.eps, K.eps, Y.eps, gY.eps
Table 1	regression_tables_baseline.do	63-111	regtableOLS1.tex
Table 2	regression_tables_baseline.do	117-188	regtableOLS2.tex
Figure 2	create_maps.R	123-136	tempMap_2010.eps
Figure 3	create_maps.R	107-116	tempMap_diff.eps
Figure 4	create_maps.R	193-240	level_effects_model.eps
Figure 5	create_figures.R	363-471	Ydecomposition_reg2.eps
Figure 6	create_figures.R	23-40	worldY_level_effects_model.eps
Figure 7	create_figures.R	476-535	worldYdecomposition_reg2.eps
Figure 8	create_figures.R	46-76	worldY_effects_comparison.eps

Figure/table number	Program	Line Number	Output file
Figure 9	create_maps.R	142-187	growth_effects_rf.eps
Table A1	create_tables.R	12-35	sumtable.tex
Table B1	regression_tables_ clustering.do	60-94	regtableOLS1_ro_cluster.tex
Table B2	regression_tables _baseline.do	194-224	regtableOLS1_nolag.tex
Table B3	regression_tables _baseline.do	231-265	regtableOLS1_resids.tex
Table B4	regression_tables _baseline.do	557-648	regtableOLS1_lagsIV.tex
Table B5	regression_tables _baseline	557-635	regtableOLS1_firststage.tex
Table B6	regression_tables _baseline.do	510-552	regtableOLS1_4lagstfp.tex
Table B7	regression_tables _mg.do	43-171	mgtable.csv
Table B8	regression_tables _baseline.do	271-304	regtableOLS1_lintrends.tex
Table B9	regression_tables _baseline.do	309-342	regtableOLS1_trends.tex
Table B10	regression_tables _baseline.do	348-381	regtableOLS1_region.tex
Table B11	regression_tables _het.do	125-164	regtable1_1990_dummy.csv
Table B12	regression_tables _lag.do	70-141	regtable_lagtemp.tex
Table B13	regression_tables _pwttfp.do	75-109	regtableOLS1_ro_rtfpna.tex
Table B14	regression_tables _het.do	171-226	regtableOLS1_hetero.tex
Table B15	regression_tables _het.do	233-288	regtableOLS1_hetero_agriculture.tex
Table B16	regression_tables _het.do	326-365	regtable_tbar.tex
Table B17	regression_tables _het.do	293-318	regtable_tempint.tex
Table B18	regression_tables _baseline	460-504	regtableOLS2_quadratictrends.tex
Table B19	regression_tables _baseline.do	387-417	regtableOLS1_kpc.tex

Figure/table number	Program	Line Number	Output file
Table B20	regression_tables_baseline.do	424-454	regtableOLS1_emp.tex
Figure C1	create_figures.R	144-214	marginal_effects_dTemp.eps, marginal_effects_Temp.eps
Figure C2	create_figures.R	220-358	tfe_tfp_reg1.eps, tfe_tfp_reg2.eps, tfe_tfp_reg3.eps
Figure C3	create_figures.R	220-358	tfe_gdppc_reg1.eps, tfe_gdppc_reg2.eps, tfe_gdppc_reg3.eps
Figure C4	create_figures.R	541-705	TrendSig1.eps
Figure C5	create_figures.R	541-705	TrendSig2.eps
Figure C6	create_figures.R	712-752	tfp8_irf.eps, tfp13_irf.eps, tfp18_irf.eps

References

Bank, The World. 2012. "World Development Indicators." http://data.worldbank.org/data-catalog/world-development-indicators.

Baum, Christopher F, Mark E Schaffer, and Steven Stillman. 2002. "IVREG2: Stata module for extended instrumental variables/2SLS and GMM estimation." Statistical Software Components, Boston College Department of Economics. https://ideas.repec.org/c/boc/bocode/s425401.html.

Dell, Melissa, Benjamin F. Jones, and Benjamin A. Olken. 2012. "Temperature Shocks and Economic Growth: Evidence from the Last Half Century." *American Economic Journal: Macroeconomics* 4 (3): 66–95. https://doi.org/10.1257/mac.4.3.66.

Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer. 2015. "The Next Generation of the Penn World Table." *American Economic Review* 105 (10): 3150–82. www.ggdc.net/pwt.

Janna, Ben. 2005. "Making regression tables from stored estimates." *The Stata Journal* 5 (3): 288–308.

———. 2007. "Making regression tables simplified." *The Stata Journal* 7 (2): 227–44.

KNMI, and World Meteorological Organization (WMO). n.d. "KNMI Climate Atlas." http://climexp.knmi.nl/plot_atlas_form.py?id=.

Miguel, Edward, Marshall Burke, and Solomon M. Hsiang. 2016. "Replication Data for: Global non-linear effect of temperature on economic production." Harvard Dataverse. https://doi.org/10.7910/DVN/LSDCAY.

Vries, Linda Arfelt de, Gaaitzen, and Pieter Woltjer. 2021. "The Economic Transformation Database (ETD): Content, Sources, and Methods." https://doi.org/10.35188/UNU-

WIDER/WTN/2021-2.