

Transportation Infrastructure and Total Factor Productivity: Development Heterogeneity and Resilience under Adverse Shocks

By Hyunseok Kim

May 2026

1. Overview

This document contains the necessary information to replicate the results for the Working Paper “Transportation Infrastructure and Total Factor Productivity: Development Heterogeneity and Resilience under Adverse Shocks.”

2. Software Requirements

The analysis was conducted using Stata 18. Required user-written commands are included locally in the package.

3. Folder Structure

The folders included in this reproducibility package are as follows:

```
Reproducibility_Package/
|
|—— code/
|   |—— ado/
|
|—— data/
|   |—— data_raw/ (original, unmodified data files)
|   |—— data_int/ (intermediate data files)
|   |—— data_final/ (final data file to be used in the analysis)
|
|—— output/
|   |—— figure/
|   |—— table/
```

4. Data Description and Sources

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.

This study assembles an annual unbalanced cross-country panel covering more than 100 countries over the period 2000–2023. Raw data are obtained from original sources and stored in the /data/raw folder. All data cleaning, harmonization of country identifiers, and variable construction are performed in Stata and saved to the /data/int and /data/final folders. Some datasets used in this study are subscription-based and were accessed through the World Bank Library.

4.1. Country Identifiers

- Country codes and names are harmonized using World Development Indicators (WDI) country codes as the baseline. Minor harmonization of country names is performed to ensure consistency across data sources (WDI, IRF, IMF). These modifications affect identifiers only and do not alter underlying data values.
- Source: World Development Indicators (WDI) <https://data.worldbank.org>
- Filename: API_NY.GDP.MKTP.KD.ZG_DS2_en_excel_v2_675.xls
- Access date: April 2026

4.2. Total Factor Productivity and Population

- Total Factor Productivity (TFP) and population data are taken from Penn World Table (PWT) version 11.0. Population data are used to convert transportation infrastructure into per capita terms.
- Source: Penn World Table 11.0 <https://www.rug.nl/ggdc/productivity/pwt/>
- Filename: pwt110.dta
- Access date: November 2025

4.3. Macroeconomic Controls

GDP per capita growth

- Obtained from World Development Indicators (WDI). GDP per capita growth is used to construct adverse macroeconomic episode indicators.
- Source: World Development Indicators (WDI) <https://data.worldbank.org>
- Filename: API_NY.GDP.PCAP.KD.ZG_DS2_en_excel_v2_674.xls
- Access date: April 2026

GDP per capita and Inflation

- GDP per capita and Inflation are obtained from the IMF World Economic Outlook (WEO), where inflation is based on the indexed average consumer price. Initial GDP per capita is defined as the average over 1998–2000.
- Source: IMF World Economic Outlook (WEO), April 2025
<https://www.imf.org/en/publications/weo>
- Filename: weoapr2025all.xls
- Access date: April 2026

Trade openness

- Defined as exports plus imports of goods and services divided by GDP. Based on WDI and included in lagged form in the empirical specifications.
- Source: World Development Indicators (WDI) <https://data.worldbank.org>
- Filename: API_NE.TRD.GNFS.ZS_DS2_en_excel_v2_731.xls
- Access date: April 2026

4.4. Transportation Infrastructure

Road Infrastructure (International Road Federation): 2000-2020

- Road network length is obtained from the International Road Federation (IRF) World Road Statistics (WRS). For the period 2000–2020, IRF WRS data were accessed through a subscription held by the World Bank Library. These data are proprietary and not publicly redistributable.
- Road length is converted into per capita terms (km per million people) and serves as the baseline transportation infrastructure measure in the analysis.
- Source: World Bank Library
- Filename: IRF_WRS_2000_2020.xlsx
- Access date: March 2026

Road Infrastructure (International Road Federation): 2018-2023

- For 2018 onward, updated IRF WRS releases are accessible to registered users, subject to IRF terms of use.
- Source: International Road Federation, World Road Statistics (WRS)
<https://datawarehouse.worldroadstatistics.org/>

The data were obtained by generating country comparison outputs within the WRS platform and downloading the associated datasets.

- Filename: IRF_WRS_2018_2023.csv • Access date: May 2026 **Important Reproducibility Note:**

Raw IRF data files are not included in this replication package due to licensing restrictions. Users with access to IRF WRS data—either through institutional subscriptions (e.g., the World Bank Library) or direct arrangements with IRF—can place the raw files in /data/raw following the filenames specified in the data-import scripts.

Rail Infrastructure

- Rail network length (total route-km) is taken from World Development Indicators (WDI). Rail infrastructure is used in robustness checks, combined with road infrastructure to construct a broader transport infrastructure measure.
- Source: World Development Indicators (WDI) <https://data.worldbank.org>
- Filename: API_9_DS2_en_excel_v2_12617.xls
- Access date: April 2026

4.5. Governance Indicators

Governance quality is measured using the Worldwide Governance Indicators (WGI).

- Six dimensions are used: Government Effectiveness, Regulatory Quality, Rule of Law, Control of Corruption, Voice and Accountability, and Political Stability and Absence of Violence/Terrorism. Initial governance is defined as the average over 1996–2000 to capture predetermined institutional conditions.
- Source: Worldwide Governance Indicators (WGI), 2025 release <https://www.worldbank.org/governance/wgi>
- Filename: wgidataset_with_sourcedata-2025.xlsx
- Access date: April 2026

4.6. Reproducibility Statement

All publicly available data sources are included in the replication package or can be freely downloaded by replicators. Road infrastructure data from the International Road Federation (IRF) are proprietary and are therefore excluded due to licensing restrictions, but their use is fully documented. All data processing, harmonization, and variable construction steps are fully reproducible using the provided Stata scripts, conditional on authorized access to the IRF data.

5. Instructions to Replicate the Results

5.1. Master Script

All results reported in the paper are generated by running the Stata master script.

- 00_master.do

The master script sets project-specific directory paths, loads required user-written packages, and sequentially executes all sub-scripts needed to import data, construct analysis datasets, and generate all figures and tables.

5.2. Execution Order and Code Structure

The master script executes the following Stata do-files in order:

Data import and construction

- 01_import_raw.do
Imports and cleans raw source datasets from publicly available sources, harmonizes country identifiers, and prepares intermediate data files.
- 02_build_database.do
Merges data from different sources and constructs the main analysis database.
- 03_generate_regression_data.do
Generates regression-ready panel datasets used in the empirical analysis.

Figures

- 04_figure_1.do
- 04_figure_2.do
- 04_figure_S1.do • 04_figure_S2.do
- 04_figure_S3.do

Tables

- 05_table_2.do • 05_table_3.do • 05_table_4.do
- 05_table_5.do
- 05_table_6A.do
- 05_table_6B.do
- 05_table_6C.do
- 05_table_S1.do • 05_table_S2.do
- 05_table_S3.do

Table 6 in the paper consists of three panels (Panels A, B, and C), which report distinct sets of implied effects. For transparency and modularity, each panel is generated using a separate Stata script: 05_table_6A.do, 05_table_6B.do, and 05_table_6C.do. Together, these scripts reproduce the complete Table 6 as reported in the paper.

5.3 Output Generation

Running 00_master.do reproduces all figures and tables reported in the paper.

- Figures are saved to:
/output/figure/ •
Tables are saved to:
/output/table/

The output folders are provided empty in the replication package and are populated automatically when the master script is executed.

5.4 Instructions for Replicators To

reproduce the results:

1. Unzip the replication package.
2. Open Stata (version 18 or later).
3. Open 00_master.do.
4. Update the root directory path in the global ROOT definition if necessary.
5. Run the master script.

All results are generated programmatically. No manual intervention is required.

Reproduction of results using IRF road infrastructure data is conditional on authorized access to the International Road Federation (IRF) World Road Statistics (WRS), which are not included in the package due to licensing restrictions.

6. Computational Requirements

The file “master.do” was last run on Windows 11 OS with 32 GB RAM and took approximately 3 minutes to run.