



Infrastructure Foundations: From Current Assets to Future Growth

Third Submission: FR_WLD_2026_601

Maria Reyes Retana, Eric Sabiti, Nihaa Sajid

reproducibility@worldbank.org

June 1, 2026

This review verifies the reproducibility of the exhibits included in the paper *"Infrastructure Foundations: From Current Assets to Future Growth"*.

Contents in this review:

1. Main findings
2. List of exhibits and reproducibility status
3. Reproduction Environment

Main findings

- **Every exhibit of the paper has been reproduced accurately.**
- The code was successfully executed on a new computer after:
 1. Adjusted the root file path to the local machine path in all main script files (.do, .R, and .ipynb).
- The output demonstrates consistent stability across multiple runs. Specifically, executing the code 2 times consecutively yielded identical results.
- The code takes approximately 17 minutes to run.
- We conducted our reproducibility analysis based on the manuscript PDF shared by the authors by email on May 7th.
- **Verification Process and Data Handling:**
 - The reproducibility package relies on 4 types of data: open, accessible (public but not included in the reproducibility package), limited-access (available to World Bank Staff), and restricted data.
 - Open data is included in the public reproducibility package.
 - Reviewers used forthcoming/accessible/restricted data provided directly by the authors to conduct the reproducibility verification, and this is not included in the public reproducibility package.
 - Reviewers verified that the limited-access data matches what was provided by the authors; results of this verification are documented in the file *comparison_report.csv*.
 - *data_hash_report.csv* lists the SHA256 hashes of all files in the Data folder to support data integrity checks. Users who acquire the forthcoming/accessible/limited-access/restricted data can use this file to verify that the data has not been altered.
- **Reproducibility Summary:**

- **Data:** Some data is restricted and not included in the reproducibility package.
- **Code:** All code files (from cleaning to analysis) are included in the reproducibility package.
- **Outputs:** Some outputs are not generated by code (created manually in Excel, ArcGIS, etc), but instructions for producing these outputs are included in the README file.
- **Dependencies environment:** The reviewers created a new environment using the latest versions of dependencies available at the moment of the review.

List of exhibits and reproducibility status

Results in the Main Section of the Paper

1. Executive summary

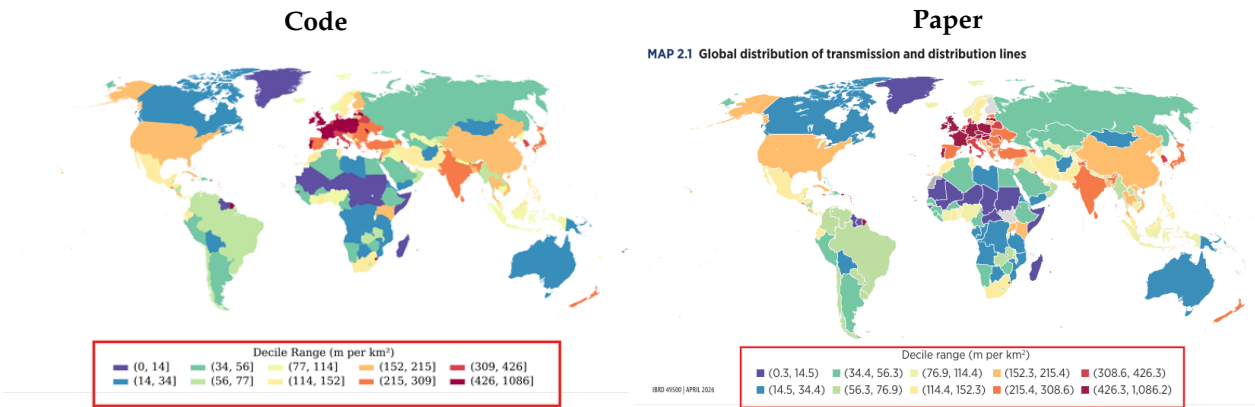
- **Figure ES.1 Reproduced.**
- **Map ES.1 Reproduced.**
- **Map ES.2 Reproduced.**
- **Map ES.3 Reproduced.** Difference in value of median in the legend
- **Figure ES.2 Reproduced.** Identical output to Figure 2.7.
- **Figure ES.3 Reproduced.** Identical output to Figure 2.12.
- **Figure ES.4 Reproduced.** Identical output to Figure 3.3.
- **Figure ES.5 Reproduced.** Identical output to Figure 3.4.
- **Figure ES.6 Reproduced.** Identical output to Figure 3.5.

2. Chapter 1

- **Figure 1.1 Reproduced.**
- **Table 1.1** Does not apply.
- **Table 1.A.1 Reproduced, manual steps.** table generated by running *CompileDataset.ipynb*

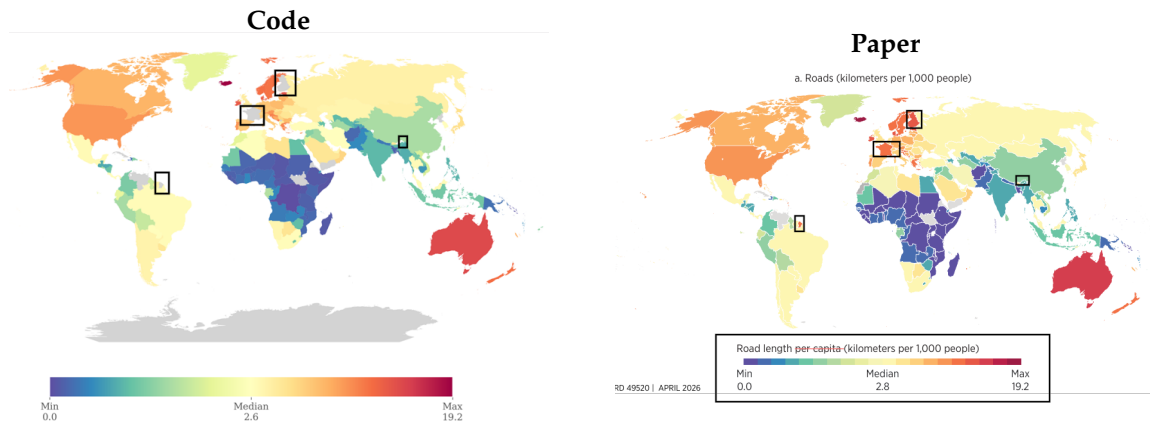
3. Chapter 2

- **Figure 2.1 Reproduced.**
- **Figure 2.2 Reproduced.**
- **Map 2.1 Reproduced, manual steps.** The figure is generated by the code and saved as Figure2.3.png, with underlying data in Figure2.3.csv. The final map was produced by the Bank's Map Unit using this output; visual differences are expected and noted in the README.



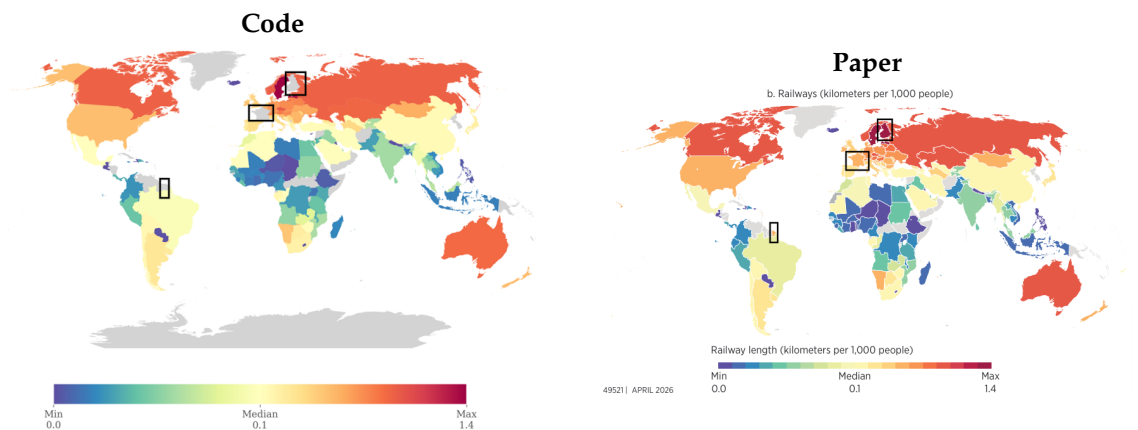
- **Map 2.2a Reproduced, manual steps.**

The figure is generated by the code and saved as Figure2_4a_Roads_per_capita.png, with underlying data in Figure2_4a.csv.csv. The final map was produced by the Bank's Map Unit using this output; visual differences are expected and noted in the README.



- **Map 2.2b Reproduced, manual steps.**

The figure is generated by the code and saved as Figure2_4a_Roads_per_capita.png, with underlying data in Figure2_4a.csv.csv. The final map was produced by the Bank's Map Unit using this output; visual differences are expected and noted in the README.

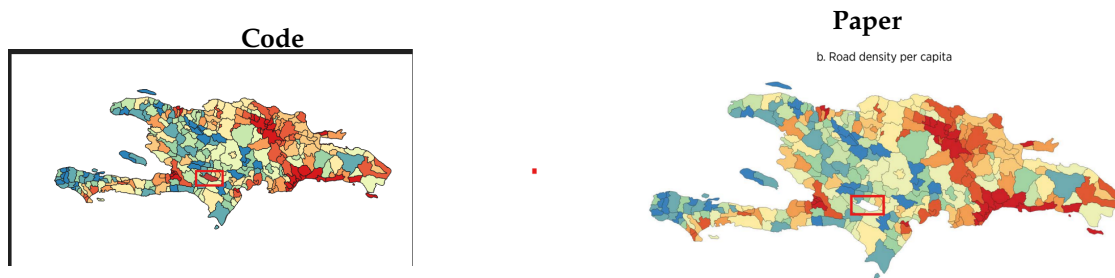


- **Figure 2.3 Reproduced.** This figure relies on restricted data; users won't be able to reproduce it, but the figure created by the replicators using the restricted data is included in the package.
- **Maps 2.3a & 2.3b Reproduced, manual changes** The map was reproduced using a point-and-click approach in QGIS, using the data provided by the authors and following the steps outlined in the README. The map is therefore marked as reproduced. Note that the data are proprietary and stored under NDA, so this map cannot be reproduced by external users.
- **Figure 2.4 Reproduced.**
- **Figure 2.5 Reproduced.**
- **Figure 2.6 Reproduced.**
- **Figure B2.1.1 Reproduced.**
- **Figure 2.7 Reproduced.**
- **Figure 2.8 Reproduced.**
- **Figure 2.9 Reproduced.**
- **Figure 2A.1 Reproduced.**
- **Figure 2A.2 Reproduced.**

4. Chapter 3

- **Figure 3.1 Reproduced.**
- **Figure 3.2 Reproduced.**
- **Figure B3.1.1 Reproduced.**
- **Figure 3.3 Reproduced.**
- **Figure 3.4 Reproduced.**
- **Table 3.1 Reproduced.**
- **Table B3.2.1 Reproduced.**
- **Figure B3.2.1 Reproduced.**
- **Table 3.2 Reproduced, manual changes** The last column in the reproduced output needs to be subtracted from 100 to obtain the values in the manuscript.
- **Figure 3.5a Reproduced.**
- **Figure 3.5b Reproduced.**
- **Figure 3.6 Reproduced.**
- **Figure 3.7 Reproduced.**
- **Table 3.3 Reproduced.** Some rounding differences but that does not impact reproducibility.
- **Table 3A.1 Reproduced.**
- **Figure 3A.1 Reproduced.**
- **Map S3.1.1 (a) Reproduced, manual changes.** Road network data and shapefiles for Hispaniola are provided in Data/FigureData/Figure 3.S1. Final maps are assembled manually in QGIS by loading the provided layers.

- **Map S3.1.1 (b) Reproduced, manual changes.** Minor visual difference. Road network data and shapefiles for Hispaniola are provided in Data/FigureData/Figure 3.S1. Final maps are assembled manually in QGIS by loading the provided layers.



- **Map S3.1.1 (c) Reproduced, manual changes.** Road network data and shapefiles for Hispaniola are provided in Data/FigureData/Figure 3.S1. Final maps are assembled manually in QGIS by loading the provided layers.

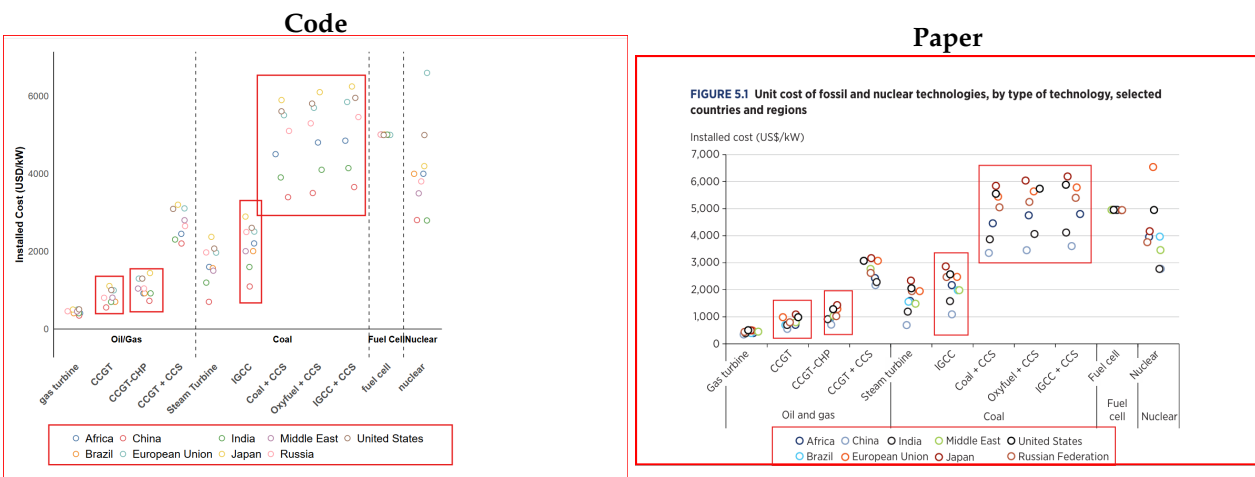
5. Chapter 4

- **Figure 4.1** Does not apply. The exhibit does not show analysis.
- **Figure 4.2** Reproduced.
- **Figure 4.3** Reproduced.
- **Map 4.1** Reproduced. Minor difference in the median value mentioned in the legend.
- **Map 4.2a** Reproduced.
- **Map 4.2b** Reproduced.
- **Map 4.3a** Reproduced.
- **Map 4.3b** Reproduced.
- **Map 4.4a** Reproduced. Minor difference in the median value mentioned in the legend.
- **Map 4.4b** Reproduced. Minor difference in the median value mentioned in the legend.
- **Maps 4.5a & 4.5b Reproduced (manual changes).** Shapefiles for the subnational transport distribution maps are provided in Data/FigureData/Figure 4.8. Final maps are assembled manually in QGIS following instructions in the README file; format differences are expected.
- **Figure 4.4** Does not show analysis results.
- **Figure 4.5** Reproduced. Same as figure 2.3
- **Figure 4.6** Reproduced. Formatting differences.
- **Figure 4.7** Reproduced. Formatting differences.
- **Figure 4.8** Reproduced. Formatting differences.
- **Map 4.6 Reproduced, manual changes** The map was reproduced using a point-and-click approach in QGIS, using the data provided by the authors and following the steps outlined in the README. There are minor formatting differences in styling and symbology, but the underlying data and geographic features match the original. The map is therefore marked as reproduced. Note that the data are proprietary and stored under NDA, so this map cannot be reproduced by external users.

- **Map 4.7 Reproduced, manual changes** The map was reproduced using a point-and-click approach in QGIS, using the data provided by the authors and following the steps outlined in the README. Note that the data are proprietary and stored under NDA, so this map cannot be reproduced by external users.
- **Table 4A.1** Does not show analysis.
- **Table 4B.1** Does not show analysis.

6. Chapter 5

- **Figure 5.1 Reproduced, minor differences.** Minor differences in the positioning of data points on the scatter plot are expected and documented in the README. These arise from the use of `position_jitter()` in R, which introduces randomness in point placement. Since the underlying values are identical, the figure is considered reproducible. This is noted here for transparency.

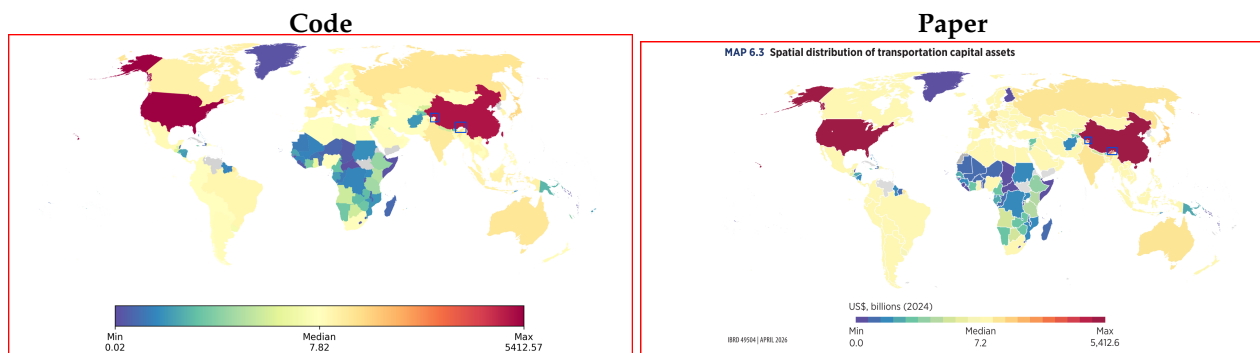


- **Figure 5.2a Reproduced.** Formatting differences
- **Figure 5.2b Reproduced.** Formatting differences
- **Figure 5.3 Reproduced.** Formatting differences
- **Figure 5.4 Reproduced.** Formatting differences
- **Figure 5.5 (a) Reproduced.**
- **Figure 5.5 (b) Reproduced.**
- **Figure 5.6 (a) Reproduced.**
- **Figure 5.6 (b) Reproduced.**
- **Figure 5.7 (a) Reproduced.**
- **Figure 5.7 (b) Reproduced.**
- **Figure 5.8 Reproduced, manual changes.** Results are shown in 5.8standardized_coefficients_road.csv and require manual processing to produce the coefficient plot; after doing so, the figure matches. The underlying data used to generate these results is restricted and shared with the reproducibility team under NDA, so an independent replicator cannot produce the figure directly. The output file generated by the replicators is available in the Outputs folder of the package.

- **Figure 5.9 Reproduced, manual changes.** Results are shown in `5.9standardized_coefficients_rail.csv` and require manual processing to produce the coefficient plot; after doing so, the figure matches. The underlying data used to generate these results is restricted and shared with the reproducibility team under NDA, so an independent replicator cannot produce the figure directly. The output file generated by the replicators is available in the `Outputs` folder of the package.
- **Figure 5.10 Reproduced.** Minor differences in legend colors and scale values, but this does not jeopardize reproducibility.
- **Figure 5.11 Reproduced.** Minor differences in legend colors and scale values, but this does not jeopardize reproducibility.
- **Figure 5.12 Reproduced.** Minor differences in legend colors and scale values, but this does not jeopardize reproducibility.
- **Figure 5.13 Reproduced.** Minor differences in legend colors and scale values, but this does not jeopardize reproducibility.
- **Table 5.A** Does not apply.
- **Table 5.B** Does not apply.
- **Table 5.A2a** Does not apply.

7. Chapter 6

- **Map 6.1 Reproduced (manual changes).**
- **Figure 6.1 Reproduced.** Minor differences in legend colors
- **Figure 6.2 Reproduced.** Minor differences in legend colors and scale values
- **Map 6.2a Reproduced.**
- **Map 6.2b Reproduced.**
- **Map 6.2c Reproduced.**
- **Map 6.2d Reproduced.**
- **Map 6.3 Reproduced, manual steps.** The underlying data (`Chapter6_Figure4.csv`) was used to produce the output by the Bank's Map Unit. Visual differences are expected and noted in the README, documented here for transparency.



- **Figure 6.3 Reproduced.**

- **Figure 6.4 Reproduced.**
- **Maps 6.4a, 6.4b & 6.4c Reproduced, manual steps.** The figure was recreated using QGIS with the information in the README.
- **Map 6.5 Reproduced, manual steps.** The figure was recreated using QGIS with the information in the README.
- **Map 6.6 Reproduced.** Minor differences in legend colors and scale values
- **Figure 6.5 Reproduced.** Minor differences in legend colors
- **Figure 6.6 Reproduced.** Minor differences in legend colors
- **Map 6.7a Reproduced.**
- **Map 6.7b Reproduced.**

Results in the Annex

- **Table 7.1** Does not apply. *Results taken directly from Foster, Gorgulu, Jain, Straub, and Vagliasindi (2025); not computationally reproducible within this package.*
- **Table 7.2** Does not apply. *Results taken directly from Foster, Gorgulu, Jain, Straub, and Vagliasindi (2025); not computationally reproducible within this package.*
- **Table 7.3** Does not apply. *Results taken directly from Foster, Gorgulu, Jain, Straub, and Vagliasindi (2025); not computationally reproducible within this package.*
- **Figure S7.1.1** Does not apply. *Results from Gerrits et al. (2025); not computationally reproducible within this package*
- **Map S7.1.1** Does not apply. *Results from Gerrits et al. (2025); not computationally reproducible within this package.*
- **Figure S7.1.2** Does not apply. *Results from Gerrits et al. (2025); not computationally reproducible within this package*
- **Table S7.1.1** Does not apply. *Results from Gerrits et al. (2025); not computationally reproducible within this package*
- **Figure S7.2.1 Reproduced.** Minor differences in legend colors
- **Figure S7.2.2 Reproduced.**

Reproduction Environment

Paper exhibits were reproduced on a computer with the following specifications:

- OS: Windows 11 Enterprise (Build 10.0.22631)
- Processor: Intel(R) Xeon(R) Platinum 8562Y+
- Memory available: ~128 GB
- Software version: Stata 19 MP; R 4.4.0; Python 3.9