



Beyond The Urban Sweet Spot: Firm-Level Evidence Of Over-Agglomeration In Mongolia's Capital

First Submission: RR_MNG_2026_637

Camila Ayala

reproducibility@worldbank.org

May 27, 2026

This review verifies the reproducibility of the exhibits included in the paper "*Beyond The Urban Sweet Spot: Firm-Level Evidence Of Over-Agglomeration In Mongolia's Capital*".

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. Opening the R project "Replication_code_Agglomeration_WP.Rproj"
 2. Restoring the environment by running `renv::restore()` and following the prompts.
 3. Opening the `replication_agglomeration.R` and running the code.
- The output demonstrates consistent stability across multiple runs. Specifically, executing the code two times consecutively yielded identical results.
- The code takes approximately 5 minutes.
- We conducted our reproducibility analysis based on the paper shared by the authors by email on May 14th.
- **Verification Process and Data Handling:**
 - The reproducibility package relies on three types of data: open, accessible and restricted data.
 - Open data is included in the public reproducibility package.
 - Reviewers used restricted data provided directly by the authors to conduct the reproducibility verification, and this is not included in the public reproducibility package.
 - `data_hash_report.csv` lists the SHA256 hashes of all files in the Data folder to support data integrity checks. Users who acquire the restricted data can use this file to verify that the data has not been altered.

- **Reproducibility Summary:**

- **Data:** Some data is restricted and has not been included in the reproducibility package. For more details, please refer to the README file.
- **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) 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

- **Figure 1 Reproduced.** Verified against Figure1.xlsx. The figure was created in Excel; we downloaded the data, replicated the calculations, and successfully reproduced the figure.
- **Figure 2A Reproduced.** Verified against Figure2A.png.
- **Figure 2B Reproduced.** Verified against Figure2B.png.
- **Figure 3 Reproduced.** Verified against Figure3.png.
- **Figure 4 Reproduced.** Verified against Figure4.png.
- **Table 1 Reproduced.** Verified against table1.pdf.
- **Table 2 Reproduced.** Verified against table2.pdf. The figures in the code output are exact, while those reported in the paper are rounded to millions; the underlying values are consistent across both.
- **Table 3 Reproduced.** Verified against table3.txt.
- **Table 4 Reproduced.** Verified against table4.txt.

Results in the Annex

- **Table A1 Reproduced.** Verified against tableA1.txt.
- **Table A2 Reproduced.** Verified against tableA2.txt.
- **Table A3 Reproduced.** Verified against tableA3.txt.
- **Table A4 Reproduced.** Verified against tableA4.txt.

Reproduction Environment

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

- OS: Windows 11 Enterprise
- Processor: Intel(R) Xeon(R) Gold 6226R CPU @ 2.90GHz
- Memory available: 32.0 GB
- Software version: Stata 19.5 MP