



Closing Data Gaps in Fertilizer Subsidy Analysis in Bangladesh: A Survey-to-Survey Imputation Approach

First Submission: RR_BGD_2026_620

Innocent Bushayija Sano, Eric Sabiti

reproducibility@worldbank.org

April 22, 2026

This review verifies the reproducibility of the exhibits included in the paper “Closing Data Gaps in Fertilizer Subsidy Analysis in Bangladesh: A Survey-to-Survey Imputation Approach”.

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 stata file "oo_master_stage1_harmonization.do" updating the directory and running the code
 2. opening the R project "fertilizer_s2s_reproducibility.Rproj"
 3. Restoring the environment by running `renv::restore()`
 4. opening the R script "03_figure1_kernel_density.R", updating directory, and running the script
 5. opening the R script "05_s2s_imputation.R", updating directory and running the script
 6. opening stata file "oo_master_stage3.do" updating the directory, installing the required packages 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 30 minutes to run.
- We conducted our reproducibility analysis based on the paper shared by the authors
- **Verification Process and Data Handling:**
 - The reproducibility package relies on 3 types of data: open, limited-access and restricted data (available only from the authors).
 - 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 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 has not been included in the reproducibility package. For more details, refer to the README file.
 - **Code:** All code files (from cleaning to analysis) are included in the reproducibility package.
 - **Outputs:** All outputs are generated by code included in the reproducibility package.
 - **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

- **Table 1** Reproduced.
- **Table 2** Reproduced.
- **Table 3** Reproduced.
- **Figure 1** Reproduced.
- **Figure 2** Does not apply. Non-analytical output
- **Figure 3** Reproduced. Figure 3 was manually prepared in excel
- **Figure 4** Reproduced. Figure 4 was manually prepared in excel
- **Figure 5** Reproduced. Figure 5 was manually prepared in excel
- **Figure 6** Reproduced.
- **Figure 7** Reproduced.
- **Figure 8** Reproduced.
- **Figure 9** Reproduced.
- **Figure 10** Reproduced.
- **Figure 11** Reproduced. Figure 11 was manually prepared in excel
- **Figure 12** Reproduced. Figure 12 was manually prepared in excel
- **Figure 13** Reproduced. Figure 13 was manually prepared in excel

Reproduction Environment

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

- OS: Windows 11 Enterprise
- Processor: Intel(R) Core(TM) i5-1145G7 CPU @ 2.60GHz
- Memory available: 15.7 GB
- Software version: Stata 18.0 MP