

When Aggregation Misleads: Bias in Unit-level Small Area Estimates of Poverty with Aggregate Data

Author: Paul Corral Rodas

Institution: World Bank – Poverty and Equity Global Practice

Contact: pcorralrodas@worldbank.org

Overview

This repository contains Stata replication code for the simulations and figures presented in the paper:

Corral Rodas, P. A. (2025). *When Aggregation Misleads: Bias in Unit-level Small Area Estimates of Poverty with Aggregate Data*. The World Bank, Poverty and Equity Global Practice.

The paper investigates how and why **unit-context models**, which rely only on **aggregate auxiliary data** to model for household level welfare, produce **biased poverty and welfare estimates**. The analysis demonstrates that the bias is driven by these models' failure to approximate **within-area welfare variance across households**, which is critical for distribution-sensitive indicators like poverty.

Contents

- [source of bias.do](#)
Main Stata simulation script. Generates synthetic populations, applies unit-level and unit-context models, and calculates poverty measures under different thresholds.
 - [source of bias figures.do](#)
Script to generate the key figures from the paper (e.g., bias across percentiles, bias vs. variance ratios, etc.). Figures, however, are produced in Tableau.
 - [misleading_aggregation.pdf](#)
Full paper outlining the theory, methods, and findings.
-

Requirements

To run the simulations and figures:

- **Stata 17 or higher** (MP or SE recommended due to simulation volume)
 - Adequate memory (100k+ obs and 1,000+ simulation runs)
 - Scripts are self-contained and do not require external datasets
-

Running the Code

1. Generate Simulation Results

Open and run:

```
do "1.source of bias.do"
```

This will:

- Simulate 1,000 populations of 500,000 households across 100 areas
- Apply unit-level and unit-context models
- Compute area-level poverty under 99 percentiles
- Export summary matrices to disk

2. Generate Figures

Once the simulation is complete, run:

```
do "2.source of bias figures.do"
```

This will:

- Load saved matrices
- Reconstruct key tables from the paper
- Output data to Excel for figure production in **Tableau**

Citation

If using this code or adapting the methodology, please cite:

```
Corral Rodas, P. A. (2025). When Aggregation Misleads: Bias in Unit-level Small Area Estimates of Poverty with Aggregate Data. The World Bank, Poverty and Equity Global Practice.
```

APA:

```
Corral Rodas, P. A. (2025). When aggregation misleads: Bias in unit-level small area estimates of poverty with aggregate data. The World Bank, Poverty and Equity Global Practice.
```

License

This code is shared under the MIT License. Please acknowledge the author when reusing or modifying.