

SDI Caseloads

Analysis of healthcare provider caseloads using Service Delivery Indicators (SDI) data from 10 African countries.

Project Structure

```
sdi-caseloads/
├── ado/                    # User-written Stata packages
├── constructed/           # Intermediate analysis datasets
├── data/                  # De-identified country data files
├── do/                    # Stata do-files
│   ├── 0-sourcedata.do    # Create de-identified data from Box (requires PII
│   │                       access)
│   ├── 1-makedata.do      # Construct analysis variables
│   ├── 2-simulations.do   # Queue simulations and optimization calculations
│   ├── 3-tables.do        # Generate main text tables
│   ├── 4-figures.do       # Generate main text figures
│   └── 5-appendix.do      # Generate appendix tables and figures
├── outputs/
│   ├── main/              # Main text outputs (tables, figures)
│   ├── appendix/          # Appendix outputs
│   └── temp/              # Temporary graph files
└── runfile.do             # Master runfile
```

Data and Licensing

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. Appropriate permission are documented in the LICENSE.txt file.

This replication package includes de-identified country files in `data/`. Each file is derived from SDI source data and harmonized for this analysis.

General SDI collection page (WBG Microdata Catalogue): <https://microdata.worldbank.org/catalog/sdi>

Dataset inventory (analysis files, survey-year coverage, and access route):

Analysis file	Dataset name (country, year used in analysis)	Approx. date accessed	Public URL / access route
data/kenya.dta	SDI Health Survey - Kenya (2018)	Jun 2026	https://microdata.worldbank.org/catalog/3872
data/madagascar.dta	SDI Survey - Madagascar (2016)	Jun 2026	https://microdata.worldbank.org/catalog/3884
data/malawi.dta	SDI Survey - Malawi (2019)	Jun 2026	Restricted (not publicly posted in SDI catalog)

Analysis file	Dataset name (country, year used in analysis)	Approx. date accessed	Public URL / access route
			as of access date). See restricted-access instructions below.
data/mozambique.dta	SDI Health Survey - Mozambique (2014)	Jun 2026	https://microdata.worldbank.org/catalog/3876
data/niger.dta	SDI Health Survey - Niger (2015)	Jun 2026	https://microdata.worldbank.org/catalog/3222
data/nigeria.dta	SDI Health Survey - Nigeria (2013)	Jun 2026	https://microdata.worldbank.org/catalog/2559
data/sierra_leone.dta	SDI Health Survey - Sierra Leone (2018)	Jun 2026	https://microdata.worldbank.org/catalog/4038
data/tanzania.dta	SDI Health Survey - Tanzania (2016)	Jun 2026	https://microdata.worldbank.org/catalog/3873
data/togo.dta	SDI Survey - Togo (2012)	Jun 2026	https://microdata.worldbank.org/catalog/2560
data/uganda.dta	SDI Health Survey - Uganda (2013)	Jun 2026	https://microdata.worldbank.org/catalog/2750

Notes:

- Survey years above are the years present in the harmonized analysis files.
- Public catalog records may label years differently in some cases (for example, fieldwork windows shown as 2013-2014).
- Malawi SDI details can be found here: Malawi - Harmonized Health Facility Assessment : 2018-2019 Report (Vol. 1 of 4) : Main Report (English). Washington, D.C. : World Bank Group. <http://documents.worldbank.org/curated/en/417871611550272923>

Restricted-data access and verification:

- The script `do/0-sourcedata.do` builds de-identified country files from restricted source files stored in a controlled WBG Box archive (`.../pii/SDI/SDI/raw/endline/DataSets/Intermediate`).
- These restricted source files are not redistributed in this reproducibility package.
- For independent verification, restricted datasets were shared with the WBG Reproducibility team under internal access controls.
- Potential replicators seeking restricted data should request access through the WBG SDI/Microdata process, including project purpose and institutional affiliation, and comply with country-specific licensing terms before access is granted.

For countries with publicly available microdata, replication results may differ slightly because anonymization protocols can trim distributions, convert some continuous variables into categorical

values, and remove outliers to ensure data privacy.

Running the Code

1. Open `runfile.do` and set paths if needed
2. Toggle which do-files to run (0/1 flags)
3. Run `runfile.do` in Stata

Note: `2-simulations.do` is slow (queue simulations and bootstrap). Pre-computed results are saved in `constructed/capacity-optimized.dta`.

Outputs

Main Tables

- `t1-summary.xlsx` - Per-provider outpatient daily caseload summary statistics by country
- `t2-regressions.xlsx` - Regression results for caseloads, facility characteristics, and provider competence
- `t3-optimization.xlsx` - Simulation: Impacts of optimal provider allocation on quality of care

Main Figures

- `f1-summary-capacity.png` - Daily outpatient caseloads faced by providers and patients, by country
- `f2-queue-simulations.png` - Variation in provider workload and patient waits in queuing simulations with varying caseloads [Note: requires running `2-simulations.do`]
- `f3-distribution-pat.png` - Share of patients seen by most busy and most competent providers, by country
- `f4-optimization.png` - Correlations between provider competence and caseloads, as observed and optimally allocated

Appendix Tables

- `at2.tex` - Caseload summary statistics by country, excluding large facilities
- `at3.tex` - Regression results for caseloads, excluding large facilities
- `at4.tex` - Reallocation results, excluding large facilities

Appendix Figures

- `a2-facility-capacity.png` - Providers serving outpatients per facility, on roster and sampled at facility
- `a3-caseload.png` - Daily outpatient caseloads per facility, by country and sector

- a4-doctors.png - Outpatient caseloads per provider, adjusted for doctor presence
- a5-queue-clumping.png - Demonstrative queuing patterns over the course of a 30-patient, six-hour working day
- a6-docs-upskill.png - Quality improvement by training non-doctors to doctor equivalence
- a7-optimization-pat.png - Competence per patient, observed and optimally allocated
- a8-optimization-change.png - Improvements in provider competence per patient after optimal reallocation, by country