# FSSIM–Dev Data

Marco Rogna

European Commission, Joint Research Centre (JRC-Seville), Calle Inca Garcilaso 3, 41092, Sevilla.

Pan–African Network for Economic Analysis of Policies:
Accra 14–16 September 2022

# FSSIM–Dev in a Nutshell

- FSSIM–Dev = **Farming System Simulator for Developing Countries**
- **Farm household model** for use in the context of developing countries
- Attempt to reproduce as faithfully as possible (PMP) **farming conditions** and **farm choices** of rural households
- "Shocks": **Variations of economic conditions** and/or different **agri–food policy** options
- Results (of policies or of changed economic conditions) evaluated on the light of:
  - **Food security**
  - **Poverty alleviation**
  - **Productivity changes**

# FSSIM–Dev Data Requirements I

- Three macro–types of variables:
  - **Agriculture**
  - **Livestock**
  - **Exogenous**
- Each set of variables can be further subdivided:
  - Agriculture and Livestock: **Quantities**, **Prices** and **Characteristics**
    * Quantities and Prices further subdivided in **inputs** and **outputs**
- Exogenous variables: Household composition, non–farming income, non–farming activities and labour (used to compute availability of farming labour)

# FSSIM–Dev Data Requirements II

| **Agriculture** | | **Livestock** | |
|---|---|---|---|
| ***Input*** | ***Output*** | ***Input*** | ***Ouput*** |
| Labour | Crop Yield | N. Livestocks owned | N. Livestocks sold |
| Inorganic Fertilizers | | N. Livestocks risen | Milk |
| Organic Fertilizers | | N. Livestocks bought | Milk derivatives |
| Agrochemicals | | Feed | Meat |
| Seeds | | Water | Eggs |
| Land | | Anti–Parasite treatment | |
| | | Other care | |

# FSSIM-Dev Agricultural Variables

FSSIM-Dev Main Agricultural Variables

| | |
|---|---|
| Plant (HH wd/ha) | Pesticide (kg/ha) |
| Growth (HH wd/ha) | Herbicide (kg/ha) |
| Harvest (HH wd/ha) | Fungicide (kg/ha) |
| Plant (Hired wd/ha) | Other Phyto. (kg/ha) |
| Growth (Hired wd/ha) | Org. fert. [HH waste (kg/ha) |
| Harvest (Hired wd/ha) | Org fert [Animal waste] (kg/ha) |
| Urea (kg/ha) | Manure (kg/ha) |
| NPK (kg/ha) | Yield (kg/ha) |
| DAP (kg/ha) | Crop sold (kg/ha) |
| Phosphate rock (kg/ha) | Crop self cons. (kg/ha) |
| Seeds (kg/ha) | Crop area |

# Activities and Characteristic Variables

**Activity**: A unique pair of **crop type** and **methods of cultivation/characteristics** inside each farm with a specific **gross margin**

Variables Defining Methods of Cultivation/Characteristics

| Variable | Values | Variable | Values |
|---|---|---|---|
| Irrigation | Yes | Cropping | Single |
| | No | | Multicropped |
| Soil type | Sandy | Topology | Flat |
| | Silty | | Valley |
| | Glacis | | Hill |
| | Clay | | Mild slope |
| Preparation | No labour | | Stiff slope |
| | Manual | Soil quality | Poor |
| | Animal | | Average |
| | Mechanical | | Good |

# Modularity of Activities

- Type of crops (approx 40 crops present) + 6 Methods of cultivation/Characteristics = **7 sets to define Activities**
- Approximately **38400 Activities**!!! Not always interesting to have such a refined level of detail
- Eliminating one (or more) Methods of cultivation/Characteristics requires to re–aggregate inputs and outputs quantities and prices
- Automatic process to perform this task: Selection of Methods of Cultivation/Characteristics to be kept and, accordingly, re–aggregation of the whole dataset

# EHCVM Data

- **EHCVM:** Enquête Harmonisée sur le Conditions de Vie des Ménages
- **Covered Countries:** 8 Western African countries (CFA adopters): Benin, Burkina Faso, Guinea Bissau, Ivory Coast, Mali, Niger, Senegal, Togo
- **Covered Period:** 2018/19, no panel, single agricultural season
- **Covered Topics:** Life conditions of rural communities with emphasis on economic attainments and on agricultural / husbandry production
- **Covered Sample:** Representative at national level
- **FSSIM-Dev Focus:**
  - **section 16**: Agricultural production
    1. Subsection 16_a: Infos at plot level
    2. Subsection 16_b: Inputs expenditures
    3. Subsection 16_c: Output
  - **Section 17:** Husbandry

# LSMS–ISA Data

- LSMS–ISA: Living Standard Measurement Study–Integrated Surveys on Agriculture
- **Covered Countries:** 8 African countries: Burkina Faso, Ethiopia, Malawi, Mali, Niger, Nigeria, Tanzania, Uganda
- **Covered Period:** From 1 to 4 waves (panel data) depending on country, with first wave in 2008/09 (Tanzania). Single or two agricultural seasons depending on country
- **Covered Topics:** Life conditions of rural communities with emphasis on economic attainments and on agricultural / husbandry production
- **Covered Sample:** Representative at national level
- FSSIM–Dev was using LSMS–ISA (2010/11 wave). Idea is to incorporate all countries missing from EHCVM: Ethiopia, Malawi, Nigeria, Tanzania and Uganda

# Ideal Data for FSSIM–Dev

- LSMS–ISA and EHCVM are the natural candidate datasets to form the database for FSSIM–Dev:
  1. Household level data
  2. Detailed info on agriculture and husbandry (inputs–outputs)
  3. Nationally representative
- First two are necessary requirements:
  - Aggregate data $\rightarrow$ **CGE or agricultural PE**
  - No info on agriculture and livestock $\rightarrow$ **no agricultural model**
- National representativeness not strictly necessary, but useful to evaluate **policies at national level**

# First Phase: Manipulation of Each Country Dataset Separately

- Remapping all variables names from the original code to a name reminding the variable content
- Computation of total hired and HH working days in different agricultural activities (planting, growth phase, harvesting) and remapping booleans
- Conversion of all quantities into Kilograms or Litres (assumed identical) and computation of expected production quantities, consumed quantities, expenditures, prices, etc.

# Conversion of Units of Measure

- Main files used are the conversion tables for consumption products released by each country statistical office
- Consumption products are different from fertilizers and pesticides
- Use of common sense to choose the appropriate conversion item. Use the same for all farmers in the same country
- For missing units of measure:
  - **cv = mean_price(local unit) / mean_price(Kg)** (if price available);
  - **cv = mean_per_ha_q(local unit) / mean_per_ha_q(Kg)** (if price not available);

# From Plots to Activities

- **FSSIM-Dev has activities as basic units of analysis**.
- Household surveys have **plots** as basic unit
- A plot may have different crops (different activities) on it: e.g a portion of the plot is irrigated and another no
- **Separation of each single activity** with its related input and output quantities (and prices)

# Outliers Detection and Re–imputation

- Detection through **interquartile range** setting the **multiplier equal to 1.2** (stricter than the conventional 1.5)
- Quantities (per–hectare) checked only for up tail outliers, prices and output (per–hectare) for both tails
- Per–hectare quantities are country–crop specific whereas for input prices only countries are considered
- Outliers are substituted with national medians
- **Boxplots** for all variables before the outliers detection, after their elimination and after the re–imputation of median values are produced for each separate country

# Imputation of Missing Prices

- For inputs with zero quantities, prices are or zero or NA. Since FSSIM–Dev requires prices, **crop–livestock medians** are taken
- FSSIM–DEV is built on the idea that farmers base their decisions on expected yields. Outputs quantities are checked for lower–tail outliers (interquartile range and manual bounds) and re–imputed with median values
- Medians are taken at:
  - Regional level
  - Country level
  - Whole dataset level but crop–livestock specific
  - Country level over all crops (not for livestock)

# Final Checks and Negative Profits Count

- Drop of unused variables by FSSIM–Dev
- Check for negative quantities and prices

Number of Activities with Negative Profits

| | ***Agricultural and Livestock Data*** |
|---|---|
| **No labour costs** | 0 / 82801 |
| **No own labour costs** | 2621 / 82801 |
| **With all labour costs** | 16723 / 82801 |
| **Husbandry - No labour** | 0 / 57435 |

# Livestock Data

- Procedures for data cleaning almost identical
- Simpler procedure since **activities coincide with animal types**
- **No data on labour** spent on rearing
- Addition of **difference between average purchasing and selling price as hypothetical gain** in value over one year

# Exogenous Variables I

- Exogenous variables are fundamental to **compute income** (poverty analysis, evaluation of redistributive properties of a policy, etc.) and to compute **availability of work from household members**
- Non–farming income and non–farming hours assumed as given
- Household farm labour = sum, for all HH members, of **maximal daily working time** per HH member minus **exogenous time requirements** per HH member

# Exogenous Variables II

FSSIM–Dev Households Variables

| | |
|---|---|
| N. Members | Weekly hours house activities |
| N. Children | Tot. hours primary job |
| N. young women | Tot. hours secondary job |
| N. adult women | Tot. income primary job |
| N. young men | Tot. income secondary job |
| N. adult men | Average hourly wage |
| N. elderly | Tot. hours self employment |
| School exp. | Tot. profit self employment |
| Other Income | |

# Farms Types

- Farms categories are optional requirements to facilitate the **assignation of parameter values** and for the **analysis of output**
- Elasticities, risk aversion, inputs substitutability, etc. may vary according to farm wealth and/or primary focus of farm
- Categories are also useful for analysing policies since they can be used to individuate groups of gainers and losers
- Three types of categories: **Agronomic/Livestock**, **Size**, **Income**

# Categories

Agricultural and Livestock Categories

| | | | | |
|---|---|---|---|---|
| | ***Cereals*** | ***Rice*** | ***Oleo–proteinous*** | ***Tuber*** |
| **Agricultural** | 6813 | 2874 | 2383 | 1931 |
| **Categories** | ***Vegetables*** | ***Industrial*** | ***Tree*** | ***Generic*** |
| | 860 | 1129 | 4785 | 3443 |
| **Husbandry Categories** | ***Dairy*** | ***Meat (Cattle)*** | ***Meat (Large)*** | ***Eggs and small animals*** |
| | 381 | 2846 | 1271 | 3975 |
| | **Total Agricultural HH** | | **Total Livestock HH** | |
| | 24218 | | 8437 | |

# What is FSSIM–Dev Useful for?

- **Ex–ante evaluation of agricultural policies**:
  1. Subsidies to agricultural inputs
  2. Minimal prices for agricultural output (e.g. Sustainable Cocoa initiative)
  3. Public works such as irrigation infrastructures
  4. Other quantitative measures (quotas, changes in agricultural practices, premiums for environmental friendly practices, etc.)
- **Evaluation of effects of economic shocks**:
  1. Prices increase or decrease (e.g. Inputs and outputs prices peaks due to Ukrainian war)
  2. Productivity losses due to extreme climatic events
  3. Introduction of a new technology

# What is Evaluated?

- Common metrics are used to evaluate policies and economic shocks:
  1. Variations in **income and poverty rates** for rural households
  2. Variations in agricultural output and consequent effects of **food security**
  3. Variations in **inputs and outputs quantities**
  4. Variations of **land allocation** between different crops
  5. Individuation of **gainers and losers** between different categories (income, farm size, farm types, etc.)
  6. Individuation of **spatial effects**

# FSSIM-Dev vs. CGE and PE Models I

## Similarities

- CGE and agricultural PE models have similar objectives: to evaluate policies and economic shocks
- Evaluation is performed on similar metrics: income changes, variations in production and consumption, etc.

## Differences

- Micro vs Macro focus
- CGE and PE models have a better ability to look at the propagation of a shock to different sectors and find the final equilibrium after all sectoral adjustments
- FSSIM–Dev looks deeper into farming households

# FSSIM–Dev vs. CGE and PE Models II

**Complementarities**

- The combined use of the two instruments offers the most complete picture
- CGE and PE models used to look at final equilibrium after a shock (final stable prices)
- FSSIM–Dev takes the final prices and looks deeper at the effects on the rural population
- Macro–adjustments and detailed effects are both captured

# Examples from Previous Studies I

- Sierra Leone (2009 – Bombali region): Evaluation of fertilizer subsidies and changes in agricultural practices (sowing dates and seeds amounts)
  - Significant productivity increases in rice cultivation for both interventions
  - Modest effects on poverty alleviation, few households are lifted out from poverty
  - Limit subsidies to farmers with 5 ha of land or less gives the highest benefit–cost ratio
- Niger (2011/12): Evaluation of small–scale irrigation schemes
  - Farm income increases of approximately 7%
  - Highest benefits for households with lowest income
  - Strong regional heterogeneity in farmers benefits (highest gains for dry regions)

# Examples from Previous Studies II

- Tanzania (2012/13): Reduction or elimination of agricultural produce tax
  - Boost in income between 2% and 21% and in output
  - Strong regional heterogeneity in benefits (Northern and Western highlands gain more)
  - Largest benefits for medium–large farms and households specialized in cash crops
  - Scarce effects on rural poverty
  - A uniform production tax at 1% for all crops has the best benefit–cost ratio
- Ethiopia (2013/14): Upscaling the Agricultural Commercialization Cluster (ACC) initiative
  - Increase in crops output between 1.8% to 62.6% with strong regional heterogeneity
  - Average gross income increase between 9% to 14%
  - Medium–large farms and households specialized in field crops are the most benefited

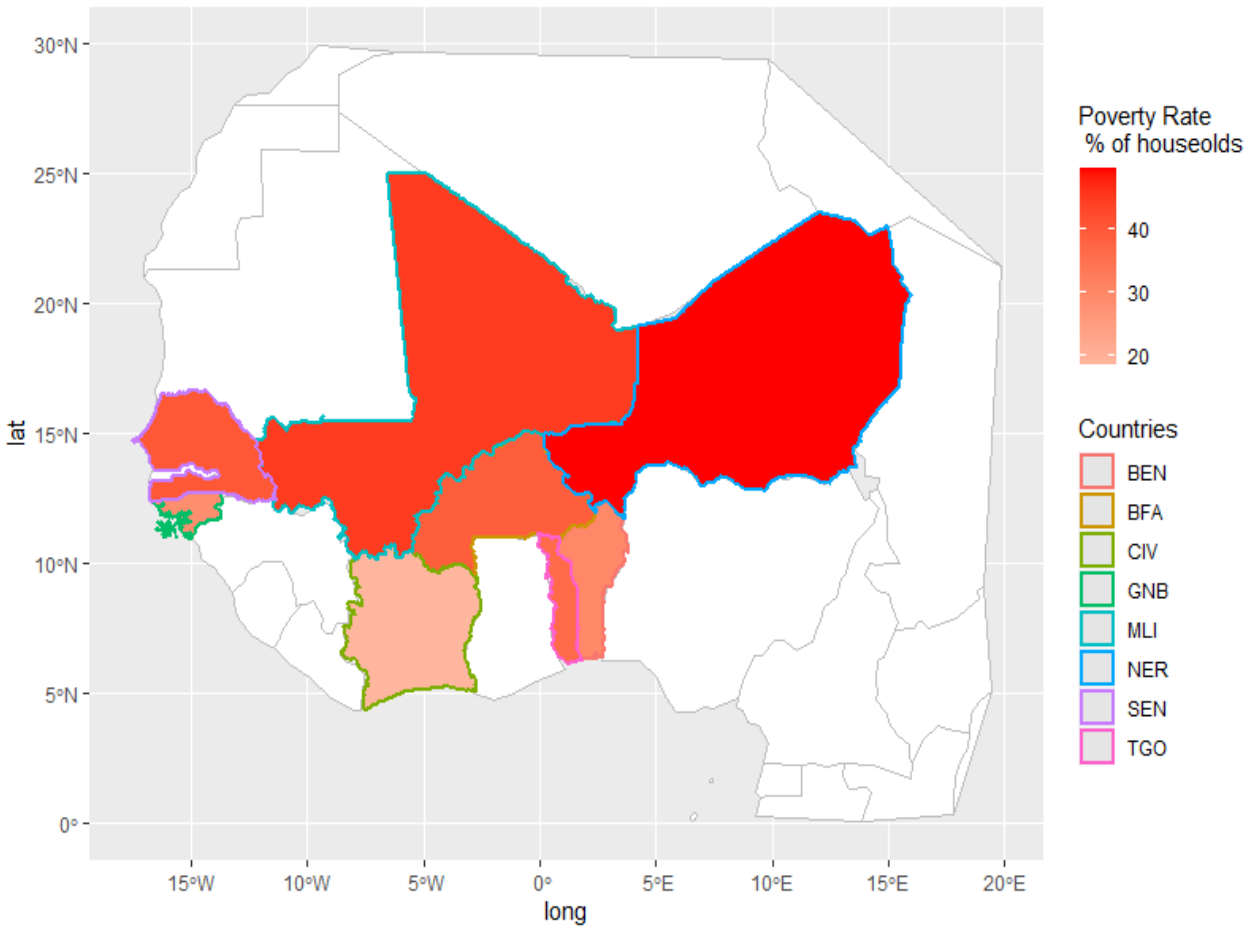
# Some Examples of Output I

Percentage of cultivated land dedicated to each crop in each country

| | BFA | CIV | GNB | MLI | SEN | TGO | BEN | NER |
|---|---|---|---|---|---|---|---|---|
| AUB | 0.04 | 0.39 | 0.08 | 0.03 | 8.04 | 0.03 | 0 | 0 |
| CAB | 0.41 | 4.57 | 0 | 0.27 | 0.47 | 0.07 | 0 | 0.12 |
| CAC | 0 | 10.82 | 0 | 0 | 0 | 0.02 | 0 | 0 |
| CAL | 0 | 0 | 0.01 | 0.09 | 0 | 0 | 0 | 0 |
| CAR | 0.03 | 0 | 0 | 0 | 0.26 | 0 | 0 | 0 |
| CAS | 0.13 | 4.18 | 1.67 | 0.14 | 1.18 | 4.66 | 7.7 | 0.23 |
| CIT | 0 | 0 | 0.13 | 0 | 0 | 0 | 0.25 | 0.21 |
| COF | 0 | 1.55 | 0 | 0 | 0 | 0.05 | 0 | 0 |
| COT | 5.35 | 9.67 | 0 | 13.69 | 0.81 | 7.96 | 9.64 | 0 |
| COW | 6.45 | 0.03 | 1.41 | 0.49 | 1.83 | 1.21 | 0.67 | 16.41 |
| CSH | 0 | 14 | 57.45 | 0.06 | 0 | 0 | 0 | 0 |
| CUC | 0.02 | 0.03 | 0.21 | 0.1 | 0.04 | 0 | 0 | 0 |
| FON | 0.21 | 0.02 | 0 | 0.15 | 0.05 | 0.07 | 0.02 | 0.25 |
| GIN | 0.01 | 0.09 | 0 | 0 | 0 | 0.05 | 0 | 0 |
| GRE | 0.03 | 0.12 | 0 | 0.05 | 0.01 | 0.06 | 0.09 | 0.09 |
| JAX | 0.02 | 0 | 0.27 | 0.03 | 0.2 | 0 | 0 | 0 |
| LET | 0.05 | 0 | 0 | 0.02 | 0.01 | 0.01 | 0.01 | 0.23 |
| MAI | 24.42 | 13.69 | 0.85 | 19.25 | 7.82 | 16.09 | 33.9 | 0.38 |
| MAN | 0.04 | 0.11 | 0.03 | 0.03 | 0 | 0 | 0 | 0.18 |
| MEL | 0 | 0 | 0 | 0.01 | 0.2 | 0 | 0 | 0.02 |
| MIL | 16.04 | 0.27 | 1.1 | 20.02 | 15.97 | 0.92 | 0.55 | 45.11 |
| MPA | 0 | 0 | 2.16 | 0 | 0 | 60.56 | 29.4 | 3.83 |
| NUT | 0.01 | 0 | 0 | 0.03 | 0.02 | 0.01 | 0 | 1.88 |
| OIL | 0 | 5.57 | 0 | 0 | 0 | 0.53 | 0.75 | 0 |
| OKR | 0.52 | 0.49 | 0.63 | 0.38 | 0.3 | 0.16 | 0.37 | 0.5 |
| ONI | 5.8 | 0.01 | 0.1 | 2.45 | 1.51 | 0 | 0.02 | 1.69 |
| PAD | 4.55 | 11.42 | 15.15 | 24.6 | 13.75 | 1.94 | 1.94 | 3.15 |
| PEA | 0.05 | 0 | 0 | 0 | 0 | 0.01 | 0 | 0 |
| PEN | 6.38 | 3.38 | 13.9 | 5.98 | 39.38 | 1.07 | 2.12 | 10.71 |
| PEP | 0.03 | 0 | 0.01 | 0 | 0.16 | 0 | 0 | 0.1 |
| PER | 0.02 | 0.31 | 0.3 | 0.02 | 0.2 | 0.69 | 2.91 | 0.16 |
| POT | 0.12 | 0.01 | 0 | 1.67 | 0.52 | 0 | 0 | 0 |
| RUB | 0 | 6.43 | 0 | 0 | 0 | 0 | 0 | 0 |
| SES | 1.94 | 0.01 | 0.13 | 0.22 | 0.23 | 0.27 | 0.01 | 1.37 |
| SOL | 0.27 | 0.03 | 0.3 | 0.07 | 1.21 | 0.01 | 0 | 0.51 |
| SOR | 25.14 | 0.31 | 1.8 | 8.98 | 1.73 | 2.66 | 2.88 | 9.74 |
| SQU | 0.01 | 0 | 0.01 | 0.02 | 0.01 | 0 | 0 | 0.01 |
| SWP | 1.12 | 0.25 | 1.24 | 0.6 | 0.36 | 0.02 | 0.07 | 0.08 |
| TAR | 0 | 0.06 | 0.22 | 0 | 0.02 | 0.05 | 0 | 0 |
| TOM | 0.21 | 0.1 | 0.39 | 0.19 | 0.22 | 0.41 | 1.7 | 0.51 |
| VOA | 0.47 | 0 | 0.02 | 0 | 0.14 | 0.02 | 0.01 | 0.43 |
| WAT | 0.07 | 0.3 | 0.06 | 0.25 | 3.37 | 0 | 0.03 | 1.81 |
| WHE | 0 | 0 | 0 | 0.03 | 0 | 0.01 | 0 | 0.3 |
| YAM | 0.06 | 11.79 | 0.35 | 0.11 | 0 | 0.36 | 4.95 | 0 |

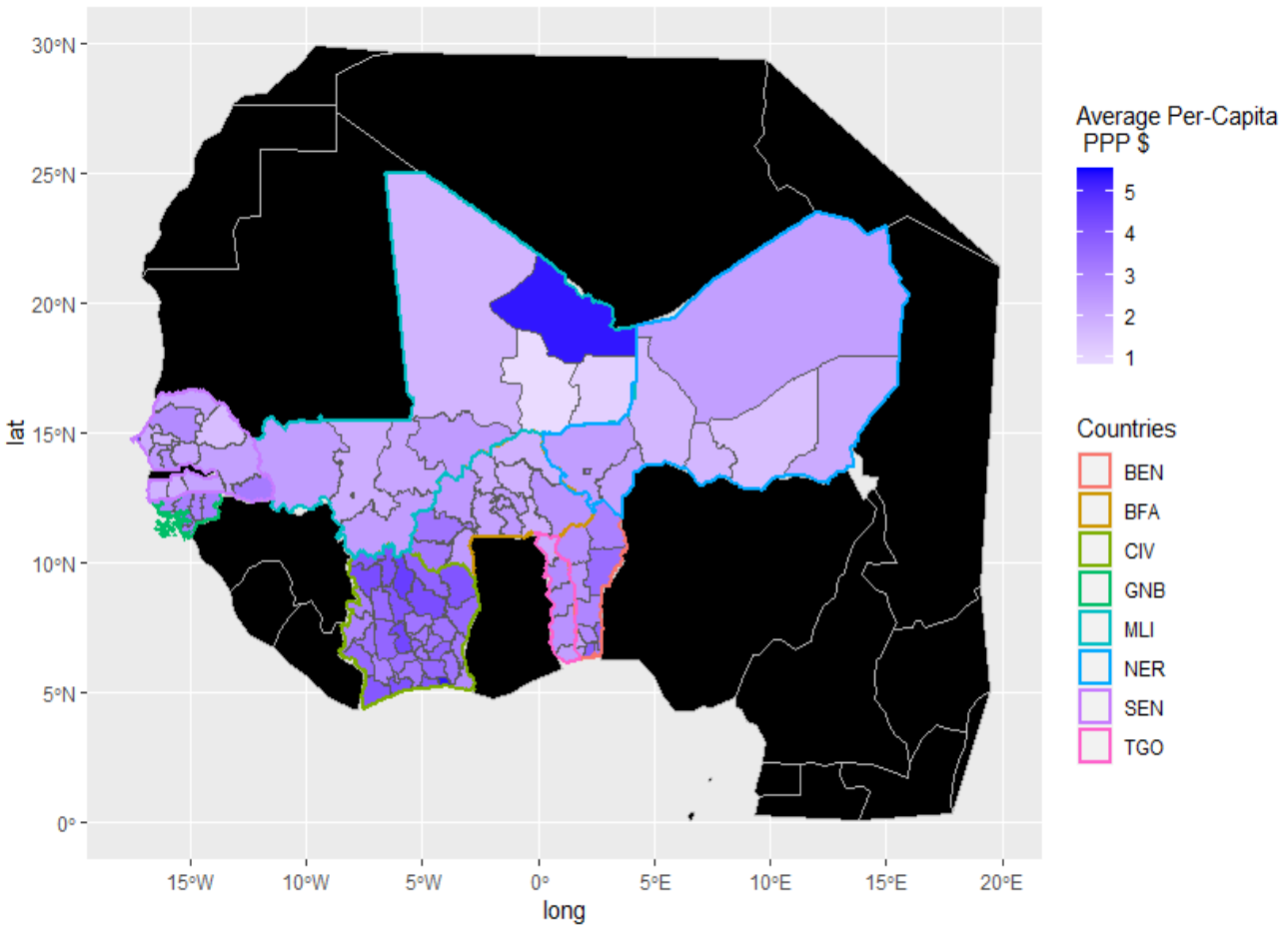
# Some Examples Output II

Spatial distribution of poverty rate (per–capita 1.9$ PPP) – National level



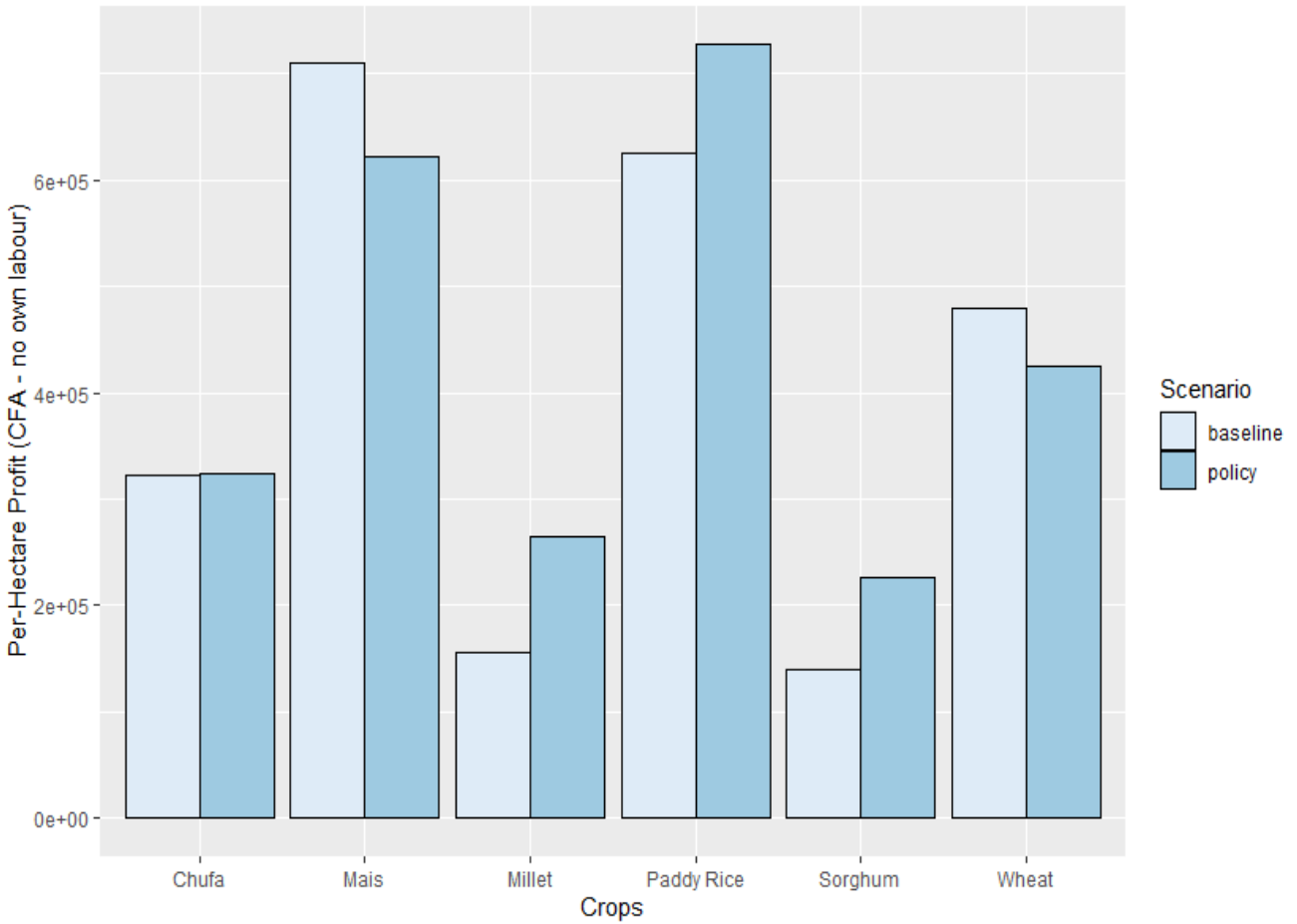
# Some Examples Output III

Spatial distribution of daily per–capita income ($ PPP) – Regional level



# Some Examples Output IIII

Comparison of profitability for selected crops between baseline and policy intervention



# Thank You!