Data Description

RWI – Leibniz Institute for Economic Research

FDZ Data Description: The Senegal Migration Panel (SMP) User Guide 1 - Overview

Cara Ebert, Bernd Beber, Niklas Murken, Nils Haveresch, Zara Riaz, Tobias Heidland, Maximiliane Sievert

March 2026

Imprint

Editor

RWI – Leibniz Institute for Economic Research e. V.
Hohenzollernstraße 1–3 | 45128 Essen, Germany

P.O. Box: 10 30 54 | 45030 Essen, Germany
Fon: +49 201–81 49-0 | E-Mail: rwi@rwi-essen.de
www.rwi-essen.de

Board of Directors

Prof. Dr. Dr. h.c. Christoph M. Schmidt (President)
Prof. Dr. Thomas K. Bauer (Vice President)
Dr. Stefan Rumpf (Administrative Board Member)
Prof. Dr. Kerstin Schneider (Member of the extended Executive Board)

RWI Data Description

Editor: Prof. Dr. Dr. h. c. Christoph M. Schmidt
Layout: Magdalena Franke, Claudia Lohkamp

FDZ Data Description:
The Senegal Migration Project (SMP)
User Guide 1 - Overview

March 2026

Project team
RWI: Cara Ebert, Bernd Beber, Niklas Murken, Nils Haveresch, Zara Riaz, Tobias Heidland, Maximiliane Sievert

Leibniz Institute for Economic Research

Leibniz Association

The RWI is funded by the federal government and the state of North Rhine-Westphalia.

Inhaltsverzeichnis

Tabellenverzeichnis

1. Introduction

This User Guide provides information about the sampling, data collection and use of the Senegal Migration Panel (SMP).

SMP is a survey of households and individuals in Senegal, focused on migration beliefs and behavior as well as climate change beliefs and adaptation.

The ***Core* of SMP** constitutes a **panel survey of households and male individuals** in **145 villages** in **rural Senegal**. So far, four *Core* data waves have been collected:

- Wave 1 (2022)
- Wave 2 (2023)
- Wave 3 (2024)
- Wave 4 (2025)

In addition, five add-on surveys were conducted as part of different waves, resulting in a total of the following surveys:

- ***Village chief*** survey in 2022 and 2024 (wave 1 & 3)
- ***Census*** of all households and their members in 2022 (wave 1) and of the same households and their members in 2024 (wave 3)
- ***Network*** survey of households in 2022 and 2024 (wave 1 & 3)
- ***Core-household*** survey in 2022, 2023, 2024, and 2025 (waves 1, 2, 3 & 4)
- ***Core-individual*** survey in 2022, 2023, 2024, and 2025 (waves 1, 2, 3 & 4)
- ***Urban*** survey of male individuals in 4 cities of Senegal in 2023 (i.e. outside of rural SMP villages)
- ***Irrigation*** survey of households in 2024
- ***Woman*** survey of female individuals in 2024 (wave 3)

The 145 SMP villages sampled from rural areas of Senegal include:

- 90 villages from three **focus departments**—Matam, Kaolack, and Sedhiou—with 30 villages sampled representatively within each department,
- 35 **national sample** villages sampled representatively from the rest of the country, excluding Dakar and selected border regions for security reasons,
- 20 **program evaluation** villages, purposively sampled for a program evaluation.

All village chiefs and households in these villages were sampled for the *Village chief* survey and the *Census* and *Network* surveys, respectively. For the *Core-household* and *Core-individual* surveys, households and individuals were selected based on prior migration experience and whether the household included a male aged 18 to 40 years in 2022.

We provide design-based sampling weights for all observations from randomly sampled villages (so excluding those in program evaluation villages) that can be used to produce nationally representative estimates of rural Senegal. We also provide separate weights for different subsamples, i.e. observations in each of the focus departments, the national sample, and program evaluation villages, which can be used to obtain representative estimates for each sampling frame.

Using these **weights, all surveys** except the *Urban* survey can be made **representative of Senegal as a whole** or of the relevant sampled **subpopulation of interest**.

Table 1: Number consented interviews and attrition rates relative to 2022 in parentheses by survey type and year

Survey	2022	2023	2024	2025
Village chief	N=144		N= 144 (100%)	
Census	N=9,665		N=8,841 (91.5%)	
Network	N=9,359		N=8,841 (94.5%)	
Core-household	N= 4,808	N=4,596 (95.6%)	N=4,514 (93.9%)	N=4,615 (95.0%)
Core-individual	N= 4,686	N=4,159 (88.8%)	N=3,771 (80.5%)	N=4,075 (87.0%)
Urban		N=1,009		
Irrigation			N=395	
Woman			N=961	

Source: Own data.

Table 1 presents an overview of the different surveys, their size by number of consented interviews, and timing. The surveys are ordered by sequence of implementation; i.e. the *Village chief* survey and *Census* were fielded first, followed by the *Network* survey which was conducted jointly with the *Core-household* and *Core-individual* surveys. In 2023, the *Urban* survey was conducted after the *Core-household* and *Core-individual* surveys were completed and in different locations. In 2024, first the *Irrigation* survey was conducted, followed by the joint implementation of the *Network*, *Core* and *Woman* surveys.

Attrition rates vary across waves and surveys. All village chiefs surveyed in 2022 were interviewed again in 2024. 91.5% and 94.5% of households listed in the 2022 *Census* and *Network* surveys were interviewed again as part of these surveys in 2024. The *Core-household* and *Core-individual* surveys have the lowest response rates in 2024, with 93.3% and 80.5% of the original surveys in 2022. Response rates of the *Core* surveys were higher in Wave 4 (2025), conducted over the phone, with 95.0% of households and 87.0% of individuals interviewed.

2. Data access

To date, Wave 1 (2022) datasets are available free of charge for scientific purposes as a Scientific Use File (SUF) Off-site. See *"SMP - User guide – 2 Wave 1 (2022) - 20260330.pdf*" for a description of the SMP Wave 1 (2022) surveys and data use.

For information about the release of future waves, please refer to: https://www.rwi-essen.de/en/research-advice/further/research-data-center-ruhr-fdz/data-sets.

3. Funding

The main funding source of the Senegal Migration Panel is the *Leibniz – Cooperative Excellence* program (grant no.: K473/2022), which funded the 2023 to 2025 *Census, Network, Core-household* and *Individual-core* surveys.

Additional funding was received by the *Policy Lab Climate Change, Development and Migration* of *the RWI – Leibniz Institute for Economic Research* for the following surveys: all SMP Wave 1 (2022) surveys, the 2023 *Urban* survey, the 2024 *Irrigation* survey and the 2024 *Village chief*

survey. Moreover, the *Urban* survey was in part funded by *the Kiel Institute for the World Economy.*

The 2024 *Woman* survey was funded by the *International Food Policy Research Institute* (IFPRI), which holds the rights to this data.

4. Description of surveys

Village chief:

- Respondent: Village chief
- Content: Village characteristics + perceived changes in weather and climate
- Type: In-person survey

Census:

- Respondent: Household head or knowledgeable person in household
- Content: Listing of all household members of age 15 and older ("Household roster") + members' basic sociodemographic background + their migration history
- Type: In-person survey

Network:

- Respondent: Household head or knowledgeable person in household (same as Core-household respondent) + only for network questions: Household head or knowledgeable person in household (same as Core-household respondent) jointly with male 18 to 40 years (same as Core-individual respondent)
- Content: Socioeconomic information of the household + social ties of the household
- Type: In-person survey

Core-household:

- Respondent: Household head or knowledgeable person in household
- Content:
  - 2022 & 2024: Socioeconomic information of the household + climate change perceptions of the household respondent
  - 2023 & 2025: Socioeconomic information of the household + migration history of household members + climate change perceptions of the household respondent
- Type: In-person survey (2022, 2023, 2024); phone survey (2025)

Core-individual:

- Respondent: Male 18 to 40 years
- Content: Detailed socioeconomic characteristics + migration attitudes and behavior + climate perceptions + wellbeing, risk preferences, etc.
- Type: In-person survey + phone mop-up (2022, 2023, 2024); phone survey (2025)

Urban:

- Respondent: Male 18 to 40 years

- Content: Detailed socioeconomic characteristics + skills and employment + migration attitudes and behavior + climate perceptions + wellbeing, risk preferences, etc.
- Type: In-person survey

Irrigation:

- Respondent: Household head or knowledgeable person in household
- Content: Information on cultivation and irrigation for 2023, an retrospectively for 2019.
- Type: Phone survey

Woman:

- Respondent: Female 18 to 64 years
- Content: Sociodemographic characteristics + migration history and intent + social and political networks + participation in community activities and local politics + women's empowerment
- Type: in-person survey

5. Sampling

This section describes the **sampling** for wave 1 **conducted in 2022**. All SMP surveys, except for the urban survey, build on this sampling. For the sampling specificities of surveys in later waves or add-on surveys, we refer to the respective User Guide of that wave.

The goal was to create samples of households, young male individuals and village chiefs in 145 rural villages that are representative for specific departments and Senegal as a whole. Therefore, the primary sampling units are villages, the secondary sampling units are households, and the tertiary sampling unit are male individuals.

Sampling of *villages*:
The sampling is based on a list of administrative units that were obtained from Senegal's National Agency for Statistics and Demography, ANSD. A village was defined as a settlement containing between 40 and 150 households. (Settlements are defined exclusively and exhaustively as "quartiers/villages" by ANSD.) Villages are ordinarily defined as containing at least several hundred residents (more than hamlets) but fewer than a couple of thousand (fewer than towns). Given an average household size of about 10, our definition meets this common notion of a village.

All urban communes and communes in the Dakar region were excluded from the sampling frame. Further, certain areas bordering the Gambia, Mali, or Guinea were excluded for security reasons (specifically, the departments of Bignona, Bounkiling, Medina Yoro Foulah, Saraya, Bakel, Salemata, and the district of Fongolembi). We also excluded ten villages from the sampling frame that had recently been enumerated as part of a related data collection effort. In total, the sampling frame covers 3,077 out of 3,477 rural villages across all regions and 35 of 41 departments outside of Dakar.

From that sampling frame, we selected 3 types of villages, with 145 villages in total:

- Type 1: Focus department villages, N=90

90 villages in 3 focus departments were selected *representatively* for the respective department. The departments are Matam, Kaolack and Sedhiou, with 30 villages sampled in each.

- Type 2: National sample villages, N=35

  35 villages were *representatively* sampled from the rest of the country, after excluding Type 1 and Type 3 villages.

- Type 3: program evaluation villages, N=20

  20 villages were *purposefully* selected for a program evaluation. This included 10 villages in which the program was implemented and 10 control villages. The control villages were selected based on a manual matching, mimicking the program selection process. In each treatment commune, we identified similar villages in terms of population using the ANSD list of administrative units. These villages were then screened through a short phone interview with the village chief. Selection criteria were absence of existing or planned irrigation infrastructure and availability of around 5 ha community owned land for a hypothetical irrigation system.

Type 1 and Type 2 villages were *randomly* sampled within location-based strata, namely within district for focus departments and within region for the national sample. A re-randomization algorithm was employed to ensure representativeness with respect to a set of characteristics for which administrative data was available (logged population and household counts, average household size, and the share of women at the village level, and the average age and share of residents recorded as absent at the level of the commune).

Sampling of *village chiefs:*
In each village the chief was sampled, resulting in 145 village chiefs to be approached for an interview (for a final sample size of 144).

Sampling of households for the *Census*:
All households (N=9,937) in the 145 villages were sampled and to be approached for an interview (for a final sample size of 9,665).

Sampling of households for the *Network* survey:
All households (N=9,937) in the 145 villages were sampled and to be approached for an interview (for a final sample size of 9,359).

Sampling of households for the *Core-household* survey:
We sampled 6,031 households in the 145 villages to be approached for a *Core-household* interview (for a final sample size of 4,808).

Households with a male in the age range of 18 to 40 years were eligible for the *Core-household* interview. Within-village household samples were drawn in proportion to village size, with a minimum target sample size of 15 households and a maximum of 45 households per village. Within villages sampling was stratified based on *Census* data of migration experiences among the target group of 18- to 40-year-old males, such that those that have migrated either domestically or internationally in the previous twelve months were undersampled roughly by a factor of two. Within each strata, household selection probabilities are proportional to the number of within-household target individuals, i.e. the relevant household size metric.

Sampling of male individuals for the *Core-individual* survey:
We targeted one male individual from each of the households sampled for the *Core-household* survey (for a final sample size of 4,686).

Male individuals in the age range of 18 to 40 years were eligible. The respondent was randomly selected among all males 18 to 40 years in the household. If the selected respondent was not available at the time of the data collection household visit, the next male on a list (with males 18 to 40 years in the household ordered randomly) was interviewed.

6. Weights

We provide design-based sampling weights that can be used to obtain estimates that are representative for particular populations. The weights and their use depend on the different village sampling strategies described in Section 5, where the national and focus department villages (Type 1 and Type 2 villages) were randomly sampled, whereas program evaluation villages (Type 3 villages) were purposively sampled.

The weights adjust for differences in village selection probabilities across strata used in the random selection of Type 1 and Type 2 villages (region strata for the national sample, and district strata for each of the three focus department samples in Kaolack, Matam, and Sedhiou), differences in household selection probabilities for *Core-household* interviews within villages, and differences in the probabilities with which target individuals for *Core-individual* interviews were selected within households. Villages in the program evaluation sample were purposively sampled, and so weights for observations in this sample do not adjust for differences in village selection probabilities.

- Variables pw_vill* give the scaled inverse of the village selection probability and can be used to ensure national / focus department representativeness when working with the *Village chief*, *Census* and *Network* data.
- Variables pw_core_hh* provide the scaled inverse of the overall core-household selection probability. They can be used to ensure national / focus department representativeness when working with *Core-household* data.
- Variables pw_core_ind and pw_core_ind*only provide the scaled inverse of the overall core-individual selection probability. They can be used to ensure national / focus department representativeness when working with *Core-individual* data.
- Variable pw_core_ind_in_vill is the scaled inverse of the core-individual selection probability, conditional on village selection. This variable can be used to ensure that estimates are representative of SMP villages when working with the *Core-individual* data, and is of particular interest when working with program evaluation data.

All weights are final design weights that take into account stage-specific selection probabilities of previous sampling stages. They do not need to be combined.

We provide weights that can be used to ensure national representativeness when using all observations from randomly sampled villages (i.e., national sample and focus department sample observations, but not program evaluation sample observations), as well as weights for each sample type separately (i.e., national sample, Kaolack focus sample, Matam focus sample, Sedhiou focus sample, and the program evaluation sample).

- Use weights **pw_vill, pw_core_hh**, and **pw_core_ind** to ensure national representativeness when using all observations from randomly sampled villages,
- Use weights **pw_*_national_only** when using the national sample only,
- Use weights **pw_*_kaolack_only** when using the Kaolack focus sample only,

- Use weights **pw_*_matam_only** when using the Matam focus sample only,
- Use weights **pw_*_sedhiou_only** when using the Sedhiou focus sample only, and
- Use weights **pw_*_PROGRAM_only** when using the program evaluation sample to ensure representativeness with respect to either households or individuals (males aged 18-40) within program evaluation villages.

Since the weights are design-based, villages in the program evaluation sample do not have weights that take into account village selection.

While weights are well-distributed for each sample type, note that weights pw_vill, pw_core_hh, and pw_core_ind have a relatively high variance, because the set of all randomly sampled observations combines a large number of villages from focus departments with small numbers of villages from other departments.

Strata and sampling units:

- Village selection was carried out within region strata (variable **strata_natl**) for the national sample, and within district strata (variable **strata_focus)** for the focus samples. We also provide a consolidated set of village-level sampling strata (variable **strata_vill**) for use with all observations from randomly sampled villages (i.e. not including program evaluation sample observations). Within villages, households were again selected in strata (given by variable **strata_hh**) based on prior migration exposure (available for reference in variable **strata_mig**).
- The primary sampling unit (PSU) is the village (variable **ea_anon**). Within villages, the sampling unit is the household (variable **hhid_anon**). An individual for the *Core-individual* survey is then selected within a sampled household.
- We provide variables that can be used for finite population corrections at each sampling stage, including the number of villages in each first-level stratum (variable **fpc_vill**), the number of households in each second-level stratum (variable **fpc_hh**), and the number of target individuals in each household (variable **fpc_ind**). With randomly sampled villages, unscaled overall selection probabilities are small and gains from declaring finite population corrections are negligible. For inferences conditional on village selection, however, users may wish to carefully consider whether to include them.

These details concerning weights, strata, sampling units, and finite populations can be used to declare a survey design, for example in Stata. We provide some examples below.

- For national representativeness at the individual level, i.e. rural males aged 18-40, using all observations from randomly sampled villages:
  ```
  svyset ea_anon [pweight = pw_core_ind], strata(strata_vill) fpc(fpc_vill) || hhid_anon, strata(strata_hh) fpc(fpc_hh) || _n, fpc(fpc_ind) singleunit(certainty)
  ```
- For national representativeness at the household level, i.e. rural households that include a male aged 18-40, using all observations from randomly sampled villages:
  ```
  svyset ea_anon [pweight = pw_core_hh], strata(strata_vill) fpc(fpc_vill) || hhid_anon, strata(strata_hh) fpc(fpc_hh) singleunit(certainty)
  ```
- For national representativeness at the village level, e.g. when working with chief data:

```
svyset ea_anon [pweight = pw_vill], strata(strata_vill)
fpc(fpc_vill) singleunit(certainty)
```

- For representativeness at the individual level in Kaolack, using only observations from the Kaolack focus sample:
  ```
  svyset ea_anon [pweight = pw_core_ind_kaolack_only],
  strata(strata_vill) fpc(fpc_vill) || hhid_anon, strata(strata_hh)
  fpc(fpc_hh) || _n, fpc(fpc_ind) singleunit(certainty)
  ```
- For representativeness at the household level in Matam, using only observations from the Matam focus sample:
  ```
  svyset ea_anon [pweight = pw_core_hh_matam_only],
  strata(strata_vill) fpc(fpc_vill) || hhid_anon, strata(strata_hh)
  fpc(fpc_hh) singleunit(certainty)
  ```
- For representativeness at the individual level within the set of sampled villages, i.e. conditional on village selection and including observations in the program evaluation sample:
  ```
  svyset hhid_anon [pweight = pw_core_ind_in_vill],
  strata(strata_hh) fpc(fpc_hh) || _n, fpc(fpc_ind) singleunit(cer-
  tainty)
  ```
- For representativeness at the individual level within the set of villages included in the program evaluation sample:
  ```
  svyset hhid_anon [pweight = pw_ind_PROGRAM_only],
  strata(strata_hh) fpc(fpc_hh) || _n, fpc(fpc_ind) singleunit(cer-
  tainty)
  ```

7. Linking households and individuals across surveys and waves

Table 2 presents a summary of variables linking households and individuals across surveys and waves.

Households can be linked across *Census, Network, Core-household, Core-individual, Irrigation* and *Woman* surveys, and waves using the variable *hhid_anon*.

Note that respondents of the household level surveys may not be the same across waves. However, respondent identifiers are recorded and can be linked across waves (if they were the same) and to the *Census*.

The male individual respondent of the *Core-individual* survey can be linked across waves of the *Core-individual* and *Census* surveys using the variable *ind_id_anon*.

Female respondents of the *Woman* survey can be linked to the household member roster of the *Census* survey. They can also be linked to their husbands by matching the woman-survey variable *husband_link* to the *Core-individual* variables *ind_id_anon.*

Household members listed in the roster of the *Census* survey can be linked across waves using the variable *ind_id_anon*.

Village chiefs can be linked at village level to all other surveys using *ea_anon*. While village chief interviews cannot be linked at household level to the other surveys, village chief interviews can be linked across waves using the variable *hh_chief_id*.

Table 2: Summary of variables linking households and individuals across surveys and waves

What to link:	Across surveys:	2022	2023	2024	2025
Household	Village chief	hhid_chief_anon		hhid_ chief_anon	
	Census	hhid_anon		hhid_anon	
	Network	hhid_anon		hhid_anon	
	Core-house-hold	hhid_anon	hhid_anon	hhid_anon	hhid_anon
	Core-individual	hhid_anon	hhid_anon	hhid_anon	hhid_anon
	Irrigation			hhid_anon	
Male respondent of Core-indi-vidual survey	Census Core-individual	ind_id_anon ind_id_anon	 ind_id_anon	ind_id_anon ind_id_anon	 ind_id_anon
Household members	Census	ind_id_anon	ind_id_anon	ind_id_anon	ind_id_anon
Female respondent of Woman survey	Census Woman Core-individual	ind_id_anon		ind_id_anon womanid husband_link	

Source: Own data.

Table shows variables that link households and individuals across waves and surveys.