Data Documentation: Syrian Refugee Life Study (S-RLS) - Phone Survey

Hasan Ebussuutoglu[*] Mansi Kalra[†] Edward Miguel[‡] I. Bailey Palmer[§]
Sandra V. Rozo[¶] Emma Smith[‖] Sarah Stillman[**] Abdulrazzak Tamim[††]

April 2024

Contents

1 Team and Acknowledgements **3**

2 Introduction **3**
2.1 Aims and Objectives of the Study . . . 3
2.2 Overview of S-RLS Survey Instruments . . . 3
2.3 Timeline of Activities . . . 4
2.4 Principal Investigators . . . 4

3 S-RLS Sample and Tracking Design **4**
3.1 Panel Sample Selection . . . 4
3.2 Panel Tracking Design . . . 5

4 S-RLS Questionnaire Design **5**
4.1 Interview Details and Respondent Burden . . . 5
4.2 Survey Modules . . . 5
4.2.1 Pre-Interview Information and Consent . . . 5
4.2.2 Confirmation of Identity . . . 5
4.2.3 Contact information . . . 5
4.2.4 Household Rosters . . . 5
4.2.5 Dwelling Characteristics . . . 6
4.2.6 Meals Consumption and Economic Outcomes . . . 6
4.2.7 Migration . . . 6
4.2.8 Health . . . 6
4.2.9 COVID-19 . . . 6
4.2.10 Conclusion . . . 6

5 S-RLS Data Collection **6**
5.1 Field Staff . . . 6
5.2 Data Quality Management . . . 6

[*]University of California, Berkeley, email:has10@berkeley.edu
[†]Center for Effective Global Action, email: mkalra@berkeley.edu
[‡]University of California Berkeley, email:emiguel@berkeley.edu
[§]University of California Berkeley, email:bailey.palmer@berkeley.edu
[¶]World Bank, Research Group, email: srozovillarraga@worldbank.org
[‖]Harvard University, email:esmith@g.harvard.edu
[**]Center for Effective Global Action, email: sstillman@berkeley.edu
[††]University of California, Berkeley, email:abdulrazzak.tamim@berkeley.edu

6 Data Issues **7**
6.1 Fertility Section 7

7 Changes to Dataset **7**
7.1 Removed Variables 7

1 Team and Acknowledgements

The Syrian Refugee Life Study (S-RLS) is a joint effort between Dr. Edward Miguel (UC Berkeley), Dr. Sandra Rozo (World Bank), and Emma Smith (Harvard University). This work has depended heavily on the support and guidance of numerous researchers and field staff, including Dr. Samuel Leone (McKinsey), a founding member of S-RLS. Our research coordinator, Sarah Stillman, and research associate Mansi Kalra, have been invaluable assets to the project and its success. The project has also benefited from theoretical and technical contributions from Bailey Palmer and Abdulrazzak Tamim, as well as excellent research assistance from Hasan Ebussuutoglu, Joaquin Fuenzalida, John Gorham, Peter Flo Grinde-Hollevik, Yazen Kashlan, Meghana Kumar, Charlotte McClelland, Benjamin Shenouda, and Andy Theocharous. Our competent field team directed by Mindset has been crucial to the success of the project. We thank Mohammad Qardan, Rana Samara, Majd Haddad, Feryal Ghazi, Mohammad Qaryouti, and the team of enumerators. We also thank members of the academic and policy community who gave valuable feedback during workshops and seminars.

2 Introduction

The Syrian Refugee Life Study (S-RLS) is among the first systematic efforts to survey a large, representative refugee sample and follow that sample over time. The sample includes approximately 2,500 Syrian refugee households residing in camps and in host communities. S-RLS currently collects information on Syrian refugees' economic well-being, health, integration, migration patterns, and intention to return and will track changes to these outcomes over time.

2.1 Aims and Objectives of the Study

This study serves as the first long-term panel of Syrian refugees in Jordan (to our knowledge), and one of the first globally, providing unprecedented spatial and temporal coverage of a population that is considered extremely hard to successfully track. More broadly, this comprehensive and representative data can provide policymakers with insights into the long-run effects of refugee crisis responses, and offer credible evidence for formulating more effective policies in future crises. The panel nature of the data also significantly expands the type of research designs available to researchers, and the interdisciplinary scope of the instrument will likely prove useful for many different lines of inquiry.

2.2 Overview of S-RLS Survey Instruments

The S-RLS consists of two sub-projects: a Randomized Control Trial (RCT) and a panel. We are committed to creating a lasting public good for the research and policy communities through publication of all deidentified data on Harvard Dataverse, enabling scholars and decision-makers to partner with us on new projects or independently pursue projects utilizing our data.

- **Panel** The panel began with an in-person survey module that rolled out in February, 2020. The panel sample consists of a representative sample drawn from the universe of Syrian refugees registered with the United Nations High Commissioner for Refugees (UNHCR) in Jordan. The February 2020 survey module is extremely detailed and gathers information about household finances, housing, consumption, education, labor, marriage, fertility, attitudes, migration history, etc. Only 468 in-person surveys (20 percent of the sample) were surveyed before in person surveying was halted due to the COVID-19 pandemic. Data from this sub-sample is currently available to the public on the Harvard Dataverse.

  Following the study's halt, the entire panel sample was resurveyed using an abridged phone survey from December 2020 to March 2021. This shorter survey focused on housing quality, basic labor and credit outcomes, a household roster, child school attendance, and COVID-19 topics. This round collected 2,576 surveys in total (62% of sample). In December 2021, we began in person surveying of the panel sample. This round collected 2,601 surveys (63%). Another follow up phone survey was launched in Summer of 2023, and a projected final round of the panel study is anticipated to be launched in Spring of 2024.

This research team also conducted a RCT evaluating the impact of a housing program run by the Norwegian Refugee Council, called the Shelter Program. The RCT sample differs from the panel sample, but was first surveyed using the same abridged phone survey as the panel sample.

2.3 Timeline of Activities

Panel

Nov. 2019	UNHCR/University of California, Berkeley (UCB) Data Use Agreement signed
Jan. 2020	UNHCR shares deidentified universe of refugees
Jan. 2020	UCB selects representative sample
Jan. 2020	Survey piloting
Feb. 2020	UNHCR shares initial batch of cases
Feb. - Mar. 2020	In-Person Survey rollout (halted due to COVID-19)
Dec. 2020 - Mar. 2021	Phone Survey rollout with full sample
Dec. 2021 - Mar. 2022	In-Person Survey rollout with full sample
Jul. - Aug. 2023	Phone Survey 2 rollout with full sample

RCT

Aug. 2019 - Mar. 2020	Shelter Program applications open
Feb. 2019	NRC/UCB Memorandum of Understanding signed
Aug. 2019 - Mar. 2020	RCT randomization
Aug. 2019 - Aug. 2021	RCT implementation
Jul. - Sep. 2020	Midline Survey rollout (Phone Survey 1)
Sep. - Nov. 2021	Endline Survey rollout (In-Person Survey)
Jul. - Sep. 2022	Neighbors Survey rollout
Nov. 2022- Feb. 2023	Follow-up Survey rollout (Phone Survey 2)

2.4 Principal Investigators

S-RLS was designed by PIs Samuel Leone, Edward Miguel, Sandra Rozo, and Emma Smith. Samuel Leone is a management consultant with McKinsey & Co. Edward Miguel is a Professor of Economics at the University of California, Berkeley. Sandra Rozo is a research economist with the World Bank's development research group. Emma Smith is a PhD candidate in Public Policy at Harvard University.

3 S-RLS Sample and Tracking Design

3.1 Panel Sample Selection

After establishing a Data Use Agreement between UC Berkeley and the United Nations Commissioner for Refugees (UNHCR), the UNHCR Jordan country office shared the deidentified universe of registered Syrian refugees in Jordan ($\sim$ 655,000 total observations) with the research team. The team then randomly selected 6,997 adult individuals, stratifying on age, gender, and location in Jordan (governorate), to create a representative sample.

UNHCR shared contact details for our sample of 6,997 adults. These details included locality-level geographic information and between one and four phone numbers per case. For cases where we were unable to contact the designated respondent (due to out-of-service or wrong numbers), we requested additional contact details from UNHCR. 23% of our sample did not have a single working phone number. As such, even though there has been a relatively low level of refusals, we have faced a lower-than-optimal tracking rate. Working with UNHCR, we were able to determine that some of our randomized sample was already out of contact with UNHCR before the study began (which boosts our tracking rate 6 percentage points per round). We also found that unreachable cases were disproportionately less vulnerable and more likely to own a passport, indicating that the "missing" people may be more economically mobile people who were able to leave the country of their own volition.

3.2 Panel Tracking Design

In the first interview with each respondent, we collect very detailed contact information. This includes an email address, Facebook account and two phone numbers where they may be contacted. We also collect detailed contact information for two other individuals who will reliably know where the focus respondent lives.

To schedule an interview, we contact each phone number listed for each household up to five times (i.e., up to 20 points of contact), calling on different days, at different times, and with different enumerators to boost our response rate. We also keep track of travelling cases, deceased cases (which are replaced with other household members), and cases whose phones are turned off (whereby we may call > 5 times over the course of several weeks in case the phones are turned back on). For those cases who do not answer or whose phone numbers were no longer working, we engage in intensive, in-person tracking. Our detailed location data (collected orally and via GPS) allow us to visit the respondent's last known address.[1]

4 S-RLS Questionnaire Design

The S-RLS questionnaire was designed to capture a full picture of the focus respondent's individual and household-level life outcomes. We survey focus respondents (FRs) on housing quality, household consumption, asset ownership, credit, aid, and transfers. We also ask individuals about income generating activities, migration history, health, and marriage.

4.1 Interview Details and Respondent Burden

The questionnaire was translated to Modern Standard Arabic by staff translators at Mindset, with revisions translated by our bilingual field coordinator, Mohammad Qardan and Research Associate Abdulrazzak Tamim. Where appropriate, we used Arabic-validated questions and modules. Respondents are compensated for their time with a payment of 5 Jordanian Dinar ($7 USD).

4.2 Survey Modules

The survey has 10 sections in total, with no breaks given that it was conducted over the phone.

4.2.1 Pre-Interview Information and Consent

This section gathers basic information about the focus respondent and collects consent.

4.2.2 Confirmation of Identity

This section collects full name, birth date, parent's full names and information of identification documents.

4.2.3 Contact information

This section collects numerous forms of contact information for the focus respondent and for two other individuals who would know how to contact the FR if they moved away, and who would not move with the FR. We also collect information on any likely moves in the future.

4.2.4 Household Rosters

We first collect information on the current household roster. For each member in the household, we ask name, whether they slept in the home last night, gender, age, relationship to FR, highest level of education completed, primary occupation, school attendance for children under 18, citizenship, and whether they joined the household after January 2020.

[1] For our phone survey round, we also attempted to send small incentives that would ensure phones were working prior to interview, but this resulted in only a few additional responses.

4.2.5 Dwelling Characteristics

This section collects data on characteristics of the home, including number of rooms, number of families occupying the dwelling, building materials, access to and quality of utilities, rent/own status and information of rent/mortgage payments, and time use on chores.

4.2.6 Meals Consumption and Economic Outcomes

Section 6.1 collects information on food consumption and food security. Section 6.2 measures economic outcomes prior to the COVID-19 pandemic, during, and after. It looks specifically at lending, hours worked per week, and total number of adults working in the household.

4.2.7 Migration

This section collects location as of January 2011, any moves since 2019, and tracks any potential moves in the future.

4.2.8 Health

These sections capture various aspects of physical and mental health. Section 8.1 captures prevalence of common symptoms and illnesses, doctor visits and medical expenditures in the past month, including skipped visits. It also collects information regarding the COVID-19 pandemic. We capture disability using the Washington Group Short Set on Disability. Section 8.2 evaluates mental health and well being. We measure depressive symptoms using using the CESD-10. Both of these instruments are validated. In section 17.3 we survey respondents perceived stress.

4.2.9 COVID-19

This section elicits information about the impact of COVID-19 on schooling, health, and levels of discrimination.

4.2.10 Conclusion

In the conclusion section, we ask about data reliability, fluency with Arabic, and provide the FR with information for UN help lines.

5 S-RLS Data Collection

5.1 Field Staff

We partner with Mindset, a Jordanian market research firm, to collect the S-RLS data and aid with data processing. Under the S-RLS's team direction, Mindset prepares electronic survey materials; hires, trains, and manages a team of local enumerators; oversees survey implementation and quality assurance; and conducts basic data preparation of the survey after enumeration. Mohammad Qardan serves as our Jordan-based project coordinator. Throughout the life of the project, we have worked with over 50 enumerators and 15 supervisors, with team size fluctuating between 10 and 35 at a given time.

5.2 Data Quality Management

Data quality management procedures for the S-RLS include in-field monitoring where at least 20% of interviews are monitored by a supervisor. Each interview's GPS location is plotted on a satellite map to confirm that the interview was conducted in the respondent's address, and survey administration time is monitored to ensure outliers are rectified.

Mindset also conducts back-checks to assure data quality after the interview is completed. Mindset staff call at least 20% of the interviewees afterwards to confirm that the respondent was actually interviewed and ask five random questions from a list of 20 questions from the survey. If more than half of the callback data

differs from the original data collected, the entire survey is recollected. If the mismatch rate is less than 50%, enumerators correct the questions which differ but otherwise do not resurvey the household.

6 Data Issues

We faced difficulties due to the poor quality of contact information data from UNHCR that serves as the sampling frame for the panel – 23% of our sample did not have a single working phone number. We combated this by utilizing Whats App to contact respondents. We also worked with UNHCR to determine which of these cases were unreachable from the beginning, such that we may be able to reconfigure our sample slightly while maintaining its representativeness.

6.1 Fertility Section

In this round of the fertility section, there was an incorrect skip pattern, and therefore for children who had passed away their birth date was not collected. Data users should be mindful of this when assessing data.

7 Changes to Dataset

7.1 Removed Variables

The following variables have been removed from the dataset

- 2011 location and location of first arrival into Jordan (only governorate level location variables and above will be shared).
- Likelihood that Syrian conflict will end in next two years
- Camp perceptions
- Migration history in Jordan (including time spent in camps)

These variables will be included in the final published dataset for this round, however, they are excluded for the time being due to vested interest of existing S-RLS researchers.