Data Documentation: Syrian Refugee Life Study (S-RLS) - Full In-Person Round

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 **2**

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

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

4 S-RLS Questionnaire Design **4**
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 Schooling History . . . 5
4.2.5 Household Rosters . . . 5
4.2.6 Dwelling Characteristics . . . 5
4.2.7 Food Consumption . . . 5
4.2.8 Non-Food Purchases . . . 5
4.2.9 Durables . . . 5
4.2.10 Self-Employment: Current and History . . . 6
4.2.11 Transfers . . . 6
4.2.12 Interaction with NGOs and Government Programs . . . 6
4.2.13 Savings and Credit . . . 6
4.2.14 Risk Preferences, Time Preferences, Ambiguity, Trust, Raven's Tests and Happiness . 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

4.2.15 Migration . . . 6
4.2.16 Health, Nutrition, Sleep, Stress, and Grit . . . 6
4.2.17 Child Strengths and Difficulties and Child Sleep . . . 7
4.2.18 Religion . . . 7
4.2.19 Host Community Relations . . . 7
4.2.20 Community Groups, Social Capital and Political Attitudes . . . 7
4.2.21 Safety and Crime Victimization . . . 7
4.2.22 Marriage . . . 7
4.2.23 Fertility . . . 7
4.2.24 Conclusion . . . 7
4.3 Travelling Cases and Respondents in Syria . . . 8

5 S-RLS Data Collection **8**
5.1 Field Staff . . . 8
5.2 Data Quality Management . . . 8

6 Data Issues **8**
6.1 Poor Tracking . . . 8
6.2 Survey Length . . . 8
6.3 Migration . . . 9
6.4 Fertility . . . 9
6.5 Migration . . . 9

7 Changes to Dataset **9**
7.1 Removed Variables . . . 9
7.2 Other Changes . . . 9

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. - 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

For the face to face survey, our field team conducted up to 5 trial phone calls with each number received per household, calling on different days and different times of day with different enumerators. Furthermore, the team obtained updated contact info for the sample, which resulted in new phone numbers to use for tracking purposes. The team utilized in person tracking for those with detailed location data (i.e. those who completed the phone survey previously). Lastly, we conducted social media (Facebook) tracking for 50 unreachable cases.

4 S-RLS Questionnaire Design

The S-RLS in-person 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 including self and wage-employment, migration history and intent to return, risk and time preferences, raven's tests, social attitudes and community relations, health, marriage, and fertility.

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 24 sections, with 5 minute breaks scheduled after sections 9 and 21.

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 Schooling History

This section collects information on each level of schooling the FR has entered. For each level we collect public/private status, year last attended, country, and length of time spent in the level.

4.2.5 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 were a member of the household in January 2011.

We then collect data on individuals who lived in the household in 2011 but do not currently.

4.2.6 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.

4.2.7 Food Consumption

This section collects information on food consumption in 9 food categories. We disaggregate consumption into assistance, home production and market purchases. We also ask several food security questions.

4.2.8 Non-Food Purchases

This section collects monthly and annual household expenditures in 9 categories. The categories include utilities, water, infant needs, household and hygiene items, debt repayment, linens, clothing, reusable household items, and "school fees, taxes, insurance and phone bills".

4.2.9 Durables

This section collects annual expenditure and ownership of 8 categories of durable goods.

4.2.10 Self-Employment: Current and History

The self-employment section collects detailed financial information on all self-employment activities in the past 12 months. This includes industry and occupation, months active, regularity, hours worked (typical or last week), employment of workers, revenues, expenses and profits.

The employment section captures the three most important jobs held in the last 12 months. For each job, we ask respondents about the start date, occupation, industry, contract status, working patterns, hours, salary, taxes, additional compensation, and commute time. We ask unemployed workers about search patterns, and all focus respondents about occupation, industry, own income, and household income in 2011.

4.2.11 Transfers

This section elicits information about transfers between households. For both sending and receiving relationships, we survey respondents about the 3 most important transfer relationships. For each relationship, we ask about location of transfer partner, mix of cash/goods, use of funds, total value over 12 months, and reciprocation. We also collect data on total transfers send/received over the past 12 months.

4.2.12 Interaction with NGOs and Government Programs

This section elicits information about receipt of cash, goods and employment aid from NGOs or community groups and government programs. This section asked about aid from specific organizations as well.

4.2.13 Savings and Credit

This section surveys focus respondents about savings behavior in banks and community groups. It also elicits information about application and receipt of formal and informal loans in the past 12 months. We ask for the total amount of purpose of all loans, and for the interest and due date on the most recent loan.

4.2.14 Risk Preferences, Time Preferences, Ambiguity, Trust, Raven's Tests and Happiness

We elicit risk and time preferences according to the Jordan version of Global Preferences Survey. We elicit ambiguity preferences using a variant of the traditional two-urn Ellsberg choice. Trust in various different groups is measured using a variant of the Dictator Game. We measure trust in other Syrian refugees, other refugees in Jordan, the average Jordanian in the same religious community, the average Jordanian in a different religious community, a Jordanian solider/police officer, a Jordanian public servant, and an employee of an international NGO. The preference elicitations are not incentivized. We measure cognitive skills with a Raven's test, and we measure happiness using a 1-7 scale.

4.2.15 Migration

The migration section asks focus respondents where they lived in January 2011, before the onset of conflict in Syria. We then ask respondents about each of their moves longer than four months between then and now. For each move, we ask respondents when they moved, to where, why they left, and with whom they moved. We also ask numerous questions about the circumstances under which they fled Syria and their arrival to Jordan. We then ask about any steps they are taking to move again, and their intent to return to Syria.

4.2.16 Health, Nutrition, Sleep, Stress, and Grit

These sections capture various aspects of physical and mental health. Section 17.1 captures prevalence of common symptoms and illnesses, doctor visits and medical expenditures in the past month, including skipped visits. We survey respondents about major illnesses or injuries since January 2011, and capture disability using the Washington Group Short Set on Disability. Section 17.2 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 amount and quality of sleep, and in section 17.4 we use the 4-item Perceived Stress Scale to measure perceived stress. Finally, we ask two questions selected from the Grit Scale to measure grit in focus respondents. The Grit Scale was developed by Angela Duckworth, a professor of psychology at the University of Pennsylvania.

4.2.17 Child Strengths and Difficulties and Child Sleep

This section first administers the validated Child Strengths and Difficulties questionnaire to a randomly selected child between the ages of 3 and 8. The questionnaire differentiates between children aged 3 and 4+ to account for normal patterns of child development. We then ask about quantity and quality of this child's sleep, including naps. Lastly, we ask about other child outcomes such as learning activities.

4.2.18 Religion

This section asks respondents their religion or denomination. We then elicit the level of religiosity in the community, and whether this level has changed in the past 12 months. We ask about religious donations, and whether nationality or sect is more important in the respondent's life.

4.2.19 Host Community Relations

This section surveys the respondent about integration with the host community and language acquisition.

4.2.20 Community Groups, Social Capital and Political Attitudes

This section asks respondents an open-form question about how describe themselves - as a Syrian, Muslim, Sunni, woman, etc. We next ask about participation in community groups, charitable giving, religious attendance, news consumption, and political engagement. We also included several questions about democracy, politics, corruption and gender issues from the Arabarometer. We conclude this section by asking respondents if they were ever arrested or imprisoned in Syria.

4.2.21 Safety and Crime Victimization

This section asks detailed questions about crime and objective measures of safety, specifically the victimization of theft and physical assault, and perceived safety.

4.2.22 Marriage

This section surveys all ever-married respondents about characteristics of each of their marriages. For each marriage, we ask respondents the first name of the spouse, age at marriage, how long they had known the spouse before marriage, whose decision it was to get married in general, and to the spouse in particular, the amount of the prompt and deferred dowry in the marriage contract, whether the marriage was registered in court, birth control usage, whether the spouse is still living and cause of death if deceased, religion, governorate of birth and education level of the spouse, when the marriage ended if applicable, and whether the marriage was polygamous.

4.2.23 Fertility

This section first asks women their age at menarche. Then all respondents are asked whether they or a partner have ever been pregnant. Then characteristics of each pregnancy are discussed. We elicit whether prenatal care was sought and where, how the pregnancy ended (live birth, still birth, miscarriage or abortion), when the pregnancy ended if not live birth, where the baby was born if a live birth, the baby's gender, birthdate, vaccination status and whether the child is still living. If the child has passed away, we ask when the child died. For children still living we ask about the child's subjective level of health and who the child lives with if they do not reside in the household. We also ask respondents who makes important child-rearing decisions in the household. See section **??** for data usage issues.

4.2.24 Conclusion

In the conclusion section, we ask about data reliability, fluency with Arabic and collect the GPS coordinates of the home.

4.3 Travelling Cases and Respondents in Syria

Our team made efforts to reach respondents that were marked as unreachable due to travel by adapting the in person survey to a shortened phone survey. This survey was constructed similarly to the one mentioned above, but without the modules for risk and time preferences, ambiguity, altruism, and Raven's games. Similarly, a shortened version of the survey was deployed as a phone survey for cases that were found to be Syria. The same modules mentioned previously were dropped from this survey as well. This was to ensure that the survey could be administered in a timely manner over the phone.

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 backchecks 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

This section outlines challenges encountered during the in person 2022 round of data collection which impact data structure or quality.

6.1 Poor Tracking

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.

6.2 Survey Length

This face to face survey was quite comprehensive and enumerators averaged one hour and a half to administer each survey. This led to burn out among respondents and enumerators, as reported by the survey staff. Breaks were built into the survey to mitigate this, however in order to capture the pre-specified outcomes the survey length could not be excessively reduced.

6.3 Migration

Approximately 40 days into surveying, a small edit was made in the migration section. Question 8 was amended to the following: "After living in [village q2e], did you live in any other neighborhood for at least four months, or another refugee camp for at least one night? Note: we only want to capture moves at least 4 months long, unless the move is to a refugee camp, in which case we want any overnight stay in a camp." This change was made to capture any moves, even under 4 months, if they were to a refugee camp. Prior to this change such moves were not captured.

6.4 Fertility

Enumeration of this section was challenging and data may be incomplete due to the long recall time horizon as well as technical issues. Specifically, there was an incorrect skip code that led to birth dates not being recorded for children who had passed away. The intention was to capture data on each pregnancy, but oftentimes household members did not recall all of the relevant details.

6.5 Migration

There was an issue with enumerators not reading important instructions to the respondents. The instruction "Confirm that the last reported residence listed in this table matches the current residence listed in Section 3 of this survey. If not, probe the FR further about their migration history" was not always read. Therefore, there are several people who do not currently live in a camp but the last reported residence is a camp.

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).
- Intentions to return and prediction of 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.

7.2 Other Changes

Sections 14, 17, and 18 have a short subset of variables that were renamed to align with the numbering of the survey. Additionally, Section 5, the roster section, was relabelled with notes to reflect which subset of the sample the variable applies to. For example, there are some variables that refer to the cumulative household members from the previous round and this round, and other variables that refer only to this round. Data users should be sure to consult the labels to determine which roster variables correspond to which round of data.