

***Access to Durable Solutions Among IDPs in Iraq:***  
**Quantitative Methodology Report**

**June 2021**

**Salma Al-Shami**  
**Georgetown University**

**Lorenza Rossi**  
**International Organization for Migration**

**GEORGETOWN**  
**UNIVERSITY**



## TABLE OF CONTENTS

<b>I. INTRODUCTION</b>	
Section 1: About this Report	3
Section 2: Study Overview	3
<b>II. RESEARCH DESIGN</b>	
Section 1: Respondents	5
Section 2: Sample Frame, Design, and Strategy	6
Section 3: Survey Questionnaire	8
Section 4: Changes to the Questionnaire Over Time	9
<b>III. FIELDWORK AND PROCESSING</b>	
Section 1: Survey Fielding and Mode	10
Section 2: Weighting	11
Section 3: Other Weighting Considerations	12
<b>IV. THE DATASET AND CODEBOOKS</b>	
Section 1: Panel Dataset	13
Section 2: Codebooks	14
Section 3: Variable Naming Conventions	14
Section 4: Missing and Not Applicable Data	15
<b>V. SAMPLE COMPOSITION</b>	
Section 1: Retention Analysis	16
Section 2: Changes in the Sample Composition	17

## LIST OF TABLES

Table 1: Household Status by Round	6
Table 2: DTM Reported IDP Population, December 2015	6
Table 3: Sample Strata	7
Table 4: Target Sample for Study	7
Table 5: Survey Mode	10
Table 6: Retention Numbers and Rates	16
Table 7: Sample Composition Over Time by Round 1 Stratum	17

## INTRODUCTION

### Section 1: About this Report

This report fully details the quantitative methodology of the panel study *Access to Durable Solutions Among IDPs in Iraq*. Its purpose is to complement information provided on the study website’s methodology page and provide more information for both those interested in how the study was conducted and to assist users of the study’s publicly available quantitative dataset.

The report is divided into five parts. This Part I proceeds with providing information on the purpose and research questions guiding the study. Part II: Research Design provides information on households selected for participation in the study, in addition to a description of the sample and the rationale underpinning its construction. Rounding out this part of the report is a discussion on the survey questionnaire and its evolution over the five rounds of data collection.

Next, Part III: Fieldwork and Processing details processes that bookend data collection, from contacting households between rounds, to logistics of fielding the survey and weighting data when it comes in. Importantly, Part III Sections 2 and 3 provide key information for data users on the intended use of the weights included with the dataset; the rationale that informed their calculation; and other possibilities for weighting the data.

Part IV: The Dataset and Codebooks discusses the construction of the panel dataset, which merges the responses of each IDP household to survey questions asked across all rounds data. Dataset users will also descriptions of information in individual codebooks, including variable naming conventions and missing data conventions applied during the data cleaning and management phase of the study.

Finally, Part V: Sample Composition provides information on the composition and sizes of the study in each round of data in addition to attrition rates across the five rounds of data collection.

### Section 2: Study Overview

Georgetown University and the International Organization of Migration (IOM) partnered to conduct a panel study of Iraqi households who were forcibly displaced by the arrival of Islamic State of Iraq and the Levant (ISIL) forces between January 2014 and December 2015. The project collects both survey and qualitative interview data from these families.

The purpose of the study is to understand the challenges that the non-camp population of Iraqi internally displaced persons (IDPs) face and the solutions they engineer to access durable solutions as defined by the Inter-Agency Standing Committee’s (IASC) *Framework on Durable Solutions Among Internally Displaced Persons*. The framework defines three “durable solutions”—sustainable return, sustainable reintegration, or sustainable resettlement— each of which depends on the fulfillment of eight criteria: long-term safety and security; adequate standard of living; access to livelihood and employment; access to effective and accessible mechanisms to restore housing, land, and property; access to personal and other documentation; family reunification; participation in public affairs; and access to effective remedies and justice.

IDPs are said to have reached a durable solution when they “no longer have any specific assistance and protection needs that are linked to their displacement and can enjoy their human rights without discrimination on account of their displacement.”<sup>1</sup>

The GU-IOM joint initiative surveyed study participants about each of IASC’s eight criteria and also asked questions related to migration history, stability, and social capital. The following inquiries guided the selection of questions for the study:

1. How do displacement and access to durable solutions among IDPs in Iraq change over time?
2. What are the needs, coping strategies, and aspirations of IDPs, and what events and factors are perceived to influence these needs, coping strategies, and aspirations over time?
3. To what extent do the experiences of IDPs in Iraq inform our conceptualization and operationalization of quasi-durable and durable solutions?

Five rounds of survey and interview data collection have been completed. The first round was collected in March and April 2016; the second in February and March 2017; the third in July and August 2017; the fourth in August to October 2018; and the fifth in October 2019 to January 2020.

In each round, the study gathers survey responses to questions asked of displaced families. To better understand how access to durable solutions changes over time, survey interviews were conducted with the same families in all rounds of the study. The research team also completed in-depth, semi-structured interviews with a subset of surveyed households, host community members, and other displacement-affected or interested populations. This document only provides information pertaining to the quantitative data collected during the mixed-method study.

---

<sup>1</sup> United Nations General Assembly. *Report of the Representative of the Secretary-General on the human rights of internally displaced persons, Walter Kälin: Framework on Solutions for Internally Displaced Persons*, A/HRC/13/21/Add.4, (9 February 2010), p. 1. Available from <http://www.un.org/Docs/journal/asp/ws.asp?m=A/HRC/13/21/Add.4>

## II. RESEARCH DESIGN

### Section 1: Respondents

Respondents participating in the panel study *Access to Durable Solutions Among IDPs in Iraq* are Iraqi IDP households displaced by ISIL from their homes between January 2014 and December 2015. At the inception of the study in 2016, these households were living in one of four governorates of displacement—Baghdad, Basra, Kirkuk, or Sulaymaniyah—and were members of the non-camp population of IDPs.

The unit of analysis for the study is an “IDP household,” where:

- An **IDP** is defined, per the 1998 Guiding Principles on Internal Displacement, as a person (or groups of persons) “obliged to flee or to leave their homes or places of habitual residence, in particular as a result of or in order to avoid the effects of armed conflict, situations of generalized violence, violations of human rights or natural or human-made disasters, and who have not crossed an internationally recognized state border.”

and

- A **household** is defined as the set of individuals living together (sharing a kitchen) on a permanent basis for at least three months prior to being displaced.

Beginning in Round 2, IDP households in the study began to move inside of Iraq. While the majority remained in the same geographic district to which they were originally displaced, some households returned to their districts of origin; some moved to districts other than the ones they had originally been displaced to without returning home.

The research team leveraged the panel design of the study to explore how geographic movement relates to attainment of durable solutions and to explore whether a more parsimonious treatment of IDPs might reveal intra-group variation that could inform more nuanced policies. The study team thus identified three different subgroups within the non-camp IDP population:

1. **IDPs:** Families who were forcibly displaced from their *districts* of origin and maintain residence in the same *district* of displacement reported in Round 1.
2. **Movers:** Families who are no longer in the same *district* they reported in Round 1 but also have not returned to their *districts* of origin.
3. **Returnees:** Families who have returned to *districts* of origin as reported in Round 1.

The number of households falling into each category is reported in Table 1 below.

Notably, household status is an operationally defined variable and not one that is internationally defined or recognized. Legally, all families included in all rounds of the survey are considered IDPs because the durable solutions framework definition of an IDP is rights-based rather than geography-based.

STATUS	ROUND 1	ROUND 2	ROUND 3	ROUND 4	ROUND 5
	N (%)	N (%)	N (%)	N (%)	N (%)
<b>IDP</b>	3852 (100)	3020 (81.1)	2883 (76.2)	2260 (62.2)	1,875 (54.1)
<b>MOVER</b>	-	250 (6.7)	275 (7.4)	275 (7.6)	296 (8.5)
<b>RETURNEE</b>	-	454 (12.2)	610 (16.4)	1100 (30.3)	1,292 (37.3)
<b>TOTAL</b>	<b>3852 (100)</b>	<b>3724 (100)</b>	<b>3718 (100)</b>	<b>3635 (100)</b>	<b>3463 (100)</b>

Household status, however, does affect the study in two important ways: First, the survey questionnaire began introducing questions that were specific to IDPs/Movers or Returnees, though the majority of questions were asked of all groups. Second, the status of the household informed the adopted weighting conventions and pursuant generalizability of the study findings. (See Part III: Fielding and Processing Section 2: Weighting for more detail.)

## Section 2: Sample Frame, Design, and Strategy

IOM’s Displacement Tracking Matrix (DTM), a system used to monitor and track the movement of IDPs, was the closest approximation of a sample frame and provided information on Iraqi IPD population parameters reported. Specifically, the DTM provided information on two geographic variables that were used in constructing the sample:

1. Governorate of Displacement: the physical location to which households were displaced
2. Governorate of Origin: the physical location where households were living just prior to being displaced.

Table 2 provides the IDP population distribution by governorate of displacement and governorate of origin as reported in the DTM in December 2015.

Governorate of Displacement	Governorate of Origin							TOTAL
	Anbar	Babylon	Baghdad	Diyala	Kirkuk	Ninewa	Salah al Din	
Baghdad	66286	1183	5295	2979	583	7643	10004	93973
Basra	434	20	42	65	124	679	384	1748
Kirkuk	20992	30	275	1360	16748	5418	16340	61163
Sulaymaniyah	14053	615	1322	3784	0	2265	1572	23611
<b>TOTAL</b>	<b>101765</b>	<b>1848</b>	<b>6934</b>	<b>8188</b>	<b>17455</b>	<b>16005</b>	<b>28300</b>	<b>180495</b>

The sample in *Access to Durable Solutions Among IDPs in Iraq* is stratified according to these two geographic variables—governorate of displacement and governorate of origin—with disproportionate allocation to the strata. All IDPs displaced during the ISIL-induced wave in Iraq originated from one of one of seven governorates of origin included in the sample: Anbar, Babylon, Baghdad, Diyala, Kirkuk, Ninewa, and Salah al Din. Though IDPs were displaced to multiple other governorates in Iraq, operational constraints alongside three main research-related

inquiries guided the research team’s decision to include only the four selected governorates of displacement. The first was IDP population density. Baghdad, Basra, Kirkuk, and Sulaymaniyah were collectively home 180,495 (34%) of all Iraqi IDP households displaced by ISIL. Second, each of these four governorates featured forms of variation that were important to capture in the study. The centrally located Baghdad governate, the most populous and urban in Iraq, is home to the capital. Basra in the country’s southernmost tip hosted IDPs displaced to several more rural areas. Sulaymaniyah is part of the Kurdistan Region of Iraq (KRI) and Kirkuk is a contested area. Finally, numerous other studies were already underway in other governates to which ISIL-displaced households resided, while those selected for inclusion were not as frequently included.

Combining both stratification variables, there are a total of 28 theoretically possible strata in the governorate of displacement/governorate of origin matrix, but one stratum, Sulaymaniyah\_Kirkuk, contained no individuals in the population. Table 3 lists the 27 strata in the sample.

Baghdad Anbar	Basra Anbar	Kirkuk Anbar	Sulaymaniyah Anbar
Baghdad Babylon	Basra Babylon	Kirkuk Babylon	Sulaymaniyah Babylon
Baghdad Baghdad	Basra Baghdad	Kirkuk Baghdad	Sulaymaniyah Baghdad
Baghdad Diyala	Basra Diyala	Kirkuk Diyala	Sulaymaniyah Diyala
Baghdad Kirkuk	Basra Kirkuk	Kirkuk Kirkuk	Sulaymaniyah Ninewa
Baghdad Ninewa	Basra Ninewa	Kirkuk Ninewa	Sulaymaniyah Salah al-Din
Baghdad Salah al-Din	Basra Salah al-Din	Kirkuk Salah al-Din	

The strata were disproportionally allocated to ensure that four groups with a very low selection probability were represented in the final sample. Given their extremely low numbers in the population, IDP families originally from Babylon living in Kirkuk or Basra and IDP families originally Baghdad and Diyala living in Basra were purposefully oversampled to ensure their inclusion in the study. Table 4 provides the allocation for the target sample, with the four strata with low selection probabilities shaded in yellow.

Governorate of Displacement	Governorate of Origin							TOTAL
	Anbar	Babylon	Baghdad	Diyala	Kirkuk	Ninewa	Salah al Din	
Baghdad	219	247	185	181	20	187	162	<b>1200</b>
Basra	73	10	21	33	62	64	137	<b>400</b>
Kirkuk	157	15	132	129	518	134	116	<b>1200</b>
Sulaymaniyah	252	128	212	207	-	215	186	<b>1200</b>
<b>TOTAL</b>	<b>700</b>	<b>400</b>	<b>550</b>	<b>550</b>	<b>600</b>	<b>600</b>	<b>600</b>	<b>4000</b>

In January 2016, survey field researchers enrolled 4,000 randomly selected IDP households in the study. Of them, 3,852 households participated in the Round 1 survey. The sample size and composition in each subsequent round is discussed in Part V: Section 2 Changes in the Sample Composition.



### Section 3: Survey Questionnaire

The quantitative questionnaire employed in *Access to Durable Solutions Among IDPs in Iraq* contains primarily closed-ended questions that attempted to turn into measurable indicators each of the Inter-Agency Standing Committee's eight criteria on access to durable solutions: safety and security, standard of living, livelihood and employment, housing land and property, personal documentation, family separation and reunification, participation in public affairs, and access to justice. Most rounds additionally included questions related to migration history, stability, social capital, and future intentions. Finally, some rounds included modules on special topics such as the experiences of female-headed households, mental and physical health, and agriculture.

In Round 1, the survey questions were categorized into 11 sections. Section 1 contains demographic information on each *household member* in the surveyed households. In contrast, the unit of analysis for all other sections of the survey is the IDP household. Sections 3 through 8, 10, and 11 ask questions about each of IASC's eight criteria for a durable solution. Sections 2 and 9 ask questions on migration history, stability, and social capital. The sections of the survey in Round 1 are as follows:

- Section 1: Roster of Household Members
- Section 2: Migration and Movement History
- Section 3: Employment, Livelihood, and Financial Security
- Section 4: Adequate Standard of Living
- Section 5: Long Term Safety, Security, and Freedom of Movement
- Section 6: Family Separation and Unification
- Section 7: Loss and Replacement of Documentation
- Section 8: Housing, Land, and Property
- Section 9: Social Capital
- Section 10: Preferences and Intention for Resettlement
- Section 11: Perception of Stability

Survey Section 1: Roster of Household Members was completed only in Rounds 1 and 4 of the study. Given that this section was the only one to collect information on all household members, including children under the age of 18, the section is not available for public release to protect the privacy and security of participating households. Instead, limited information from this section about the household head or about the individual answering the survey questions in the remaining sections of the survey have been merged into the publicly available dataset. Please see Part IV: The Dataset and Codebooks below for more information on this limited information.

## Section 4: Changes to the Questionnaire Over Time

In Round 2 onwards, sections 12 through 15 were added to the survey as follows:

- Section 12: Participation in Public Affairs (Rounds 2,3,4 and 5)
- Section 13: Effective Remedies, Including Access to Justice (Round 2,3,4, and 5)
- Section 14: Health Status- PROMIS Measures (Rounds 3 and 4 only)
- Section 14: Agriculture Module (Round 5 only)
- Section 15: Questions for Subpopulations (Round 5 only)

Some questions initially asked in Round 1 Section 9: Social Capital were moved in Round 2 onwards to Section 12: Participation in Public Affairs. Furthermore, some health-related questions from Section 14: Health Status-PROMIS (Patient-Reported Outcomes Measurement Information System) from Rounds 3 and 4 were moved to Section 4: Standard of Living in Round 5. Finally, beginning in Round 2, the survey included specific questions for study-defined subpopulations designated based on the location of IDP households. In Rounds 2 through 4, these questions were dispersed throughout the survey according to substantive themes. In Round 5, however, these questions were all combined into Section 15: Questions for Subpopulations.

Beginning in Round 4, questions that previously had numerous response categories (often upwards of eight categories) were revised to have fewer categories and facilitate analysis. These revisions were carried out in a way to ensure comparability across rounds of data collection, primarily by retaining the “Other” category in the response set. But while respondents were previously asked to specify “other” in an open-response question, they were no longer asked to do so starting in Round 4. Questions and response sets asked in each round of the survey can be found in the corresponding codebooks. Please see II. About the Dataset and Codebooks below for more information on this limited information.

While the study attempted to maintain the question number, question structure, and response categories of repeated questions as much as possible over the multiple rounds of the study, users are advised that in some cases, the number, structure, or response categories for the same question text did vary over time. **Whenever possible, codebooks provide notes alerting users to changes in questions repeatedly asked in multiple rounds of the study; however, users are strongly encouraged to rely on more than just question numbers when checking to see if the same or comparable questions were asked in more than round. Users also should verify the comparability of response categories and question structures of questions asked in multiple rounds.**

As conditions on the ground evolved and new trends were discovered in the data, new question batteries were added across the survey’s main sections. For example, while the baseline survey in Round 1 heavily focused on the immediate aftermath of displacement, the surveys in Round 4 and Round 5 featured questions that probed topics like living expenses, compensation, healthcare, and children’s education. Users can consult the codebooks for all available questions asked in each round.

### III. FIELDWORK AND PROCESSING

#### Section 1: Survey Fielding and Mode

In all rounds of data collection, IOM-Iraq’s field researchers fielded the quantitative survey in Arabic, Kurdish, or Turkmen to participating households. The language in which the survey was administered is documented in the dataset beginning in Round 4. Prior to each round of data collection, teams participated in extensive trainings to review best practices and to provide feedback on the questionnaire. Whenever possible, the same interviewer recontacted the same households in each round of data.

Interviews were conducted with IDP heads-of-households, who often had input from or consulted with a spouse. In some instances, particularly where the household head was female, a relative of the household head participated in the interview. Beginning in Round 3, the study documented the household member with whom the interview was primarily conducted and limited demographic information about this individual has been included in the dataset. Please see II. About the Dataset and Codebooks below for more information.

Each interview took between 30 and 60 minutes to complete, depending on family size in Rounds 1 and 4 when Section 1: Roster of Household Members was administered. Prior to implementing the survey in Round 1, the survey was pre-tested in March 2016.

Surveys and qualitative interviews were conducted exclusively face-to-face in Round 1 using Computer Assisted Personal Interviewing (CAPI). Beginning in Round 2, face-to-face interviews continued, but changes in on-the-ground circumstances mandated a number of interviews needed to be conducted on the phone using Computer Assisted Telephone Interviewing (CATI). Specifically, an increasingly large share of IDP households returned to their governorates of origin or moved to governorates inaccessible to IOM teams. Data for the mode of interview was not collected in Round 2 but was collected in subsequent rounds and is displayed in Table 5 below. A variable denoting the mode of survey distribution—in phone or in person—was added to the data file starting in Round 3.

	<b>In Person N (%)</b>	<b>By Phone N (%)</b>
Round 1	3852 (100)	-
Round 2	NA	NA
Round 3	3555 (95.6%)	163 (4.4%)
Round 4	3463 (95.3%)	172 (4.7%)
Round 5	2880 (83.2%)	583 (16.8%)

From the beginning of the study until mid-2018, between rounds and just prior to data collection, IOM Iraq’s field researchers used the TextIt System, a text messaging platform, to maintain monthly contact with IDP families participating in the study and to track their movement. A monthly mass text was sent to households to verify their then-current location. When households indicated they had moved, field researchers called to verify the new location and to update the

database. For costs incurred for their participation in the study, respondents were paid mobile phone credits.

From Round 4 to Round 5, the field researchers contacted participants every two months via phone calls to identify their locations and movements. After Round 5, the field teams contacted the families once in August 2020 to update their locations. Because Round 6 was done exclusively over the phone, teams updated the families' locations in parallel with data collection

## Section 2: Weighting

Sample weights representing the inverse probability of selection for each stratum in Table 2 have been calculated for each round of data collection. Weights for the stratified sample have been calculated according to the following scheme:

$$w_h = \frac{1}{\frac{n_h}{N_h}} = \frac{N_h}{n_h}$$

Where,

$w_h$  is the weight for stratum  $h$

$n_h$  is the sample size for stratum  $h$  in a given round (See Table 6 Below)

$N_h$  is the population reported in the DTM for stratum  $h$  (See Table 2 Above)

Beginning in Round 2, some households returned to their governorates of origin (study-defined Returnees). Others moved to governorates not covered by the original sample of the study (some of the study-defined Movers who moved to governorates other than Baghdad, Basra, Kirkuk, and Sulaymaniyah). These households could be weighted according to the scheme above because they no longer belong to one of the 27 initially defined sample strata. In this approach, which weights based on geographic location, these households no longer represent the population from which they were initially selected. As such, study-defined Returnee households and study-defined Mover households who fall into this category have been assigned a weight of 1.

**Consequently, the weights provided with the dataset and the ones employed in all the Yearly Reports, Thematic Reports, Policy Reports, and Presentations available on the study website allow survey findings to generalize the Iraqi IDP population that was displaced to Baghdad, Basra, Kirkuk, of Sulaymaniyah as of December 2015 and that stayed in one of those governorates for the duration of the study.** These reports always analyze study-defined Returnees separately and make clear the analyses apply only to the sample and not to a broader population. The research team recommends that the included weights can be used to generalize to this population only in two cases:

Case 1: The analysis is conducted using data only from Round 1, where all households have the study-defined household status of "IDP."

Case 2: The analysis is conducted using data that **excludes** returnees and movers who moved to governorates not included in the initial sample. For any analyses using data from Rounds 2, 3, 4, or 5 individually or in combination with each other, users should remove households that received a weight of 1.

Additionally, the research team recommends that the included weights *should not be used* when the aim of the analysis is to describe any differences between IDPs, returnees, and movers. These unweighted analyses can provide insight on intra-sample differences, but the results cannot be generalized to any population, particularly because of the disproportional allocation of strata and because of the prospective nature of the study.

### **Section 3: Other Weighting Considerations**

Appropriate weighting schemes for datasets on displaced or otherwise mobile populations are richly debated, and ways exist to calculate weights other than those used and provided in the dataset. All of the information on the strata and geographic locations of participating IDP households in the study are available to data users, who thus have the ability to calculate weights according to schemes of their choosing.

The IASC Framework effectually leaves open to interpretation theoretical decisions of how to weight data among IDPs. On the one hand, the Framework employs a rights-based definition of IDPs: IDPs remain IDPs until they, “can enjoy their human rights without discrimination on account of their displacement.” As such, other weighting approaches might treat the panel as if it were a cohort study, thereby shifting focus from the spatial dimension to the rights-related dimension of displacement and commensurately treating all participating households by their internationally and legally recognized status. On the other hand, despite rights-based approach to differentiating IDPs and non-IDPs, the Framework’s three codified solutions to displacement—return, reintegration, or resettlement—are all geographically-based, which guided the research team’s approach to weighting.

The employed method of calculating weights described in Section 2 above placed a premium on the geographic location of households for several reasons. First, pre-existing research that was confirmed by this study suggests that among non-camp displaced populations, housing represents one of the largest expenses incurred while in displacement. Though majorities must rent while displaced, they own properties in their areas of origin. Returning home thus makes available to returned households significant funds otherwise unavailable to those still living in displacement and thus produce measurable differences in standard of living. Second, familiarity with a region alongside the presence of expanded familial and social networks also potentially alleviates pressures of displacement by providing households help with finding jobs, childcare, or access to healthcare needs. Weighting all IDPs without regard for geography, an observable and observed source of variation, could thus bias results.

Despite these acknowledged differences between IDPs living in displacement and IDPs who had returned home, there was no existing population data to determine how many IDPs displaced from one district (their origin district) to a second district (their displacement district) thereafter either returned to their origin district or moved to another district. The absence of this population data that would allow for more accurate weighting coupled with potentially biasing variation led to the decision to analyze study-defined returnees separately from study-defined IDPs.

## IV. THE DATASET AND CODEBOOKS

### Section 1: Panel Dataset

The **PANELFILE** contains IDP households' responses to all survey questions from all rounds of the study. This file is merged by an identification number, the UniqueID, assigned to each IDP household in the study. Subsequent to Round 1, if a household did not participate in a given round, the data is entered as missing (-88) for all questions. In the **PANELFILE**, the **status** variable in Rounds 2 through 5 denotes not only households that are study-defined IDPs, movers, or returnees, but also, it denotes whether or not a household declined to participate or could not be reached in a given round.

Ahead of each merged round of data, the **PANELFILE** reiterates identifier variables that include the UniqueID for each household, the stratum from which the family was sampled, the study-defined household status (IDP/Mover/Returnee), the mode of the survey, and necessary information about the governorate of origin and governorate of displacement in the current and prior round of data collection.

Additionally, ahead of merged data from Rounds 1 and 2, where information about the specific household member responding the survey *was not* recorded, users will find information limited demographic information about the household head. This includes the household head's gender, age, education level, and marital status. Users also will find information on the total number of household members per household and the total number of household members under 18 years of age. For both Rounds 1 and 2, all the included demographic data, including household size, comes from data collected in Round 1 Survey Section 1: Roster of Household Members. As discussed above in Part II Section 3: Questionnaire, the full Roster is not released to protect study participants.

Similarly, ahead of merged data from Rounds 3, 4, and 5, where information about the specific household member responding to the survey *was* recorded, users will find the specific respondent's gender, age, education level, and marital status. The total number of household members and total number of household members under 18 years of age are also included for these Rounds 3 through 5, but this time, the data comes from the Round 4 Survey Section 1: Roster of Household Members. If households dropped out of the study after completing Round 3 but before Round 4, when the Roster was updated, the demographic information on the specific household member responding and on household size was taken from the Round 1 Roster.

## Section 2: Codebooks

While the PANELFILE merges all rounds of data into one dataset, there is one, separate codebook for each round of data. The codebooks contain the questions asked in each round. Each codebook has five sections as follows:

1. About the Codebook
2. Notes About Round X Codebook
3. Round X Variable Name and Question Index
4. Round X Variable List and Coding
5. Appendix I: Governorate and District Identification Codes

For the convenience of data users, the first two sections of each codebook reiterate information on missing data, weights, and round-specific notes that are contained in this document. The **Variable Name and Question Index** is a list of variable names and the corresponding question numbers and question text from the given round of the survey instrument. The **Variable List and Coding** provides values and value labels for each survey question. The Appendix in each codebook provides the list of codes and corresponding names of Iraqi governorates and districts.

## Section 3: Variable Naming Conventions

In each codebook, variables are entered according to this scheme:

### **variableName**

response type (single choice, single choice-TEXT, single choice-INTEGER, multiple choice, multiple choice-rank, multiple choice-unranked)

Question Number (QX.X). Text of survey question  
Value Value Label

Variable names roughly correspond to question order and numbers on the survey instrument. In each variable name, the first number corresponds to the section of the survey from which it comes (see survey design below). The second number is the question number within the section. Any number or letter that follows an underscore ( ) refers to a sub-question or to disaggregated answers. Finally, the final number in each variable name designates the round of data from which that variable comes. For example, variable **q21\_b1** refers to section 2, question 1, sub-question b of the Round 1 dataset.

Starting in Round 2, there were questions that were specifically asked of IDPs/Movers only and questions asked of Returnees only. These questions are clearly marked in the codebooks and the variable names have the appendage IDP and RT attached. For example, in Round 2, Question 2.5 was asked only of returnees, and the variable name is **q25RT2**. A variable without RT or IDP in the name indicates that the question was asked to all respondents regardless of household status. In Round 5, the questions that were for specific subpopulations of the survey were moved to Section 15 rather than being dispersed throughout the other sections of the survey.

#### **Section 4: Missing and Not Applicable Data**

Unless otherwise specified, missing data are denoted in one of two ways. Responses were coded -88 if the question was not applicable, and responses were coded -99 if the respondent did not know or refused to answer the question



## V. SAMPLE COMPOSITION

### Section 1: Retention Analysis

Of the 4,000 families initially targeted for inclusion in the study, 3852 participated in Round 1; 3,724 participated in Round 2; 3,718 participated in Round 3; 3,635 participated in Round 4; and 3,463 participated in Round 5. Of the 128 families who dropped out of the survey after Round 1, 110 did so permanently while 18 families returned for Round 3. Meanwhile, 24 families completed Round 1 and Round 2 but dropped out in Round 3. The drop out number in Round 3 is offset due to families who were included in Round 1; dropped out in Round 2; and returned in Round 3. Table 6 summarizes these findings and provides attrition rates for the rounds of survey collection.

<b>Round</b>	<b>Number of Respondents (N)</b>	<b>Number of Respondents Who Dropped Out After Round</b>	<b>Retention Rate (%)</b>
<b>Round 1</b>	3852	128	
<b>Round 2</b>	3724	6	96.68 (Round 1 to 2)
<b>Round 3</b>	3718	83	96.05 (Round 1 to 3)
<b>Round 4</b>	3635	172	94.37 (Round 1 to 4)
<b>Round 5</b>	3463	-	89.90 (Round 1 to 5)

## Section 2: Changes in Sample Composition

Table 7 below provides the numbers of families in each round by both governorate of displacement and governorate of origin. Please note that this table does not reflect the location of families in the given round. Rather, it provides tabulations of the number of households still in the study by the original sample stratum from which they were selected.

<b>Stratum</b>	<b>Target N</b>	<b>Round 1 N</b>	<b>Round 2 N</b>	<b>Round 3 N</b>	<b>Round 4 N</b>	<b>Round 5 N</b>
Baghdad_Anbar	219	180	173	172	169	167
Baghdad_Babylon	247	76	75	75	75	74
Baghdad_Baghdad	185	363	358	355	350	336
Baghdad_Diyala	181	156	145	143	135	125
Baghdad_Kirkuk	20	19	17	17	17	15
Baghdad_Ninewa	187	118	116	114	107	101
Baghdad_Salah al-Din	162	140	137	136	134	123
<b>BAGHDAD TOTALS (% of the sample)</b>	<b>1200 (30%)</b>	<b>1052 (27.31%)</b>	<b>1021 (27.42%)</b>	<b>1012 (27.22%)</b>	<b>987 (27.15%)</b>	<b>941 (27.17%)</b>
Basra_Anbar	73	140	137	137	137	135
Basra_Babylon	10	10	10	10	10	10
Basra_Baghdad	21	20	20	20	20	20
Basra_Diyala	33	15	14	14	13	10
Basra_Kirkuk	62	52	52	52	52	51
Basra_Ninewa	64	229	224	225	225	225
Basra_Salah al-Din	137	134	131	131	131	129
<b>BASRA TOTALS (% of the sample)</b>	<b>400 (10%)</b>	<b>600 (15.58%)</b>	<b>588 (15.79%)</b>	<b>589 (15.84%)</b>	<b>588 (16.18%)</b>	<b>580 (16.75%)</b>
Kirkuk_Anbar	157	157	154	156	153	151
Kirkuk_Babylon	15	15	15	15	15	15
Kirkuk_Baghdad	132	50	45	44	43	39
Kirkuk_Diyala	129	140	140	139	138	131
Kirkuk_Kirkuk	518	521	509	512	511	504
Kirkuk_Ninewa	134	102	94	95	95	95
Kirkuk_Salah al-Din	116	131	127	127	127	125
<b>KIRKUK TOTALS (% of the sample)</b>	<b>1200 (30%)</b>	<b>1116 (28.97%)</b>	<b>1084 (29.12%)</b>	<b>1088 (29.26%)</b>	<b>1082 (29.78%)</b>	<b>1,060 (30.61%)</b>
Sulaymaniyah_Anbar	252	258	243	240	231	226
Sulaymaniyah_Babylon	128	131	126	127	126	108
Sulaymaniyah_Baghdad	212	92	88	86	77	66
Sulaymaniyah_Diyala	207	224	217	215	200	174
Sulaymaniyah_Ninewa	215	161	152	152	144	136
Sulaymaniyah_Salah al-Din	186	218	205	200	200	172
<b>SULAYMANIYAH TOTALS (% of the sample)</b>	<b>1200 (30%)</b>	<b>1084 (28.14%)</b>	<b>1031 (27.69%)</b>	<b>1029 (27.68%)</b>	<b>978 (26.90%)</b>	<b>882 (25.47%)</b>
<b>SAMPLE TOTAL</b>	<b>4000</b>	<b>3852</b>	<b>3724</b>	<b>3718</b>	<b>3635</b>	<b>3,463</b>