8

Sampling for Representative Surveys of Displaced Populations

Ana Aguilera, Nandini Krishnan, Juan Muñoz,
Flavio Russo Riva, Dhiraj Sharma and Tara Vishwanath

1 Introduction

As of April 2018, the United Nations High Commissioner for Refugees (UNHCR) reported that an estimated 6.6 million Syrians were internally displaced within the country, and that over 5.6 million Syrians

A. Aguilera (✉) · N. Krishnan · D. Sharma · T. Vishwanath
World Bank, Washington, DC, USA
e-mail: aaguileradellano@worldbank.org

N. Krishnan
e-mail: nkrishnan@worldbank.org

T. Vishwanath
e-mail: tvishwanath@worldbank.org

J. Muñoz
Sistemas Integrales, Santiago, Chile
e-mail: juan.munoz@ariel.cl

F. R. Riva
São Paulo School of Administration, São Paulo, Brazil

had fled to seek refuge in other countries, of which around 8% were accommodated in camps.¹ In addition to these official figures, there were anywhere from 0.4 to 1.1 million unregistered Syrian refugees in Lebanon and Jordan, and an estimated one million Syrian asylum-seekers in Europe.² In effect, more than half of Syria's pre-war population has been forcibly displaced since the beginning of the Syrian civil war.

The Syrian crisis has caused one of the largest episodes of forced displacement since World War II and some of the densest refugee-hosting situations in modern history. Syria's immediate neighbors host the bulk of Syrian refugees: Turkey, Lebanon, and Jordan rank in the top five countries globally for the number of refugees hosted—according to UNHCR data, as of June 2018, Turkey hosted 3.5 million Syrian refugees, Lebanon 0.97 million, and Jordan 0.66 million. In fact, Lebanon and Jordan hold the top two slots for per-capita recipients of refugees in the world, at 164 and 71 refugees per 1000 inhabitants, respectively (UNHCR 2019).³ The influx into these countries has also occurred at a more rapid rate than prior refugee crises. At one point in the conflict, an average of 6000 Syrians were fleeing into neighboring countries every day.⁴ Beyond the immediate impact of inflow of refugees, the host countries are also dealing with other consequences of the

¹<http://www.unhcr.org/en-us/syria-emergency.html>.

²According to a 2014 background paper on Unregistered Syrian Refugees in Lebanon, from the Lebanon Humanitarian INGO Forum, “general estimates and media reports citing unnamed Lebanese officials put the number of Syrians living in Lebanon and not registered with UNHCR between 200,000 and 400,000, although the reliability of and sources for these estimates—which do not distinguish between those in need of protection and/or assistance and those not in need—are unknown” (Lebanon Humanitarian INGO Forum 2014). The paper cites a range of estimates (from around 10 to 50%) based on data from various sources, with differing coverage and survey periods. The 2015 Jordanian census estimated 500,000–600,000 more Syrians than the numbers registered with UNHCR.

³Since these figures are based on official UNHCR registration numbers, they do not reflect the unknown number of unregistered refugees, as already noted in footnote 2. At the end of 2014, the United Nations estimated that registered Syrian refugees represented 29% of the total population in Lebanon and 9.5% of the total population in Jordan. Areas with the largest number of Syrians, such as the Bekaa Valley in Lebanon, have seen much higher proportions of refugees to local citizens.

⁴Quoted by the UN High Commissioner for Refugees in a speech to the United Nations Security Council in 2013.

Syrian conflict, including the disruption on trade and economic activity and growth and spread of the Islamic State (also called ISIS) in Iraq. While the Kurdish Region of Iraq (KRI) hosts at least 200,000 Syrian refugees, the ISIS-induced displacement from neighboring parts of Iraq means that KRI is now hosting over 2.25 million displaced persons, equivalent to approximately 40–50% of its population.

While each neighboring country has received many Syrian refugees in both absolute and relative terms, that is where the commonality ends. Each country has responded to the influx in its own way, influenced by its previous experience of handling protracted displacement situations. Given its history of encampment of the displaced Palestinian population, Lebanon has refrained from setting up camps for Syrians. There is also understandable wariness and anxiety of the impact the influx may have in the delicate domestic political power-sharing equilibrium. In KRI, the influx of Syrian refugees overlaps with a significant number of Iraqi citizens seeking a safe haven from the ISIS militants. The refugees and internally displaced people (IDPs) are located both in camps and non-camps, with a very porous camp boundary that allows its residents to move freely and work outside the camp. At the time of the survey, Jordan had an explicit policy to house refugees in camps and few refugees have legal residency and/or work permits, although a significant majority of refugees had moved outside the camps.

Creating an evidence base to frame the policies for refugees in host environment requires a sampling methodology to select a sample that represents both the host and refugee populations. There are several challenges associated with conducting a representative survey of the host community population and the forcibly displaced. In all three settings we consider, a reliable and updated sampling frame for the resident population was not available.⁵ No sample frames existed for forcibly displaced populations as they were excluded from available national sampling frames. Databases maintained by humanitarian agencies for internal programming purposes are often incomplete and out of date.

⁵The last official population census in Lebanon was in 1932 and the available sampling frames were also considerably dated in Jordan and KRI.

The displaced also have high degree of mobility and they are often unwilling to speak to surveyors. In this context, and in similar contexts of forced displacement, the selection of a representative sample of hosts and the displaced becomes a major challenge to drawing credible inferences about their socio-economic outcomes.

In this chapter, we describe the strategies that had to be devised to overcome these challenges when designing the sampling procedure for the Syrian Refugee and Host Community Surveys (SRHCS), which were implemented over 2015–2016 in Lebanon, Jordan, and the Kurdistan region of Iraq.⁶ Section 2 describes the innovative use of available information to come up with a strategy for generating representative samples of host community and refugee households in the three settings. Section 3 presents the implementation of this strategy. Section 4 concludes by highlighting implementation challenges and drawing general lessons from our experience on sampling forcibly displaced populations.

2 The Innovation

In all three settings, the main challenge to implementing a survey that would yield estimates representative of the refugee and host community populations, was the lack of an updated or comprehensive sample frame, including for hosting populations and especially for displaced populations. In general, the latter were completely missing from existing national sample frames. None of the three countries had at the time, a recent population and housing census, duly updated for population growth and movement, which could have provided the frame to choose the survey sample for the hosting community.

Each of the three contexts presented different challenges. Lebanon and Iraq have both not had a census for several decades and existing sample frames were out of date at the time of the SRHCS. In Lebanon, information from this sample frame was not available at low levels of geographic disaggregation, while in Iraq, internal displacement of

⁶The survey was conducted to support analysis on impacts of the influx on local communities in the three settings (see World Bank 2018b).

millions of Iraqis had made existing frames obsolete. In Jordan, while census exercises are undertaken every decade, data from the most recent census was not available for the SRHCS, and we had to rely on a relatively outdated sample frame based on the 2005 census. Differences in the distribution of Syrian refugees across the three contexts implied a country-specific approach as well. In Lebanon, there were no refugee camps for Syrians; in Jordan, there were two main refugee camps for Syrians; and in Kurdistan, Iraq, Syrians as well as Iraqi IDPs lived in camps but were also free to move in and out.

Defining a sampling strategy to yield representative samples of hosts and displaced populations in this context involved two key innovations. The first was the creation of a sample frame feasible for household listing operations from large geographical divisions where it did not exist. This was the case in Lebanon and among the two largest refugee camps in Jordan. In Lebanon, cartographic divisions of the country were only available for large areas, and had to be segmented and subsegmented based on satellite imagery and dwelling counts to yield geographic areas small enough for listing. These segmentations attempted to divide the larger areas into equal population size subdivisions or segments, much the same way as enumeration areas are generated. Similarly, for the two largest refugee camps in Jordan, Zaatari, and Azraq, satellite imagery was used to divide the camps into mutually exhaustive and exclusive sampling units of roughly equal population size.

The second innovation was the use of available information from different sources on displaced population prevalence which were incorporated into the sample frames of host population prevalence. In most cases, this information was only available at a geographic level higher than the smaller sampling units used in the final frame. This data allowed for the estimation of known probabilities of selection. The first stage sample selection assumed these probabilities were uniformly distributed over the larger geographic area, and in the sampling units within that area. The household listing operation in the selected small sampling units was then used to update this known (albeit incorrect) probability of selection. In Lebanon and Kurdistan, auxiliary information on spatial distribution of refugees and IDPs available from the UNHCR and the International Organization for Migration (IOM),

was merged with the sampling frame. Subdistrict level refugee and IDP prevalence information was used to stratify subdistricts by intensity of prevalence: low, middle, and high. The sample was further stratified into subgroups of interest, depending on the context. In Lebanon, the survey was representative of the host community and the Syrian refugee population. In Kurdistan, the scope of the survey was expanded to include IDPs, so that the survey was representative of the host community, Syrian refugees inside and outside of camps, and IDPs inside and outside of camps.

3 Implementation

In what follows, we detail the sampling strategy for Lebanon, which was the most complicated, and then describe the strategy for the other two contexts.

Lebanon. Conducting a representative survey in Lebanon was especially challenging. The first difficulty was that, as of 2015, there was no recent or reliable sample frame, even for Lebanese households, as the last official population census was conducted in 1932. Typically, such a sample frame consists of the universe of enumeration areas in a country, with associated estimates of population. This meant that we had to construct our own sample frame by selecting a few Small Area Units (SAUs) and then conducting a full listing operation by visiting every household within the selected SAUs and collecting basic demographic and contact information. The second difficulty was that there was no available cartographic division of the country into geographic areas small enough to be the subject of a full listing operation, which could then serve as a sampling frame for the SAUs. Circonscription Foncières (CF) were the finest level of disaggregation available; CFs are generally too large to be listed as some have populations of over 100,000. Finally, there was no available sampling frame for Syrian refugees in Lebanon, which meant that we had to depend on UNHCR data on registered Syrian refugees, combined with the estimates of Lebanese population at the CF level. Given these challenges and time and budgetary constraints, the sample was selected in multiple (four) stages as described below.

3.1 First Sampling Stage

The sample frame for the first stage is the list of 1301 CFs published by the Council for Development and Reconstruction (CDR) in 2004 and the 2014 UNHCR registration database. Each CF is identified by way of its administrative affiliation—Kaza, Qadha, and Mohafza. The UNHCR database reports the total population in each CF, as well as the number of Lebanese and Syrian population in each.^{7,8,9} The CF cartographic boundaries are described digitally in a linked Geographic Information System shape file.

The CFs were sorted into three strata depending on their ex-ante prevalence of Syrian population, as follows:

- Low prevalence: where the Syrian population accounted for less than 20% of the total population;
- Medium prevalence: where the Syrian population accounted for between 20 and 50% of the total population;
- High prevalence: where the Syrian population accounted for over 50% of the total population.

Prevalence of Syrian refugees at the CF level was defined as the number of registered Syrian refugees from the 2014 UNHCR database divided by the sum of the number of registered Syrian refugees and the 2004 Lebanese population counts from the CDR database. The first columns of Table 1 show the distribution of the CFs into strata, as well as the population in each stratum, as per the UNHCR database.

⁷Lebanese population distribution by cadasters, supplied by CDR Shapefile (2002–2003); Population estimate of Lebanese 4 million referenced in the Lebanon Crisis Response Plan (LCRP) (UNHCR 2015).

⁸Total population of Syrian refugees as reported by the UNHCR registration database as of December 2014.

⁹Total population of Palestinian refugees in Lebanon (PRL) estimated between 260,000 and 280,000 (UNRWA-AUB 2010). Database provided the population distribution by camps and gatherings. In addition, the total population of Palestinian refugees from Syria is estimated to be 43,000 according to the UNRWA; UNHABITAT UNDP study on gatherings.

Our intention was to select 75 CFs in total. The decision of how to distribute them across the 3 strata faced the classical dilemma of whether to do it in proportion to the population of the strata, which would deliver nearly optimal estimates for the country as a whole, or to allocate the same sample size (i.e. 25 CFs) to each stratum, which would deliver estimates of nearly the same quality for each of them. Since both considerations were important for the 2015 SRHCS, we opted to do it in accordance to Markwardt's rule (also known as the '50/50 equal/proportional allocation'), which is generally considered a good compromise between the two extremes. The last three columns in Table 1 show the chosen allocation, the corresponding sample sizes (in number of households), and the expected maximum margins of error.¹⁰

Within each stratum, CFs were selected for inclusion with probability proportional to size (PPS), using the total population as a measure of size, and with implicit stratification by administrative units (Kaza, Qadha and Mohafza). Some of the large CFs were selected more than once. For instance, there were 34 selections made from among the 'low prevalence' CFs (as per Table 1), and one extremely populous CF (Chiyah, located in Mount Lebanon) was randomly selected three times. As a result, the 75 selections were drawn from 71 different CFs. Annex Table 1 shows the list of sampled CFs, where the last column indicates the number of times each CFs was selected in the sample (e.g. one, two or three times depending on each case).

¹⁰More precisely, the last column of Table 1 shows the maximum expected margins of error for the estimation of a household-level prevalence P (such as the percentage of households with children, the percent of households reporting illnesses, etc.) at the 95% confidence level. These are given by $ME = 1.96 [Deff P (1-P)/n]^{0.5}$, where n is the sample size and $Deff$ is the *design effect*, basically due to the tendency of neighboring households to behave similarly in regards the indicator being observed. The column was computed for $Deff = 2$ (a value found in practice for many indicators of interest) and $P = 0.5$ (for which ME is maximum).

3.2 Segmentation of Circonscriptions Foncières (PSUs)

Given that CFs are larger in size than typical census Enumeration Areas which are roughly of 200 households each, the majority of the selected sample CFs was too large to be manageable for implementing a complete household listing operation. For this reason, these large CFs were divided into ‘super segments’ and ‘segments’ of roughly equal size within each category, using total number of households as a measure of size. The number of households in each ‘super segment’ or ‘segment’ was estimated based on observation of height of buildings and estimated population density in each area in the 2015 ESRI World Imagery¹¹ and 2015 Google Earth imagery, combined with local knowledge of these areas.

Based on the estimated measure of size, only five CFs were considered to be too large in size and hence were selected for ‘super segmentation’. At a later stage, all CFs and ‘super segments’ were divided into ‘segments’ due to their large size.

3.3 Second Sampling Stage: Super Segmentation of Circonscriptions Foncières

In the second stage, the boundaries of the ‘super segments’ in each CF were drawn using the 2015 ESRI World imagery basemap. These boundaries take into account the total estimated household count, as well as natural boundaries such as major roads, rivers, and paths that can easily be recognizable by field teams during the listing operation and implementation of the household questionnaire.

Within each super-segmented CFs, the sample ‘super segments’ were selected with equal probability, based on the assumption that each ‘super segment’ is of roughly equal size. The number of ‘super segments’ selected within each CF was the same as the number of times the corresponding CF was selected in the first sampling stage. For instance, if a

¹¹Esri, DigitalGlobe, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AEX, Getmapping, Aerogrid, IGN, IGP, swisstopo, and the GIS User Community.

CF was selected three times in the first sampling stage, we selected three ‘super segments’ within this CF. Similarly, if a CF was selected only once or twice on the first sampling stage, we correspondingly selected one or two ‘super segments’ on the secondary sampling stage.

Annex Table 2 shows the list of ‘super segments’ within selected CFs, where the ninth column indicates the number of times each CFs was selected in the sample (e.g. one, two or three times depending on each case). The column headed ‘Prob 2’ shows the probability of selecting the ‘super segment’ within each CF.

3.4 Third Sampling Stage: Segmentation of Circonscriptions Foncières

In a third stage, the boundaries of the ‘segments’ were drawn for all CFs and selected ‘super segments’ within CFs. Similar to the process of ‘super segmentation’, boundaries of segments were drawn using the 2015 ESRI World imagery basemap. These boundaries also take into account the total estimated household count, as well as natural boundaries such as major roads, rivers, and paths.

Within each CF or corresponding ‘super segment’, the sample ‘segments’ were selected with equal probability, with the underlying assumption that each ‘segment’ is of roughly equal size. Annex Table 3 shows the list of ‘segments’ for all CFs, where the last column indicates the probability of selecting the ‘segment’ within each CF in the third sampling stage.

3.5 Fourth Sampling Stage

The sample frame for the fourth stage is the full list of all households in the sample CF segments. The listing operation consisted of a full enumeration of all physical structures in the area, with each physical structure being classified as a primary or secondary residential dwelling, commercial building, school, hospital, government office, etc. The listing operation collected information about the household occupying each residential dwelling, and each household was classified as either a Syrian refugee

household or a host community household. Care was also taken to record two households living in the same unit separately.¹²

To ensure the quality and completeness of the listing operation, enumerators relied on high-resolution paper maps identifying all buildings within each segment. Each building or structure was pre-assigned with a unique identifier. Enumerators then created a record for each residential unit and household following the protocol described in the 2015 SRHCS Manual of Enumerator. The 40 households to be visited by the 2015 SRHCS in each segment (with a target of 20 Syrian refugee and 20 non-Syrian refugee households in each) was selected from the listing data by systematic equal-probability sampling.¹³

3.6 Selection Probabilities and Sampling Weights

Given the sampling design discussed in the last paragraphs, the probability p_{hizsj} of selecting household $hijzsj$ in segment $hizs$ of super segment hiz in Circonscription Foncière hi of stratum h is given by:

$$p_{hizsj} = \frac{k_h n_{hi}}{\sum_i n_{hi}} \times \frac{t_{hi}}{T_{hi}} \times \frac{g_{hi}}{G_{hi}} \times \frac{m_{hij}}{n'_{hi}}$$

where the four fractions on the right-hand side respectively represent the probability of selecting the CF in the first stage, and the conditional

¹²One segment (in the Saida Ed-Dekermane CF, segment number 61119-0-26) was dropped from the original sample since the field team could not get access to the area due to insecurity and was thus unable to implement the household listing operation. Therefore, the intended sample of 40 household in this segment was distributed among two other similar segments, selecting 20 additional households in each. The selection of these two segments was based on the household listing data and local knowledge provided by the survey firm. The two identified segments are located in Saida Al-Qadima and Mazraa 2 (Beirut) and are similar to the Saida Ed-Dekermane segment in that they have: (i) a high share of Palestinian refugees; (ii) high density of urban population; and (iii) high poverty rate.

¹³After listing, only 15 households were found in segment 31116-11. Therefore, all eligible households were selected for interviewing (full census). The total sample size was reduced by 25, for a total 2975 sample households.

probabilities of selecting the super segment, the segment, and the household in the second, third, and fourth stages, and:

- k_h is the number of CFs selected in the stratum (the fifth column in Table 1),
- n_{hi} is the number of households in the CF, as per the sample frame (the column headed ‘population’ in Table 1),
- t_{hi} is the number of ‘super segments’ to be drawn in the CF, as per the first sampling stage (the column headed ‘No. super segments selected’ in Annex Table 2),
- T_{hi} is the total number of ‘super segments’ in the CF, as per the segmentation procedure (the column headed ‘No. of super segments’ in Annex Table 2),
- g_{hi} is the number of segments to be drawn in the CF, as per the second sampling stage (the column headed ‘n_segments to draw’ in Annex Table 3),
- G_{hi} is the total number of segments in the CF, as per the segmentation procedure in the third sampling stage (the column headed ‘n_segments per SSU’ in Annex Table 3),
- m_{hij} is the total number of households identified as Syrian refugees during the household listing operation;
- m_{hizsj} is the number of households selected in the segmented CF (with a target 20 Syrian-refugee and 20 non-Syrian-refugee households in this case); or $m_{hij} = m_{hij} + (40 - m_{hij})$;
- n'_{hizs} is the number of households in the segmented CF, as per the household listing operation.

To deliver unbiased estimates from the sample, the data from each household hij should be affected by a sampling weight (or raising factor) $whzsj$, equal to the inverse of its selection probability (i.e. $whzsj = phizsj - 1$).

Kurdistan. Much of the sampling procedure in Kurdistan resembled that of Lebanon, except for one important difference: unlike in Lebanon, the frame for the first stage sample existed in Kurdistan (albeit outdated), and a subset of the enumerations areas had updated population information from the 2012 IHSES survey (which did not take into account subsequent internal displacement). A subsample of the 2012

clusters was selected for our survey, followed by a comprehensive listing exercise to update the frame for second stage sampling. Four strata based on refugee and IDP prevalence were defined as following:

- Low Syrian prevalence (<5%) and Low IDP prevalence (<15%)
- Low Syrian prevalence (<5%) and High IDP prevalence ($\geq 15\%$)
- High Syrian prevalence ($\geq 5\%$) and Low IDP prevalence (<15%).
- High Syrian prevalence ($\geq 5\%$) and High IDP prevalence ($\geq 15\%$).

In the first stage, within each stratum, enumeration areas were selected with PPS using the number of households reported from the 2012 listing exercise as a measure of size. In the second stage, 18 households per PSU were selected: six Syrian households, six IDP households, and six host community households in each PSU to the extent possible. In areas where there were less than six Syrian or IDP households, the shortfall was met by host community households. The sampling frame for second stage sampling was the complete list of households in the selected EAs from the listing exercise.

Jordan. In contrast to Lebanon and Iraq, Jordan has carried out Population and Housing Censuses on regular intervals, with the last one in late 2015. What was particularly attractive about the latest census from the perspective of sampling was that it explicitly asked about the nationality of all residents. This would have allowed stratification of areas by density of Syrians. However, the original design could not be implemented because we could not access the new sample frame based on the 2015 Jordanian census. The design was then amended to include a representative sample of the Azraq and Za'atari camps (which account for the vast majority of Syrian refugees in camps in Jordan). This sample was complemented by purposive samples of the surrounding governorates, Mafrq and Zarqa, where the sample included areas physically proximate to the camp and other areas with a high number of Syrian refugees. In Amman Governorate, a purposive sample was drawn, combining a geographically distributed sample with a sample of areas with a high prevalence of Syrian refugees per the 2015 census, as indicated by the Jordanian Department of Statistics. Analytically, this implies the insights from Jordan will be limited to camp residents, neighboring areas of the camps, and Amman governorate.

4 Implementation Challenges, Lessons Learned, and Next Steps

The three surveys described in this paper were designed to generate comparable findings on the lives and livelihoods of Syrian refugees and host communities in the three settings. The absence of updated national sample frames and the lack of a comprehensive mapping of the forced displaced within these countries posed challenges for the design of these surveys. These challenges are not unique—indeed, most developing countries face similar issues, which are exacerbated at times of large scale internal population movements or in contexts of a large localized or widespread influx of migrants. Such data challenges become particularly stark in countries hosting displaced populations or in situations of ongoing or protracted conflict as local populations move to escape violence. But exclusion of displaced persons from national sampling frames, and consequently from national surveys, provides a skewed picture of the world (World Bank 2018a). As the number of displaced persons continues to increase, it becomes all the more urgent to devise strategies to include them in representative socioeconomic surveys.

This methodology paper describes the strategy implemented in the three contexts to generate known ex-ante selection probabilities through a variety of data sources, the use of geospatial segmenting to create enumeration areas where they did not exist, and to use data collected by humanitarian agencies to generate sample frames for displaced populations. The strategies implemented in these surveys can be useful in designing similar exercises in contexts of forced displacement. Moreover, this effort shows the importance of including refugees and non-nationals in national sample frames. The move by Jordan's statistical agency to explicitly include non-nationals in the 2017/2018 household survey is a commendable step in the right direction.

Annex

See Tables 1, 2, and 3.

Table 1 Syrian Refugee and Host Community Survey: sampling strata—Lebanon

Stratum	Prevalence	Sample frame		Syrian Refugee and Host Community Survey 2015		
		No. of CFs	Population	No. of selections	Sample size (HHs)	Margin of error (%)
1. Low prevalence	<=0.20	946	3,003,958	34	1360	3.76
2. Medium prevalence	0.21–0.50	273	1,039,171	24	960	4.47
3. High prevalence	0.51–1.00	82	465,867	17	680	5.31
Total		1301	4,508,995	75	3000	2.53

Table 2 List of selected segments (enumeration areas)—Lebanon

Segment serial number	CF CAS code	CF name	Qadha name	Mohafza name	Total Syrian population (combined CF)	Total population (combined CF)	No. of polygons	Prevalence of Syrians	Stratum 1-3	Prob 1	Times associated CF selected
1	10210	Msaïtbé foncière	Beirut	Beirut	3508	93,838	1	0.04	1	0.98263	1
2	10310	Mazraa foncière	Beirut	Beirut	12,410	125,792	1	0.10	1	1.31724	2
3	10310	Mazraa foncière	Beirut	Beirut	12,410	125,792	1	0.10	1	1.31724	2
4	10650	Achrafieh foncière	Beirut	Beirut	3108	71,541	1	0.04	1	0.74915	1
5	21111	Chiyah	Baabda	Mount Lebanon	50,085	251,061	1	0.20	1	2.62901	3
6	21111	Chiyah	Baabda	Mount Lebanon	50,085	251,061	1	0.20	1	2.62901	3
7	21111	Chiyah	Baabda	Mount Lebanon	50,085	251,061	1	0.20	1	2.62901	3
8	21177	Bourj El-Brajneh	Baabda	Mount Lebanon	24,065	139,404	1	0.17	1	1.45978	2
9	21177	Bourj El-Brajneh	Baabda	Mount Lebanon	24,065	139,404	1	0.17	1	1.45978	2
10	21219	Hadath Beyrouth	Baabda	Mount Lebanon	2702	26,829	1	0.10	1	0.28094	1
11	22111	Bourj Hammoud	El Metn	Mount Lebanon	18,456	94,232	1	0.20	1	0.98676	1
12	22155	Sinn El-Fil	El Metn	Mount Lebanon	3498	38,208	1	0.09	1	0.40010	1
13	22228	Baouchriyé	El Metn	Mount Lebanon	7317	72,611	1	0.10	1	0.76035	1
14	22359	Byaqout	El Metn	Mount Lebanon	346	3753	1	0.09	1	0.03930	1
15	22611	Broummana El-Matn	El Metn	Mount Lebanon	980	8844	1	0.11	1	0.09261	1
16	23469	Aain Zhalta	Chouf	Mount Lebanon	164	1910	1	0.09	1	0.02000	1
17	25111	Jounié Sarba	Kasrouane	Mount Lebanon	775	15,489	1	0.05	1	0.16219	1
18	25211	Aajaltoun	Kasrouane	Mount Lebanon	401	4554	1	0.09	1	0.04769	1
19	26141	Aamchit	Jubail	Mount Lebanon	791	14,288	1	0.06	1	0.14962	1
20	31116	Trablous El-Haddadine	Tripoli	North	1703	53,893	1	0.03	1	0.56435	1
21	31151	Trablous El-Qobbe	Tripoli	North	10,079	65,830	1	0.15	1	0.68935	1
22	32189	Bkeftine	Koura	North	77	881	1	0.09	1	0.00923	1
23	35179	Qboula	Akkar	North	4	616	1	0.01	1	0.00645	1
24	35487	Qbayat Aakkar	Akkar	North	568	6973	1	0.08	1	0.07302	1
25	51131	Zahlé Haouch El-Oumara	Zahle	Bekaa	29	5757	1	0.01	1	0.06028	1

(continued)

Table 2 (continued)

Segment serial number	CF CAS code	CF name	Qadha name	Mohatza name	Total Syrian population (combined CF)	Total population (combined CF)	No. of polygons	Prevalence of Syrians	Stratum 1-3	Prob 1	Times associated CF selected
26	53451	Haour Taala	Baalbek	Bekaa	198	3,478	1	0.06	1	0.03642	1
27	61119	Saida Ed-Dekermane	Saida	South	3	60,366	1	0.00	1	0.63213	1
28	61183	Miyé ou Miyé	Saida	South	2453	25,610	1	0.10	1	0.26818	1
29	61489	Aanqoun	Saida	South	645	5386	1	0.12	1	0.05640	1
30	62211	Jouaiya	Sour	South	467	7364	1	0.06	1	0.07711	1
31	62276	Aabbassiyet Sour	Sour	South	2171	14,082	1	0.15	1	0.14746	1
32	71236	Sarba En-Nabatieh	Nabatiye	Nabatiye	68	799	1	0.09	1	0.00837	1
33	72143	Aain Ibl	Bint Jubail	Nabatiye	153	2734	1	0.06	1	0.02863	1
34	74111	Hasbaliya	Hasbaliya	Nabatiye	575	8310	1	0.07	1	0.08702	1
35	22375	Dbayé	El Metn	Mount Lebanon	784	3268	1	0.24	2	0.05969	1
36	23211	Chhim	Chouf	Mount Lebanon	6067	19,616	1	0.31	2	0.35826	1
37	23321	Rmeilet Ech-Chouf	Chouf	Mount Lebanon	2351	4734	1	0.50	2	0.08646	1
38	24111	Choueifat El-Aamrousiyé	Aley	Mount Lebanon	19,572	73,031	1	0.27	2	1.33381	1
39	24133	Choueifat El-Quoubbé	Aley	Mount Lebanon	5843	26,791	1	0.22	2	0.48930	1
40	24343	Bayssour Aaley	Aley	Mount Lebanon	1706	8019	1	0.21	2	0.14646	1
41	31161	Trablous et Tabbaneh	Tripoli	North	6404	26,311	1	0.24	2	0.48053	1
42	32113	Kfar Aaqqa	Koura	North	923	3778	1	0.24	2	0.06900	1
43	33111	Zgharta	Zgharta	North	3218	15,813	1	0.20	2	0.28880	1
44	34269	Aabrine	Batroun	North	447	1753	1	0.25	2	0.03202	1
45	35275	Bebrine	Akkar	North	5301	18,073	1	0.29	2	0.33008	1
46	35364	Ouadi El-Jamous	Akkar	North	1619	5924	1	0.27	2	0.10819	1
47	37231	Beddaoui	Minieh-Danieh	North	16,976	44,404	1	0.38	2	0.81098	1
48	37271	Minie	Minieh-Danieh	North	17,610	38,905	1	0.45	2	0.71054	1
49	51133	Zahlé Aradi	Zahle	Bekaa	1232	6151	1	0.20	2	0.11234	1
50	51224	Jdita	Zahle	Bekaa	2990	9242	1	0.32	2	0.16879	1

(continued)

Table 2 (continued)

Segment serial number	CF CAS code	CF name	Qadha name	Mohafza name	Total Syrian population (combined CF)	Total population (combined CF)	No. of polygons	Prevalence of Syrians	Stratum 1-3	Prob 1	Times associated CF selected
51	52224	Baaloul BG	West Bekaa	Bekaa	871	2089	1	0.42	2	0.03815	1
52	53111	Baalbek	Baalbek	Bekaa	22,898	71,504	1	0.32	2	1.30592	1
53	53167	Saaidé	Baalbek	Bekaa	761	1647	1	0.46	2	0.03008	1
54	53311	Deir El-Ahmar	Baalbek	Bekaa	2924	7442	1	0.39	2	0.13592	1
55	53445	Nabi Chit	Baalbek	Bekaa	3094	9603	1	0.32	2	0.17539	1
56	61311	Ghaziyé	Saida	South	5163	18,290	1	0.28	2	0.33404	1
57	71113	Nabatiyeh El-Fauouka	Nabatiye	Nabatiye	2568	6905	1	0.37	2	0.12611	1
58	74122	Hebbariyé	Hasbaya	Nabatiye	780	2484	1	0.31	2	0.04537	1
59	24211	Aaramoun Aaley	Aley	Mount Lebanon	9827	15,666	1	0.63	3	0.50870	1
60	31111	Trablous Ez-Zeitoun	Tripoli	North	18,633	23,529	1	0.79	3	0.76402	1
61	35111	Halba	Akkar	North	10,842	16,668	1	0.65	3	0.54123	1
62	35429	Kouachra	Akkar	North	1958	3177	1	0.62	3	0.10316	1
63	35516	Mazareaa Jabal Akroum	Akkar	North	5965	11,487	1	0.52	3	0.37300	1
64	37317	Bqaa Seffrine	Minieh-Danieh	North	2224	4271	1	0.52	3	0.13869	1
65	51125	Zahlé Maallaqa Aradi	Zahle	Bekaa	6171	10,097	1	0.61	3	0.32786	1
66	51231	Saadnayel	Zahle	Bekaa	16,293	23,393	1	0.70	3	0.75961	1
67	51234	Qabb Elias	Zahle	Bekaa	27,951	39,206	1	0.71	3	1.27308	1
68	51267	Barr Elias	Zahle	Bekaa	34,688	45,306	1	0.77	3	1.47115	1
69	51284	Majdel Aanjar	Zahle	Bekaa	16,722	24,653	1	0.68	3	0.80052	1
70	51311	Riyayq	Zahle	Bekaa	6921	10,808	1	0.64	3	0.35095	1
71	52211	Joubb Jannine	West Bekaa	Bekaa	7833	13,478	1	0.58	3	0.43765	1
72	52234	Khiara	West Bekaa	Bekaa	1577	2004	1	0.79	3	0.06507	1
73	52277	Marij BG	West Bekaa	Bekaa	15,071	18,366	1	0.82	3	0.59637	1
74	61115	Saida El-Qadimeh	Saida	South	14,641	23,658	1	0.62	3	0.76821	1
75	61453	Bissariye	Saida	South	4931	8661	1	0.57	3	0.28124	1

Table 3 List of sample super segments (for CFs divided into super-segments or secondary sampling units)—Lebanon

SN	CAS_code	CF_name	Qadha_name	Mohafza_Na	Total_popu	Super segment ID	Segment ID	n_segments per SSU	n_segments to draw	Rand (TSU)	Prob 3
1	10210	Msaïtbé foncière	Beirut	Beirut	93838	10210-7	10210-7-13	18	1	0.02851	0.05556
2	10310	Mazraa foncière	Beirut	Beirut	125792	10310-1	10310-1-18	26	1	0.01869	0.03846
2	10310	Mazraa foncière	Beirut	Beirut	125792	10310-7	10310-7-6	17	1	0.08653	0.05882
4	10650	Achrafieh foncière	Beirut	Beirut	71541	10650-0	10650-0-66	93	1	0.00334	0.01075
5	21111	Chiyah	Baabda	Mount Lebanon	251061	21111-10	21111-10-34	41	1	0.02708	0.02439
5	21111	Chiyah	Baabda	Mount Lebanon	251061	21111-5	21111-5-9	23	1	0.04097	0.04348
5	21111	Chiyah	Baabda	Mount Lebanon	251061	21111-7	21111-7-19	22	1	0.08325	0.04545
8	21177	Bourj El-Brajneh	Baabda	Mount Lebanon	139404	21177-11	21177-11-1	14	1	0.03035	0.07143
8	21177	Bourj El-Brajneh	Baabda	Mount Lebanon	139404	21177-2	21177-2-9	23	1	0.00106	0.04348
10	21219	Hadath Beyrouth	Baabda	Mount Lebanon	26829	21219-0	21219-0-6	28	1	0.10421	0.03571
11	22111	Bourj Hammoud	El Metn	Mount Lebanon	94232	22111-6	22111-6-3	21	1	0.00019	0.04762
12	22155	Sinn El-Fil	El Metn	Mount Lebanon	38208	22155-0	22155-0-66	68	1	0.00901	0.01471
13	22228	Baouchiryé	El Metn	Mount Lebanon	72611	22228-0	22228-0-49	83	1	0.02951	0.01205
14	22359	Byaqout	El Metn	Mount Lebanon	3753	22359-0	22359-0-2	6	1	0.07392	0.16667
35	22375	Dbayé	El Metn	Mount Lebanon	3268	22375-0	22375-0-4	4	1	0.21483	0.25000
15	22611	Broummana El-Matn	El Metn	Mount Lebanon	8844	22611-0	22611-0-2	10	1	0.22362	0.10000
36	23211	Chhim	Chouf	Mount Lebanon	19616	23211-0	23211-0-5	21	1	0.09593	0.04762
37	23321	Rmeilet Ech-Chouf	Chouf	Mount Lebanon	4734	23321-0	23321-0-2	5	1	0.67365	0.20000
16	23469	Aain Zhalta	Chouf	Mount Lebanon	1910	23469-0	23469-0-1	2	1	0.47936	0.50000
38	24111	Choueifat El-Aamrousiyé	Aley	Mount Lebanon	73031	24111-0	24111-0-101	102	1	0.00238	0.00980
39	24133	Choueifat El-Quoubbé	Aley	Mount Lebanon	26791	24133-0	24133-0-11	29	1	0.09931	0.03448
59	24211	Aaramoun Aaley	Aley	Mount Lebanon	15666	24211-0	24211-0-11	18	1	0.06641	0.05556
40	24343	Bayssour Aaley	Aley	Mount Lebanon	8019	24343-0	24343-0-7	10	1	0.02895	0.10000
17	25111	Jounie Sarba	Kasrouane	Mount Lebanon	15489	25111-0	25111-0-20	22	1	0.05377	0.04545
18	25211	Aajaltoun	Kasrouane	Mount Lebanon	4554	25211-0	25211-0-1	5	1	0.09509	0.20000

(continued)

Table 3 (continued)

SN	CA5_code	Cf_name	Qadha_name	Mohafza_Na	Total_popu	Super segment ID	Segment ID	n_segments per SSU	n_segments to draw	Rand (TSU)	Prob 3
19	26141	Aamchit	Jubail	Mount Lebanon	14288	26141-0	26141-0-9	14	1	0.10108	0.07143
60	31111	Trablous Ez-Zeitoun	Tripoli	North	23529	31111-0	31111-0-13	48	1	0.01400	0.02083
20	31116	Trablous El-Haddadine	Tripoli	North	53893	31116-0	31116-0-11	54	1	0.01494	0.01852
21	31151	Trablous El-Qobbe	Tripoli	North	65830	31151-0	31151-0-42	44	1	0.00794	0.02273
41	31161	Trablous et Tabbaneh	Tripoli	North	26311	31161-0	31161-0-16	27	1	0.08705	0.03704
42	32113	Kfar Aaqqa	Koura	North	3778	32113-0	32113-0-1	4	1	0.10281	0.25000
22	32189	Bkefine	Koura	North	881	32189-0	32189-0-1	1	1	0.45403	1.00000
43	33111	Zgharta	Zgharta	North	15813	33111-0	33111-0-9	18	1	0.06386	0.05556
44	34269	Abrine	Batroun	North	1753	34269-0	34269-0-1	3	1	0.08812	0.33333
61	35111	Halba	Akkar	North	16668	35111-0	35111-0-15	19	1	0.02170	0.05263
23	35179	Qboula	Akkar	North	616	35179-0	35179-0-1	1	1	0.81850	1.00000
45	35275	Bebnine	Akkar	North	18073	35275-0	35275-0-3	21	1	0.04383	0.04762
46	35364	Ouadi El-Jamous	Akkar	North	5924	35364-0	35364-0-9	9	1	0.35237	0.11111
62	35429	Kouachra	Akkar	North	3177	35429-0	35429-0-3	3	1	0.22822	0.33333
24	35487	Qbaiyat Aakkar	Akkar	North	6973	35487-0	35487-0-4	7	1	0.01762	0.14286
63	35516	Mazareaa Jabal Akroum	Akkar	North	11487	35516-0	35516-0-5	11	1	0.18676	0.09091
47	37231	Beddaoui	Minieh-Danieh	North	44404	37231-0	37231-0-50	57	1	0.02521	0.01754
48	37271	Minie	Minieh-Danieh	North	38905	37271-0	37271-0-20	40	1	0.01934	0.02500
64	37317	Bqaa Sefrine	Minieh-Danieh	North	4271	37317-0	37317-0-4	4	1	0.44794	0.25000
65	51125	Zahlé Maallaqa Aradi	Zahlé	Bekaa	10097	51125-0	51125-0-4	15	1	0.19174	0.06667
25	51131	Zahlé Haouch El-Oumara	Zahlé	Bekaa	5757	51131-0	51131-0-4	6	1	0.12081	0.16667
49	51133	Zahlé Aradi	Zahlé	Bekaa	6151	51133-0	51133-0-5	7	1	0.01805	0.14286
50	51224	Jdita	Zahlé	Bekaa	9242	51224-0	51224-0-3	11	1	0.01322	0.09091
66	51231	Saadnayeel	Zahlé	Bekaa	23393	51231-0	51231-0-16	26	1	0.10708	0.03846
67	51234	Qabb Elias	Zahlé	Bekaa	39206	51234-0	51234-0-26	35	1	0.00073	0.02857

(continued)

Table 3 (continued)

SN	CAS_code	CF_name	Qadha_name	Mohafza_Na	Total_popu	Super segment ID	Segment ID	n_segments per SSU	n_segments to draw	Rand (TSU)	Prob 3
68	51267	Barr Elias	Zahle	Bekaa	45306	51267-0	51267-0-14	48	1	0.01760	0.02083
69	51284	Majdel Aanjar	Zahle	Bekaa	24653	51284-0	51284-0-13	25	1	0.01400	0.04000
70	51311	Riyah	Zahle	Bekaa	10808	51311-0	51311-0-2	11	1	0.07445	0.09091
71	52211	Joubb Jannine	West Bekaa	Bekaa	13478	52211-0	52211-0-1	14	1	0.01374	0.07143
51	52224	Baaloul BG	West Bekaa	Bekaa	2089	52224-0	52224-0-1	2	1	0.19555	0.50000
72	52234	Khiara	West Bekaa	Bekaa	2004	52234-0	52234-0-2	2	1	0.61762	0.50000
73	52277	Marj BG	West Bekaa	Bekaa	18366	52277-0	52277-0-8	20	1	0.13774	0.05000
52	53111	Baalbek	Baalbek	Bekaa	71504	53111-0	53111-0-70	80	1	0.01073	0.01250
53	53167	Saaidé	Baalbek	Bekaa	1647	53167-0	53167-0-1	2	1	0.57735	0.50000
54	53311	Deir El-Ahmar	Baalbek	Bekaa	7442	53311-0	53311-0-5	9	1	0.16490	0.11111
55	53445	Nabi Chit	Baalbek	Bekaa	9603	53445-0	53445-0-10	10	1	0.24514	0.10000
26	53451	Haour Taala	Baalbek	Bekaa	3478	53451-0	53451-0-2	3	1	0.23547	0.33333
74	61115	Saïda El-Qadimeh	Saïda	South	23658	61115-0	61115-0-16	25	1	0.08783	0.04000
27	61119	Saïda Ed-Dekermane	Saïda	South	60366	61119-0	61119-0-26	69	1	0.01328	0.01449
28	61183	Miyé ou Miyé	Saïda	South	25610	61183-0	61183-0-1	29	1	0.10490	0.03448
56	61311	Ghaziyé	Saïda	South	18290	61311-0	61311-0-5	19	1	0.00795	0.05263
75	61453	Bissariyé	Saïda	South	8661	61453-0	61453-0-6	9	1	0.10027	0.11111
29	61489	Aanqoun	Saïda	South	5386	61489-0	61489-0-3	5	1	0.19827	0.20000
30	62111	Jouaiya	Sour	South	7364	62111-0	62111-0-4	9	1	0.20830	0.11111
31	62276	Aabbassiyet Sour	Sour	South	14082	62276-0	62276-0-1	18	1	0.00890	0.05556
57	71113	Nabatiyeh El-Fauouka	Nabatiye	Nabatiye	6905	71113-0	71113-0-2	9	1	0.18614	0.11111
32	71236	Sarba En-Nabatieh	Nabatiye	Nabatiye	799	71236-0	71236-0-1	1	1	0.59953	1.00000
33	72143	Aain Ibi	Bint Jubail	Nabatiye	2734	72143-0	72143-0-1	3	1	0.32534	0.33333
34	74111	Hasbailya	Hasbailya	Nabatiye	8310	74111-0	74111-0-2	8	1	0.04804	0.12500
58	74122	Hebbariyé	Hasbailya	Nabatiye	2484	74122-0	74122-0-1	3	1	0.06554	0.33333

References

- Lebanon Humanitarian INGO Forum. (2014). *Background Paper on Unregistered Syrian Refugees in Lebanon*. Available at [http://lhif.org/uploaded/News/d92fe3a1b1dd46f2a281254fa551bd09LHIF%20Background%20Paper%20on%20Unregistered%20Syrian%20Refugees%20\(FINAL\).pdf](http://lhif.org/uploaded/News/d92fe3a1b1dd46f2a281254fa551bd09LHIF%20Background%20Paper%20on%20Unregistered%20Syrian%20Refugees%20(FINAL).pdf).
- The Data Blog. (2017). *A First Look at Facebook's High-Resolution Population Maps* [Online]. Available at <http://blogs.worldbank.org/opendata/first-look-facebook-s-high-resolution-population-maps>. Accessed 6 November 2017.
- UNHCR. (2015). *Lebanon Crisis Response Plan 2015–16*.
- UNHCR. (2019). *Global Trends: Forced Displacement in 2018*.
- UNRWA-AUB. (2010). *UNRWA-AUB Socio-Economic Survey of Palestine Refugees in Lebanon*. <https://www.unrwa.org/newsroom/press-releases/unrwa-aub-socio-economic-survey-palestine-refugees-lebanon>.
- World Bank. (2018a). *Poverty and Shared Prosperity Report: Piecing Together the Poverty Puzzle*. Washington, DC.
- World Bank. (2018b). *Syrian Refugees and Their Hosts in Jordan, Lebanon and the Kurdistan Region of Iraq: Lives, Livelihoods, and Local Impacts*. Unpublished Manuscript.

The opinions expressed in this chapter are those of the author(s) and do not necessarily reflect the views of the International Bank for Reconstruction and Development/The World Bank, its Board of Directors, or the countries they represent.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 3.0 IGO license (<https://creativecommons.org/licenses/by/3.0/igo/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the International Bank for Reconstruction and Development/The World Bank, provide a link to the Creative Commons license and indicate if changes were made.

Any dispute related to the use of the works of the International Bank for Reconstruction and Development/The World Bank that cannot be settled amicably shall be submitted to arbitration pursuant to the UNCITRAL rules. The use of the International Bank for Reconstruction and Development/The World Bank's name for any purpose other than for attribution, and the use of the International Bank for Reconstruction and Development/The World Bank's logo, shall be subject to a separate written license agreement between the International Bank for Reconstruction and Development/The World Bank and the user and is not authorized as part of this CC-IGO license. Note that the link provided above includes additional terms and conditions of the license.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.