

Schiefer, David; Düvell, Franck; Sağıroğlu, Ali Zafer; Mann, Lena

Article — Published Version

## Sampling refugee groups: The example of Syrians in Turkey

International Migration

**Provided in Cooperation with:**

John Wiley & Sons

*Suggested Citation:* Schiefer, David; Düvell, Franck; Sağıroğlu, Ali Zafer; Mann, Lena (2023) : Sampling refugee groups: The example of Syrians in Turkey, International Migration, ISSN 1468-2435, Wiley, Hoboken, NJ, Vol. 61, Iss. 5, pp. 231-247, <https://doi.org/10.1111/imig.13130>

This Version is available at:

<https://hdl.handle.net/10419/288104>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*


*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<http://creativecommons.org/licenses/by-nc-nd/4.0/>

ORIGINAL ARTICLE

# Sampling refugee groups: The example of Syrians in Turkey

David Schiefer<sup>1</sup>  | Franck Düvell<sup>2</sup> | Ali Zafer Sağiroğlu<sup>3</sup> | Lena Mann<sup>4</sup>

<sup>1</sup>German Center for Integration and Migration Research (DeZIM-Institute), Berlin, Germany

<sup>2</sup>Institute for Migration Research and Intercultural Studies (IMIS), University of Osnabrück, Osnabrück, Germany

<sup>3</sup>Ankara Yıldırım Beyazıt University (AYBU), Ankara, Turkey

<sup>4</sup>University of Jena, Jena, Germany

## Correspondence

David Schiefer, German Center for Integration and Migration Research (DeZIM-Institut), Mauerstraße 76, 10117 Berlin, Germany.  
Email: [schiefer@dezim-institut.de](mailto:schiefer@dezim-institut.de)

## Funding information

Bundesministerium für Familie, Senioren, Frauen und Jugend

## Abstract

Quantitative surveys among refugees not only advance our scientific understanding of forced migration, but they also provide political and civil society actors with the information they need to take measures to improve their protection and living conditions. These actors often need to acquire such information within a short period of time. However, the methods to access and survey this group to gain valid findings are subject to scientific debates. Based on a survey among Syrian refugees in Turkey, this article demonstrates a study design that generated valid knowledge about this population within a limited amount of time and resources. As a unique feature of the article, we compare our sample with four other sources of information to scrutinize how well it resembles the target population. As we can show, although point estimations differ to some extent, the internal distributions of sample characteristics are comparable across sources confirming the success of our sampling method.

## KEYWORDS

access, hard-to-survey populations, methodology, refugee surveys, sampling, Turkey

## INTRODUCTION

The number of refugees worldwide is greater than ever before (UNHCR, 2021). Quantitative surveys among refugees not only advance our scientific understanding of forced migration, they also provide political and civil society actors

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2023 The Authors. *International Migration* published by John Wiley & Sons Ltd on behalf of International Organization for Migration.

with valuable information which they need to take appropriate and proactive measures to improve their protection and living conditions and to predict and manage refugee movements (Jacobsen & Landau, 2003). Such knowledge is also politically relevant, for example, for international negotiations (Düvell, 2018). Given the relevance, survey findings can have it is necessary to base them on rigorous scientific methods that ensure the plausibility of findings. On the other hand, political developments are often rather dynamic and evoke questions to scientists that need to be responded to in a timely manner. However, surveying refugee populations poses numerous methodological, logistical and ethical challenges. They are therefore often considered examples of so-called 'hard-to-survey' populations (Tourangeau, 2014; but see Aljadeeh, 2022 for a critique of this term in the context of refugees) and the proper methods for surveying these groups within given time frames and available resources are subject to ongoing scientific debates (Bloch, 2007; Jacobsen & Landau, 2003; Müller-Funk, 2021; Reichel & Morales, 2017; Vigneswaran & Quirk, 2013). Based on a cross-sectional survey among Syrian refugees in Turkey, this article contributes to this debate by demonstrating a study design that generated valid knowledge about this population within a limited amount of time and resources. As a unique feature of the article, we compare the structure of our sample with four other sources of information to scrutinize how well our sample resembles the target population. Such a comparison is rare in the literature on sampling refugee populations yet important to appraise sampling strategies.

## BACKGROUND OF THE STUDY

Since the outbreak of the devastating violence in 2011, Syria has seen one of the largest refugee exoduses in world history. Although statistical numbers are somewhat inconsistent, estimates say that meanwhile up to 5.6 million Syrians have left the country due to the civil war (World Bank, 2019). Among these, around 4 million fled to Turkey. According to official numbers, 3.6 million Syrian refugees have been registered in Turkey until 2018 (DGMM, 2018). The country hosts by far more Syrian refugees than any other country in the world (Düvell, 2019; World Bank, 2019). Until 2015, Turkey had pursued an open-door policy towards Syrians granting them temporary protection. In March 2016, however, EU and Turkey arranged an agreement on joint effort to prevent refugees from moving on to Europe but also to improve the living conditions of refugees in Turkey. In addition, Turkey closed its southern border to reduce further inflows from Syria. As a result, both the migration flows of Syrians into Turkey and from Turkey to Europe have strongly decreased. Syrians in Turkey have formal access to education, health care and the labour market, and Turkey has been financially supported by the EU to accommodate and integrate them. However, their situation is still rather precarious (Baban et al., 2017) and Turkish politics and also the society have shifted from initial openness towards hostility (Erdogan, 2019).

Against this backdrop, there has been a high need for reliable knowledge about the living conditions of Syrians in Turkey as well as their aspirations to migrate onward, among political as well as civil society agents in Turkey and the EU, especially after 2016 when measures had been implemented to reduce onward migration and enhance the living conditions for Syrians in Turkey. The present study took up this need and investigated the living conditions and migration aspirations of Syrian refugees in Turkey. It was conducted by the German Center for Integration and Migration Research (DeZIM-Institute, Berlin) in collaboration with the Migration Policy Research Centre at Ankara Yildirim Beyazit University in Turkey and the Aalborg University in Denmark. The project was funded by the German Federal Ministry for Family Affairs, Senior Citizens, Women and Youth. All data collections were conducted with permission from the Directorate General of Migration Management in Turkey (DGMM). The study comprised a quantitative survey of around 1800 Syrians interviewed between November 2018 and March 2019 in six major Turkish cities.<sup>1</sup>

## SURVEYING SYRIAN REFUGEES IN TURKEY: METHODOLOGICAL CHALLENGES

The target population of our study were adult Syrian refugees living in Turkey. This comprises individuals aged 18 and older who had been permanent residents and/or citizens of the Syrian Arab Republic (in the following: Syria) before

the Civil war in Syria had started in 2011 and who had migrated or fled to Turkey between 2011 and the end of 2018 (before the beginning of data collection). Most of them have been registered as individuals under temporary protection by the Turkish government. In our attempt to gather valid knowledge about this target population by means of a survey we faced several challenges due to characteristics of the group, the conditions of the field and project-specific limitations.

Firstly, the official statistics on the target group are scarce. The only database that covers nearly the whole population is held by the Turkish Directorate General of Migration Management (DGMM). It covers all Syrian citizens who have been formally registered as persons under temporary protection in Turkey (DGMM, 2018; IOM, 2018d). The DGMM publishes up-to-date statistics only with regard to places of residence, age and gender. Furthermore, the information on places of residence is only available on province level and in some cases on city level but not on district level, and information on age and gender is only available on the national level. Apart from availability, the plausibility of such statistics – in terms of the size and regional distribution of the target group has been questioned (Düvell, 2019). Most importantly, however, the Turkish authorities do not grant access to this address database. This renders it impossible to construct a probability sample based on a person-based sampling frame.

Secondly, by the time of our study Syrians represented <5 per cent of the total population of Turkey (DGMM, 2018), yielding it as a rare population (c.f. Kalton, 2009). Thirdly, Syrians in Turkey are very unevenly dispersed across Turkish provinces. At the time of the survey, proportions varied from 0.02 per cent in Artvin to provinces such as Hatay (around 28%) or Kilis with an extremely high share of around 89 per cent (DGMM, 2018). Within provinces, Syrians often concentrate in certain cities, and within cities in certain districts. This applies also (or especially) to subgroups such as those who speak Kurdish, Turkish or other languages as mother tongue. If one wants to also include these, only a limited number of regions (which in turn have certain characteristics themselves) qualify as sampling units. This is again aggravated by the fact that for reasons of political sensitivity in Turkey, there is no official statistical information on these groups on the regional level, not even for the local population.

Fourthly, Syrians in Turkey live in several ways in precarious conditions. Despite a regular immigration status, they are large de facto excluded from legal employment and often live at the margins of society, physically and socially (for an overview see, e.g., Baban et al., 2017: 46). Furthermore, whereas Turkey initially welcomed Syrian refugees rather openly, by the time of the survey their presence had become a politically sensitive topic whilst the political climate in Turkey had generally become more oppressive (European Commission, 2019). The tense political atmosphere vis-à-vis refugees intensified during the time of the local elections in 2019 (Erdogan, 2019), which in turn coincided with the survey. All this required the research team to take certain precautions; for example, interviewers had to be trained as to not violate any legal requirements but also to consider mistrust and fear among potential survey participants and interdependencies and power dynamics vis-à-vis the research staff (Krause, 2017). Another problem that resulted from the precarious living conditions was that it was difficult to clearly identify domiciles and households.

Finally, the study received only a limited budget, which for administrative reasons had to be spent within a specified time frame of a few months. In addition, given that the living conditions and migration aspirations of Syrians in Turkey were politically highly relevant in 2018 (e.g., for EU migration policies), we attempted to come up with policy-relevant insights within as short a period as possible.

## SAMPLING AND SURVEYING HARD-TO-SURVEY POPULATIONS

In order to gain knowledge about a target group, the gold standard in quantitative social science would be to draw a probabilistic sample using techniques such as random or stratified random sampling that give every member of the population a chance to be drawn. That way findings obtained from the sample can be generalized to the total population. This requires a sampling frame, a list of all members of the population (individuals, households, etc.) based on which a sample can be selected. Once the selection probability for each participant is known, one can apply weights to correct deviations from the population structure (Levy & Lemeshow, 1999; Reichel & Morales, 2017). However,

as outlined in the previous section, such a sampling frame is not available for Syrians in Turkey and knowledge about the demographic structure of the population is scarce. Since this is often the case when doing research with forced migrants (Bloch, 2007; Reichel & Morales, 2017; Vearey, 2013), other sampling techniques have been developed that aim at still generating probability samples. Reviewing all these methods in detail goes over the scope of this report (for more detailed overviews see, e.g., Erens, 2013; Kalton, 2009; Marpsat & Razafindratsima, 2010; Reichel & Morales, 2017). However, a few techniques shall be sketched in order to embed our approach within the methodological landscape and to discuss them against the background of the specific conditions the project faced.

In absence of a sampling frame, one option is to randomly select households or individuals from the general population within a pre-defined space (country, a city, etc.) and first screen for members of the target group (Erens, 2013; Kalton, 2009; Reichel & Morales, 2017). This can be done, for example, through random dialling of phone numbers or random selection of addresses from an address register. Individuals or households identified as belonging to the target group (via screening questions) then constitute the sampling frame from which a random selection is subsequently made for the actual survey. Although specific techniques (e.g., focused enumeration) can increase efficiency such a procedure is still fairly time-consuming and costly, especially when the target group is small in size and unevenly distributed across regions (Erens, 2013; Kalton, 2009). This method was therefore deemed to be unrealizable in the present study. Also, using addresses for the screening sample was no option because official address databases were not accessible.

Given that population screening was unrealizable, we considered sampling techniques suitable for small and unevenly dispersed groups. A common approach is to focus on specific areas where the group is known to reside in high proportion. For example, one first randomly selects a number of regions (e.g., provinces), and within these selected regions one samples individuals and households proportionally to the size of the target population. Alternatively, one preselects a limited number of areas known to be inhabited by a high number and share of the target group. While this can be considered a purposeful selection of areas that creates some selection bias, principles of randomization and stratification can still be applied. For example, the population within these areas can be divided into strata and within each strata households can be randomly selected. Such a pre-selection is always a trade-off of the associated cost benefits versus the drawbacks in terms of sampling bias (Cibea et al., 2012; Erens, 2013; Kalton, 2009; McKenzie & Mistiaen, 2007). Nevertheless, if one aims at probabilistic samples that are representative for a small sub-population within a nation, the area-based approach still requires substantial effort, because the number of selected areas need to be sufficiently high and need to cover the whole country. In our case, we knew that we were only able to manage a survey in a limited number of provinces in Turkey within a limited amount of time. Furthermore, the approach requires knowledge on the exact prevalence of the target group in the regions. As we have explained above, such statistics are only available on the province level or (with regard to age) on the national level. Moreover, they are based on places of registration and there is evidence for high levels of irregular onward mobility among Syrians in Turkey which has raised concerns regarding the plausibility of the statistics (Düvell, 2019).

On the local level, various sampling techniques have been developed that aim at generating probabilistic samples in the absence of a proper sampling frame. One is random route (or random walk), a procedure where geographical starting points (e.g., a street intersection) are randomly chosen from which interviewers walk along a route based on a pre-defined walking scheme and contact households along the way (see, e.g., Ersanilli et al., 2011; Vearey, 2013). A further method is to randomly select buildings from regional maps, areal images or geo databases (see, e.g., Singh & Clark, 2013; Vearey, 2013). Within buildings, households and subsequently individuals can then be randomly selected. Yet another method is the so-called time-location sampling, where individuals are approached randomly at locations (e.g., public squares) which members of the target group are known to visit frequently (Baio et al., 2011; Marpsat & Razafindratsima, 2010; Muhib et al., 2001). All these local sampling techniques in principle allow to calculate selection probabilities and apply weighting procedures to the data, thus allow for unbiased estimations within an area. However, they require accurate information on the area, for example, on the number of buildings or of places where members of the target group gather, which, in turn, requires thorough mapping, extensive ethnographic fieldwork or knowledge of local experts (Reichel & Morales, 2017; Singh & Clark, 2013). Again, in the present study time and

budgetary restrictions limited our ability to fully exploit these methods. We were, for example, not able to collect the necessary information to estimate the selection probabilities of buildings, households and individuals.

Taken together, various techniques meanwhile exist that attempt to obtain probability samples of hard-to-survey populations.<sup>2</sup> In practice, however, the choice of the sampling strategy is always a matter of weighing the available options and restrictions. Many studies consider sampling options that at least partly abandon the claim of a random probability sample yet would still be methodologically rigorous and generate a sample that is “representative enough” (Vearey, 2013: 161) for the study’s purpose (see also Bloch, 2007; Jacobsen & Landau, 2003; Vigneswaran & Quirk, 2013). Researchers often decide to pursue multi-stage sampling approaches that combine probability with non-probability techniques on the national, regional and local level (see, e.g., Jacobsen & Landau, 2003; Misago & Landau, 2013; Singh & Clark, 2013; Vearey, 2013). As we will describe in detail in the following sections, this was also the case in our study.

## THE STUDY’S APPROACH

We developed a sampling strategy that comprised a combination of purposive selection and random sampling methods on the regional, local and household level. With this strategy, we aimed at gaining a sample that resembles the Syrian population in Turkey as adequate as possible despite the conditions and restrictions described in the previous sections. Figure 1 summarizes the selection and sampling stages, respectively. The first stage comprised an a priori exclusion of parts of the population. In the second stage, we selected provinces and cities within provinces, in the third stage, we selected districts and neighbourhoods and in the fourth stage, we sampled households and individuals.

### Stage 1: A priori exclusions

Among all 81 Turkish provinces, the 10 provinces with the highest share of Syrians together encompass around 79 per cent of the total population of Syrians under temporary protection in Turkey in December 2018 (DGMM, 2018). If one adds the unregistered Syrians assumed to live in Istanbul (Daily Sabah, 2019; Erdogan, 2019) this figure would be even higher. Even though there is little hard evidence — all available official data is on province level only — all sources seem to suggest more than 90 per cent lived in urban areas at this time, notably the respective provincial capital (Kiniklioglu, 2020). Furthermore, according to official statistics nearly all Syrian refugees are living in private accommodations (often improvised dwellings); only around 3 per cent reside in refugee camps (IOM, 2018d). Against the background of this uneven regional distribution and following an area-based sampling approach described above, those living in provinces with an overall Syrian population of less than 80,000 were a priori excluded from the sample. Also excluded were Syrian refugees held or hosted in open or closed camps. The main reason for this second exclusion was that access to the sites is highly restricted and was considered impossible by the team in Turkey. Although this step excluded parts of the target group that are worth being considered in terms of life circumstances, their exclusion was deemed justifiable given their small proportion of the overall target population. We identified 13 provinces which hosted 80,000 or more Syrian refugees. These hosted 3.1 million of the 3.6 million or 86 per cent of the Syrian refugee population in Turkey in December 2018 (DGMM, 2018).

### Stage 2: Selecting provinces and cities

Among the 13 remaining provinces, we made a theoretically driven yet pragmatic preselection of six provinces and nine cities, respectively, in several steps (a table showing characteristics of the selected cities and provinces can be requested from the first author):

- We first selected provinces based on their geographic location vis-à-vis the Syrian border and the EU, respectively. We identified three clusters of provinces hosting large numbers of Syrians (see UNHCR, 2022, for a map): those in

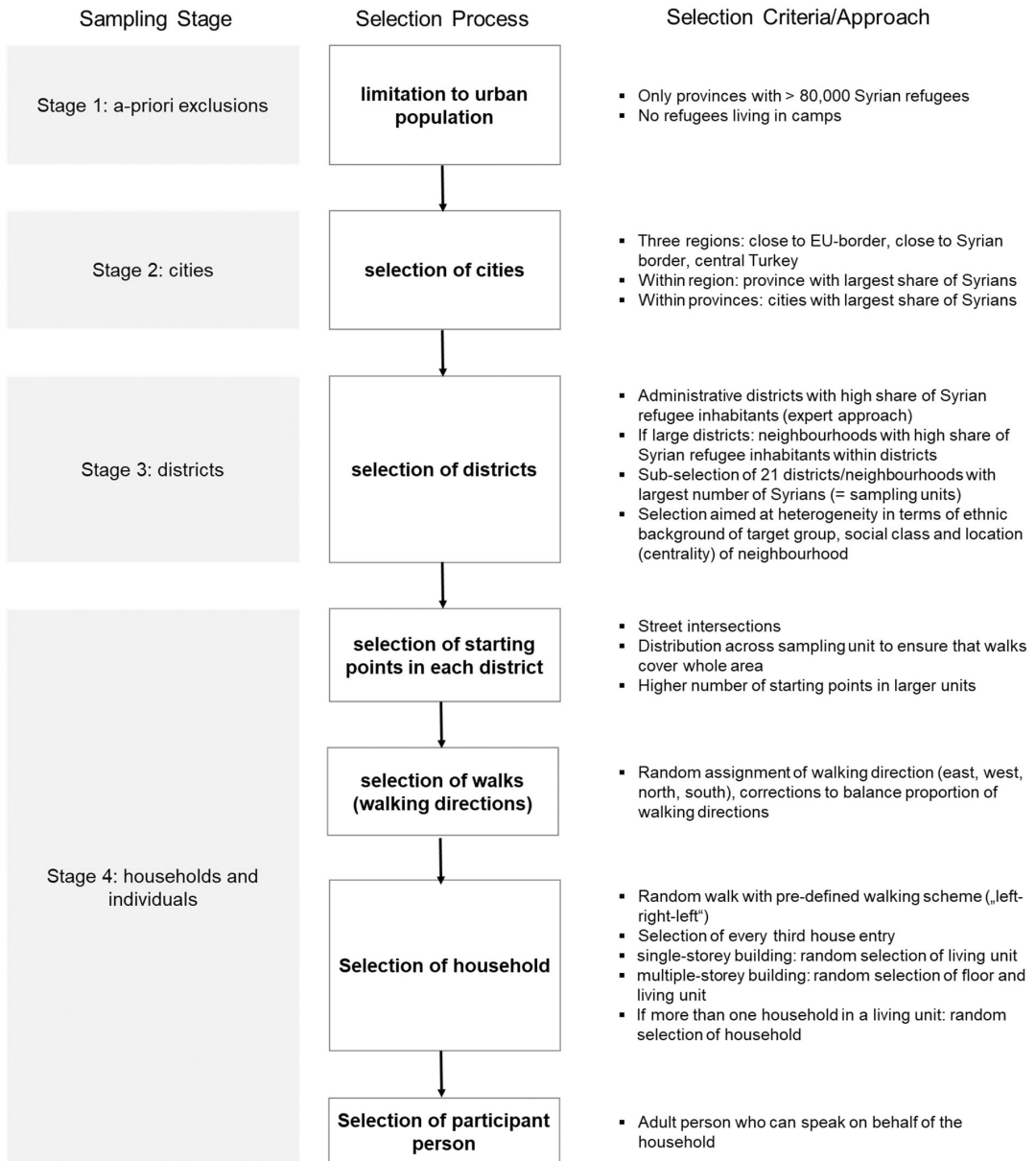


FIGURE 1 Multi-stage sampling procedure.

proximity to the Syrian border (Sanliurfa, Gaziantep, Kilis, Hatay, Adana, Mardin, Mersin, Kahramanmaraş), those in proximity to the EU (Istanbul, Bursa, Izmir) and those in-between in Central Anatolia (Ankara, Konya).

- Within these clusters, we chose the six provinces with the largest number of Syrian refugees (Istanbul, Ankara, Konya, Hatay, Gaziantep, Sanliurfa). These hosted a total of 2.07 million or at least 57 per cent of the Syrian refugee population in Turkey in December 2018 (DGMM, 2018).
- Within these six provinces we again chose the places where Syrians were known to be concentrated. This is usually the province capital which carries the same name as the province. In three provinces – Gaziantep, Hatay and Sanliurfa – there was more than one urban centre hosting Syrians. Therefore, we chose additional towns or regions within these provinces (Nizip, Reyhanli, Suruç).

The nine cities in the six provinces represent a theoretically sound and carefully selected subset which covers around 2 million or just under 60 per cent of the Syrian population in Turkey. There are no official statistics on age distribution at the provincial level. Thus, relying on the assumption that the proportions of the adult population in the provinces are similar to that on the national level, the six provinces cover around 1.9 million or approximately 60 per cent of our target group (adult Syrians in Turkey).

The provinces show similarities to other provinces in the same cluster, for example, with regard to infrastructures for onward migration (well developed in the north-west cluster) or return (better opportunities in the south-east cluster), populations (e.g., larger Arab-speaking community in the south-east cluster). Furthermore, given that Syrians primarily live in urban environments, the selected provinces (or the cities within the provinces) resemble other regions populated by Syrians in that they are primarily urban regions. Only in the provinces Hatay, Gaziantep and Sanliurfa, there is a larger number of Syrians in rural areas. We account for this by including smaller towns in rural areas there.

Apart from these similarities, our selection also made sure to cover different kinds of living environments for Syrian refugees in Turkey. Here we follow migration theory which suggests that migrants and refugees chose their destination due to its utility (Fafchamps & Shilpi, 2013) such as economic opportunities, infrastructures, social networks or geographic features. Accordingly, the sample includes Istanbul as a cosmopolitan “mega city” (see United Nations, 2018: 2) that is economically very attractive but also an important hub with a migration infrastructure facilitating onward migration to the EU. Besides Ankara as the capital and Konya as an ethnically less diverse industrial city, both in central Anatolia, it also includes provinces close to the Syrian border with traditionally close economic, family and cultural links with neighbouring Syria. The share of Syrians in these provinces also ranges substantially from <2 per cent in Ankara to 28 per cent in Hatay (DGMM, 2018), providing for different levels of community support. The six provinces also cover Syrians from different regions in Syria, according to expert ratings. Finally, the provinces are economically thriving though with diverse industries offering different opportunities in business and trade, textile, transport, service or agriculture.

### Stage 3: Selecting districts and neighbourhoods

The selected nine cities and towns altogether comprise 132 official districts (Turkstat, 2022) which constitute the sampling frame for the district level. Here, we again followed the approach of making a selection of districts based on the expected strike rate for members of the target group (Erens, 2013). To this end, we adapted a method previously applied in another study in Turkey (Ersanilli et al., 2011):

- We first identified those districts with a high concentration of Syrians. Because no official statistics on the local distribution of Syrian are available, we applied an expert-based selection approach (see, e.g., Kalton, 2009; Misago & Landau, 2013). We consulted local informants (local scholars of Turkish and Syrian origin) and searched for information on the internet, in newspapers or via municipal administrations. As a result, we compiled a list of 24 official districts. Because many of these had a large size and an unequal distribution of Syrian inhabitants, we chose one or more neighbourhoods (*mahalle* in Turkish) within these districts. Furthermore, in three cases two neighbouring small districts were merged into one. As a result, we obtained an overall number of 45 districts or neighbourhoods, respectively.
- Among these 45 districts/neighbourhoods, we selected 21 for the data-collection. We selected those with the largest proportion of Syrians within the city. Apart from the population share of Syrians, we made sure that districts/neighbourhoods represent different social classes or include Syrians of different mother tongues (classification based on expert ratings) and different geographical locations (central vs. peripheral districts/neighbourhoods). The latter was deemed relevant in terms of access to the housing and labour market infrastructure.
- If there were two or more districts with the same characteristics, we chose the one with the highest number of inhabitants.



Again, detailed information about the selected districts and neighbourhoods can be requested from the first author. Overall, since Syrian residences are concentrated in certain districts in each city, the selection strategy covered major parts of the Syrian population within these cities.

#### Stage 4: Sampling households and individuals

Within each selected district/neighbourhood, individual households were sampled based on a random-walk procedure (Hubbard et al., 2016; Reichel & Morales, 2017). Each walk was to comprise a batch of around eight interviews. We aimed at 250 interviews per city, thus, planned for 30 walks per city. Since the exact number and share of the Syrian population in districts was not known, it was impossible to assign the number of walks proportionally to the size of the population. We therefore started off by assigning equal numbers of walks to the districts but partially also adapted the number of walks according to the geographical size of the districts. The walks were prepared prior to the field work, using digital city maps. Starting points (usually street intersections) were selected evenly over the district area. Depending on the geographical size of the district, the number of starting points ranged from 3 (Gaziantep Yukaribayir) to 12 (Ankara Caldiran) with a mode value of four starting points. Depending on the number of starting points within a district, one to four walks were assigned to each starting point. Each walk had a pre-defined initial walking direction and a pre-defined walking scheme (including specific instructions, e.g., for streets leading out of a district) to ensure random sampling of addresses.

Interviewers were instructed to contact every third household they pass on their way. However, the project team knew from previous fieldwork in Turkey that Syrians often do not live in conventional houses or apartments but also share single rooms or live in basements, backyard sheds or tents or even open spaces between houses. We therefore conceptualized a household as a family or another group of people who share any delineable space as a dwelling. Hence, interviewers were instructed to approach every third place they pass and identify as a dwelling. Doors of businesses were not chosen. In case a building consisted of more than one living unit or storey, or when more than one household lived in a single apartment,<sup>3</sup> interviewers chose households following a prescribed randomization principle where possible. For each household contacted only one (adult) individual was interviewed. Interviewers asked the present household members who would be able to speak about the situation of the household as a whole.

### FIELD WORK

Budgetary restrictions did not allow to have a professional polling agency conduct the survey. In each city, the interviews were jointly conducted by a male and a female interviewer, one of Turkish, the other of Syrian origin. All interviewers received detailed training and instructions.

The cross-sectional survey took place between November 2018 and February 2019. In some cases, the walks did not result in the desired number of interviews, due to low density of Syrians or a general low house density. In these cases, additional walks in the same district/neighbourhood were prepared. In three cases (one in Ankara, one in Gaziantep and one in Konya) the district as a whole did not seem to render it possible to gather enough interviews within the target group, therefore, new walks were prepared in another district. In Konya, 100 additional interviews had to be conducted in May 2019 due to improper assessment of some questions in the previous interviews. The same walks like before were used but different households were contacted. Four additional walks were conducted in another district that was not part of the previous list. The new interviews were added to the original sample. Because no major changes of the conditions in Turkey were noted between February and May 2019, we assume that this did not add a bias to the sample.

Altogether, 236 walks in 23 districts were conducted. The number of walks per district ranged from two to 25 with an average of around 10 walks. On average eight interviews per walk took place. To take into account potentially varying working times of household members, the walks were conducted at different days of the weeks (including the weekend) and at different times of the day (see Misago & Landau, 2013). Around 30 per cent of the interviews were

taken in the morning (before 12 p.m.), around 40 per cent in noon time (12–2 p.m.) and around 30 per cent in the afternoon (after 2 p.m.). No interviews were conducted after sunset, this was due to security issues.

In case a contact attempt was not successful, the households were re-contacted twice at later times. If this was also not successful, the address was omitted. Households inhabited by a Turkish family were not interviewed. Among all Syrian households contacted, the response rates were rather high; they ranged from 78 per cent (Konya) to 95 per cent (Istanbul). In two cities, Gaziantep and Sanliurfa, interviewers did not record response rates properly but reported very high rates as well.<sup>4</sup>

## SAMPLE DESCRIPTION AND WEIGHTING

Table 1 documents the sample characteristics based on the raw data. According to the multi-stage selection approach described above, and provided that population fractions of adults are similar on the national and province level, the sample covers the adult Syrian population within the six provinces; around 1.8 million adult Syrians or 57 per cent of the adult Syrian refugee population in Turkey. As can be taken from the table, sample characteristics somewhat differ between provinces. Whereas the age distribution is very similar, the gender distribution is more balanced in Konya and Istanbul than in the other provinces. Furthermore, participants from Gaziantep are more likely to speak Turkish as their mother tongue than those from other provinces, and Kurdish is spoken more often by participants in Istanbul and Sanliurfa. The proportion of lower-educated participants (none or only primary school degree) is particularly high in Ankara and Gaziantep (60% and 59%), and the proportion of higher-educated participants (secondary or university degree) is highest in Hatay. On the regional level, participants from the central region (Ankara, Konya), are on average somewhat less educated than participants from the north-west (Istanbul) and south-east (Gaziantep, Hatay, Sanliurfa). Participants from the centre furthermore only rarely speak Kurdish or Turkish. Further regional differences (not documented in Table 1) regard, for example, to aspirations to migrate onward to Europe. These are much stronger among participants in the north-west (Istanbul; 47.4%) compared to the south-east (12.6%) and the centre (9.2%). This was to be expected given that Istanbul offers a much better infrastructure for onward migration

TABLE 1 Sample characteristics.

	Sample characteristics						
	Total	Ankara	Gaziantep	Hatay	Istanbul	Konya	Sanliurfa
N <sup>a</sup>	2,065,841	87,553	418,802	441,913	558,115	103,915	455,543
n	1886	297	305	289	315	390	290
Female (%)	59.0	65.3	65.9	61.5	53.0	48.7	62.9
Age (mean, SD)	35.6 (12.1)	35.3 (12.0)	34.6 (11.6)	37.3 (12.1)	35.2 (11.3)	34.8 (12.7)	36.7 (12.4)
Arabic mother tongue (%)	83.9	99.3	76.1	99.0	72.3	98.7	54.3
Kurdish mother tongue (%)	12.0	0.3	5.2	0.3	23.9	0.3	45.7
Turkish mother tongue (%)	4.0	0.3	18.7	0.7	3.8	1.0	0.0
Educational level							
No school degree (%)	12.3	9.5	12.8	8.3	7.9	16.2	17.9
Primary school (%)	36.8	50.7	46.2	25.6	34.6	34.9	29.0
Middle school (%)	30.3	21.0	27.9	33.2	35.6	31.0	32.8
Second school (%)	10.3	13.5	6.9	15.9	8.6	8.0	10.3
University degree (%)	10.3	5.4	6.2	17.0	13.3	10.0	10.0

Note: Unweighted data.

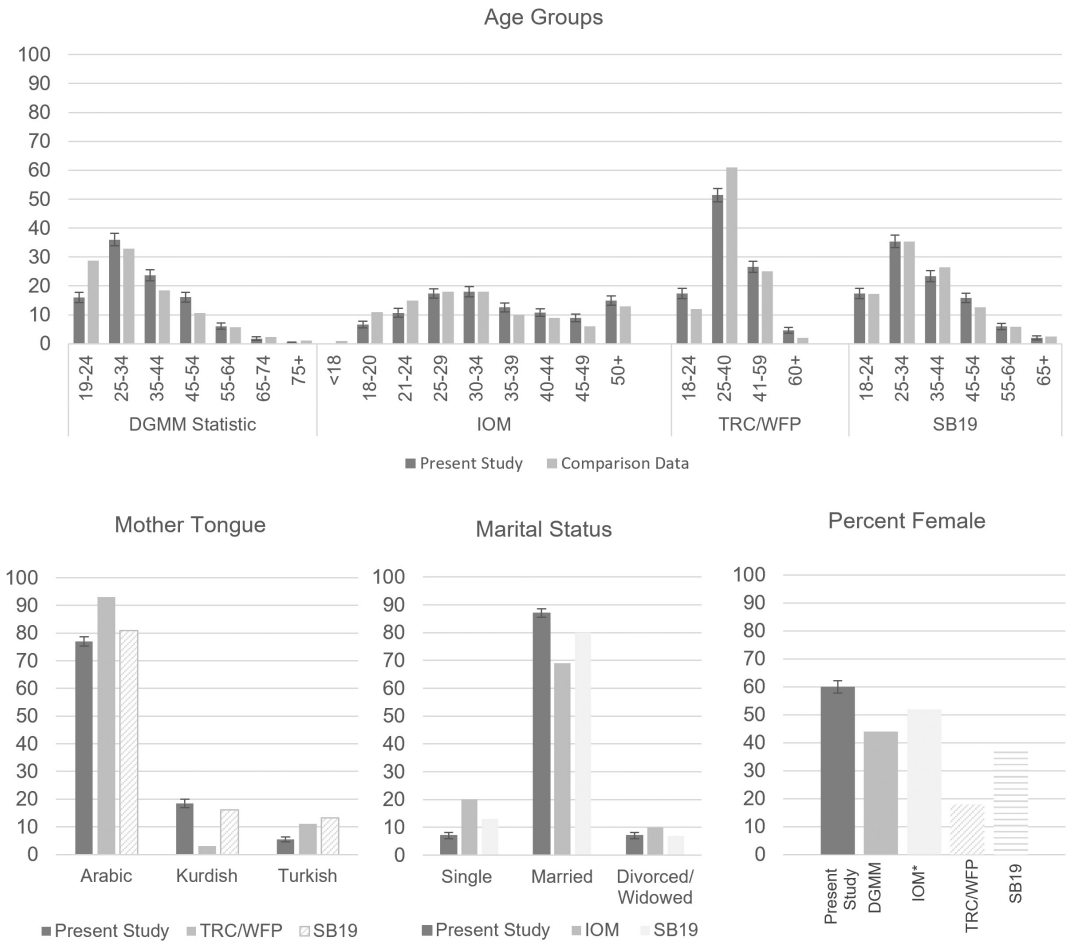
<sup>a</sup>Population numbers taken from DGMM (2018). Numbers refer to all ages. Information on the adult population on the province level is not available.

and thus attracts individuals who aspire that. Regional differences also exist regarding participants' satisfaction with their living conditions. For example, participants from the north-west are more likely to be satisfied with their chances to find a job (37.1%) than those in the centre (26.1%) and the south-east (20.1%).

These regional differences are relevant when considering the fact that the Syrian population is unequally distributed across the six provinces, whereas our sub-samples of each province are rather similar in size (around 300 participants per province). For example, the 315 participants from Istanbul represent around 560,000 Syrians under temporary protection registered in Istanbul, whereas the 297 participants from Ankara represent only around 90,000 (DGMM, 2018). Consequently, Syrians from certain provinces (e.g., Ankara) are over-represented in the sample, while those from other provinces (e.g., Istanbul) are under-represented, compared to the actual population fractions in the six provinces. To correct for these unequal population fractions, a design weight was computed as the inverse of the ratio of sample and population size within provinces (for details see, e.g., Lohr, 2019). Given a sample size  $n$  and a population size  $N$  in the province the sample-population-ratio  $\pi$  in province  $P$  reads  $\pi_P = \frac{n_P}{N_P}$ . Accordingly, the design weight for person  $i$  in sample  $P$  reads  $d_{iP} = \frac{1}{\pi_P}$ . The resulting weights for each person are then applied to the raw data when estimating mean values, proportions or other parameters. That way, the sample more accurately resembles the population fractions of Syrians in the six provinces. Note, however, that the plausibility of official numbers on Syrians in Turkish provinces have been questioned (Düvell, 2019, see also the introduction of this report). Thus, the weighting procedure involves some uncertainties.

## SAMPLE EVALUATION

In order to examine how well the sample resembles the Syrian population within Turkey, we compare it with data from four other sources. One is the official statistic on Syrians registered under temporary protection by the Turkish Authorities (DGMM, 2018). This source provides distribution of gender and age among the target population on the national level. Data from December 2018 were used, the month during which the majority of interviews were conducted. In addition, we compare our sample to those of three other surveys among Syrian refugees. These were conducted in approximately the same time period and attempted to cover the Syrian population in Turkey as comprehensively as possible. Facing the same challenges in terms of access to the target group, they chose different sampling strategies. Compared to our study, they were equipped with better resources (e.g., in terms of preparatory field work or access to sampling frames). The first study is the Flow Monitoring Survey conducted by the International Organization of Migration (IOM) between December 2017 and February 2018 in 14 provinces in Turkey (IOM, 2018b). This survey is part of IOM's Displacement Tracking Matrix (DTM) and comprised around 3000 refugees who had arrived in Turkey in 2016 or later, 28 per cent of the participants were of Syrian origin. The provinces were selected based on information on population characteristics taken from a baseline assessment (part of the DTM) which combines official statistics as provided by the DGMM with those gathered by IOM through interviews with local experts in villages and neighbourhoods (see IOM, 2018a). This combination of statistical information and local field work, which also tracks local mobilities, was assumed to provide a rather accurate basis for sampling (IOM, 2018a). According to the report, participants were sampled proportional to the size of the Syrian population in the provinces (IOM, 2018b: 25). On the local level, participants were approached in public places or via house visits (IOM, 2018c). The survey sample was created "in a manner which ensured maximum equal representation of parameters such as age, gender, geography, and education" (IOM, 2018b: 25). The second study is the Syrians Barometer 2019 which comprised a survey among 1400 Syrians in 15 cities (Erdogan, 2020). The sample is described as "the most comprehensive study in its field" (2020: 46) and as being diverse in terms of regional distribution, socio-economic status, gender and age. Cities in border regions as well as metropolitan (e.g., Istanbul, Ankara, Izmir) and non-metropolitan cities were selected. The report does not provide details on the selection principles (random vs. non-random) on the province and local level or details on weighting procedures. It states that the survey findings can only be generalized to the Syrian population in Turkey in a limited manner (Erdogan, 2020: 46). The third study was conducted by the Turkish Red Crescent and World Food Programme (TRC & WFP, 2019). The researchers were granted access to a data basis that lists all applicants to the Emergency Social Safety



**FIGURE 2** Distribution of demographic characteristics in the present sample and comparison data. *Note:* Bar represent proportions in percent. Vertical brackets indicate the 95%-confidence-intervals for the estimates based on the present data. In the present data age categories were generated in different ways depending on the respective comparison data. \*The IOM study differentiates between adult female (52%), adult male (47%) and children (1%). *Sources:* Present study (weighted data), DGMM (2018), IOM (2018b), TRC/WFP (2019), Erdogan (2020).

Net (ESSN), a program that provides financial support to refugees in Turkey since 2016. The researchers selected 19 provinces in three regions in Turkey (north, south, west), and within these provinces, they drew a random household sample from the ESSN application list. Population weights were calculated for analyses at the national and province level. The sample is thus representative for ESSN applicants in the 19 provinces, covering around 413,000 households or 2.4 million Syrians (TRC & WFP, 2019: 7–8).

A comparison of the three survey studies with the DGMM statistics reveals that the age and gender distribution of the IOM data show only minor deviations from the DGMM statistics. The TRC/WFP sample, in turn, comprises a much lower proportion of females, and in both the TRC/WFP and the Syrian Barometer sample participants are somewhat differently distributed across age categories than the DGMM statistics. For example, those from the lowest age category (18–24) are comparably under-represented (Figure 2).

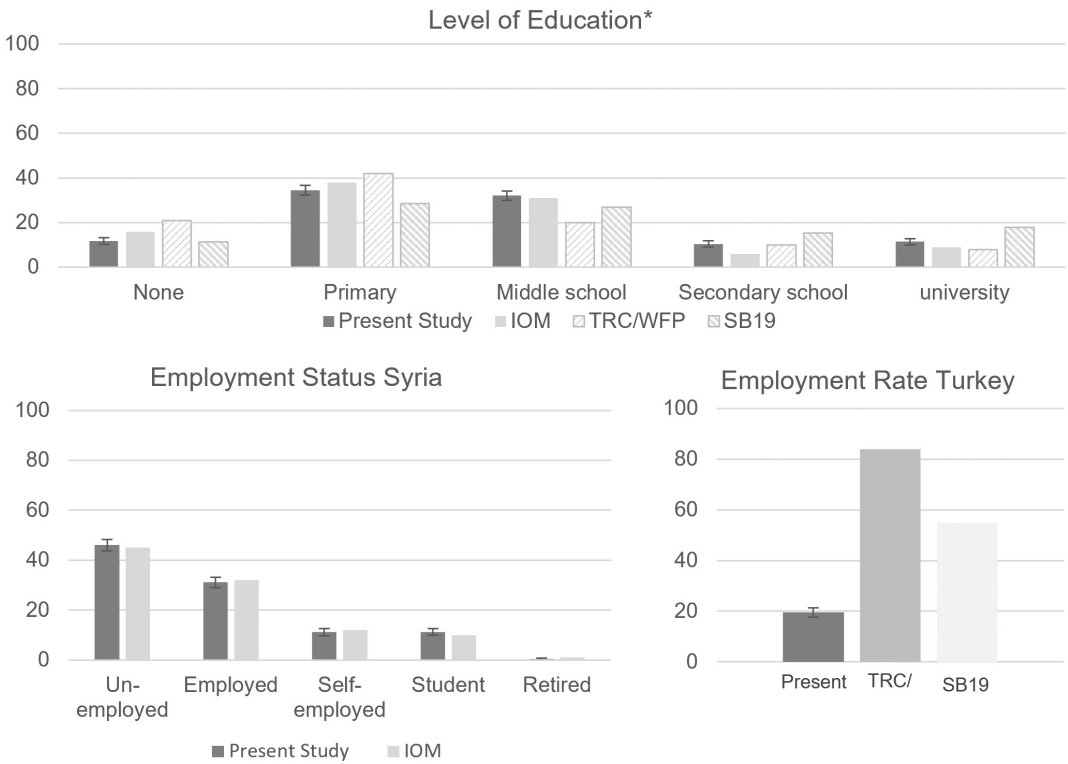
When comparing the present study's sample to the other data sources, the sample weight described in the previous section is applied. Furthermore, we selected only those sample characteristics that are comparable to the other sources in terms of content and question format (e.g., response categories). This primarily regards to

demographic characteristics, but we were also able to compare two attitudinal measures. Not all characteristics, however, were available from all studies. Regarding the age range covered, the data sources are comparable. The TRC/WFP study and the Syrian Barometer surveyed adult participants only, as was the case in our study. The IOM study additionally surveyed a small group of underage participants (1 per cent of the sample) for which results are not reported separately in the study's report. The DGMM statistic covers all ages. Here, we consider only individuals older than 18.<sup>5</sup>

Overall, the comparisons reveal striking similarities but also certain deviations between sources. For example, age is very similarly distributed in all data sources (Figure 2) with only a few more substantial deviations regarding particular age categories in the TRC/WFP study (age group 25–40) and the DGMM statistic (age group 19–24). Gender proportions, in contrast, vary to a large extent. The DGMM gender statistics are likely to be the most accurate. In the present sample, the proportion of female participants is much higher compared to the DGMM statistics, and the Syrian Barometer and the TRC/WFP data comprise much lower proportions (Figure 2). These differences are most likely due to different recruitment strategies. In the present study, we interviewed person who were at home and who felt eligible to speak on behalf of the household. Due to predominant patterns of economic behaviour and cultural habits men might have been more likely to be at work and thus not at home or spend time outside the household at the time our interviews were conducted (see Harvey et al., 2013; UNHCR, 2014). In the TRC/WFP, in turn, the interview was conducted with the “primary breadwinner” of a household (TRC & WFP, 2019: 8). Regarding marital status, the distributions are rather similar, however, the IOM sample comprises less married participants. The TRC/WFP and Syrian Barometer furthermore sampled a higher proportion of participants with Turkish as their mother tongue, and the TRC/WFP study a smaller proportion of Kurdish participants. Length of stay in Turkey distributes among participants of the TRC/WFP in a very similar way like in the present study (no figure).

Regarding socio-economic status, the distributional patterns deviate to some extent (Figure 3). The TRC/WFP sample comprised more participants with a lower level of education. The Syrian Barometer, in contrast, surveyed comparably higher-educated individuals. Deviations regarding employment status at the time of the survey are large, which can partly be due to different ways of assessment. In the present sample, 19.5 per cent stated that they “currently work”. In contrast, 84 per cent of the TRC/WFP sample stated that they had “at least four days of formal or informal work in the previous month” (TRC & WFP, 2019: 20). According to the Syrian Barometer 55 per cent are “currently working in an income-generation job” (Erdogan, 2020: 137). The IOM study additionally asked for employment status before departure from Syria, which was also assessed in the present study using the same response options. Here the distribution patterns are very similar. Distributions of English language skills, another indicator of socio-economic status, are also rather similar across samples (no figure).

The final comparison was made for two attitudinal measures that were assessed in similar ways in one of each of the comparison studies (Figure 4). The present study as well as the IOM study asked participants about their intended country of residence. The overall distribution patterns are similar: The vast majority intended to stay in Turkey, followed (with great distance) by Europe and Canada. However, in the IOM study, the proportion of those who intended to stay in Turkey is much higher than in the present study. This might be due to different ways of assessment as well: In the present study, participants were asked whether they would like to stay in Turkey in the future. In the IOM study, participants were asked what their intended country of destination was at the time of the survey. Furthermore, the Syrian Barometer 2019 assessed whether participants perceive problems regarding various life domains in Turkey. The responses can be compared to those of the current study, where participants were asked how satisfied they are with various aspects of their life in Turkey. Five life domains were comparable (work, education, health, accommodation and attitudes of Turkish people). Although the response scales differed, both studies show similar patterns. Participants perceive the health and education system most positively, and the employment conditions most negatively.

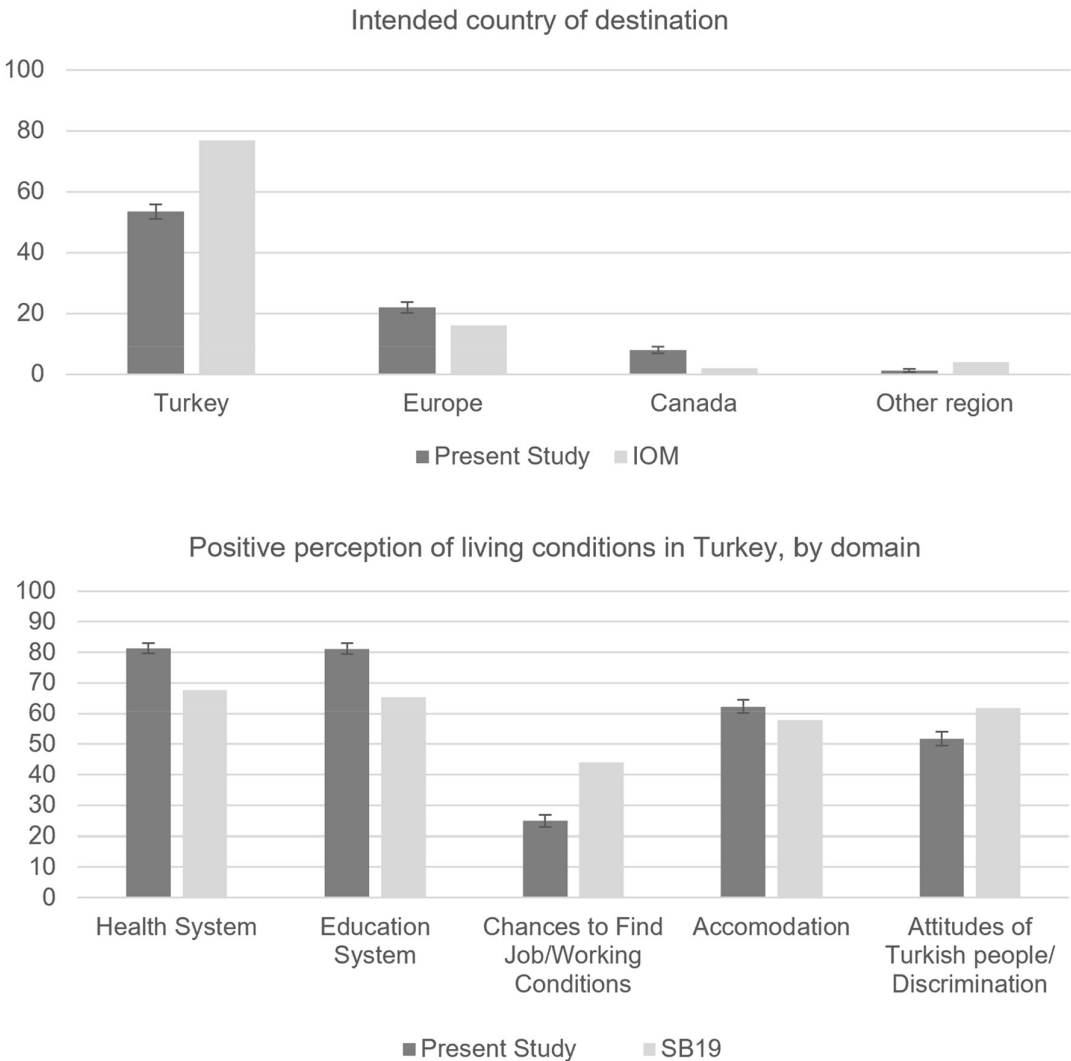


**FIGURE 3** Comparison of education and employment. *Note:* Bar represent proportions in percent. Vertical brackets indicate the 95%-confidence-intervals for the estimates based on the present data. *Sources:* Present study (weighted data), IOM (2018b), TRC/WFP (2019), Erdogan (2020). \*Denominations of categories differ between studies.

**DISCUSSION**

Given the current high numbers of refugees around the globe the need for reliable knowledge about their life circumstances, well-being, perceptions and aspirations is very high. It plays a crucial role for the scientific understanding of forced migration but also for the concrete promotion of their protection, well-being and economic subsistence by state and non-state actors. However, gaining such knowledge via opinion polls among this group is methodologically challenging given its often small proportion of the overall population in a country, its uneven geographical distribution, often social and ethnic diversity and a lack of access to reliable information about the group that would be necessary for drawing representative samples. In such hard-to-survey populations (Tourangeau, 2014) textbook sampling techniques such as random probability sampling are hardly applicable, hence, choosing a “next best solution” is often the only alternative (Müller-Funk et al., 2019: 9–10) to generate a sample that is “representative enough” (Vearey, 2013: 161). Proper sampling methods are subject to ongoing scientific debate.

Based on a survey among Syrian refugees in Turkey, this report contributes to this debate by demonstrating a study design which attempted to generate valid knowledge about the target population despite methodological challenges and a limited amount of time and resources. The study's core approach was a multi-stage selection and sampling strategy that combined probability with non-probability techniques depending on the prevailing circumstances on the different geographical levels. Provinces, cities and city districts were purposively selected in a way that ensured the inclusion of diverse sub-groups and contextual conditions and a high overall coverage of the target population. On the local level random sampling techniques were applied. Furthermore, the lack of administrative



**FIGURE 4** Comparison of migration aspirations and perceptions of living conditions. *Note:* Bar represent proportions in percent. Vertical brackets indicate the 95%-confidence-intervals for the estimates based on the present data. With regard to perceptions of living conditions, participants of the present study were given a response scale that ranged from one (not satisfied at all) to seven (very satisfied). Participants who chose the values five to seven were categorized as having a positive view of the respective life domain. In the SB19 study the scale had five response options from “not experiencing problems at all” to “experiencing a lot of problems”. Participants who chose the two positive responses (not experiencing problems [at all]) were categorized as having a positive view of the domain. *Sources:* Present study (weighted data), IOM (2018b), Erdogan (2020).

data on the regional and local level was compensated through the consultation of local experts. This refers to the local distribution of the target group but also to the specific accommodation situation, which, for example, lead us to adapt our household sampling strategy. Altogether, we gained a sample that covers around 60 per cent of the target population and that is diverse with regard to geographical location and socio-economic living conditions in Turkey.

A comparison with official statistics and with three other surveys conducted under similar conditions at around the same time, but with better resources, revealed that, although point estimations differ to some extent, the internal

distributions of sample characteristics are quite comparable across data sources. We conclude that our sample resembles the target population of adult Syrian refugees in Turkey very well under the given circumstances.

Nevertheless, the fact that the sample also differs from other sources regarding certain characteristics points to the fact that, albeit general tendencies and distributions can be reliably assumed, precise point estimates can be biased to a certain extent. Hence, given the fact that refugees are a vulnerable group and their situation is often a politically sensitive subject, policy recommendations based on the results of surveys that were conducted under such difficult conditions always need to be made with some caution and need to consider potential biases when interpreting findings. This is clear, for example, from the very different gender distributions in the compared studies, which need to be taken into account. Furthermore, when data are collected in a limited number of regions (like in the present case) findings need to be contextualized (Reichel & Morales, 2017). Thus, studies need to be transparent with regard to all details of the sampling and survey methods they applied (Jacobsen & Landau, 2003; Reichel & Morales, 2017). At the same time, actors who use the results of studies for their political or practical work must be aware of the limitations of such studies and critically reflect on them. Whenever possible, findings of multiple studies using different methodological approaches need to be compared and integrated to obtain an overall picture of the target group's situation.

The study corroborates conclusions made from previous studies on hard-to-survey populations that, in the absence of proper conditions for random probability sampling, comprehensive information about the target population and the regional circumstances plays a central role for the development of an adequate study design, be it from administrative data, local experts or extensive preparatory field work. Based on that knowledge, probability sampling techniques can be integrated in an overall sampling strategy as much as possible (Reichel & Morales, 2017; Singh & Clark, 2013; Vearey, 2013). For this, in turn, it is important to involve local researchers, staff and other experts in the research project.

The literature meanwhile provides various options to sample hard-to-survey populations such as refugees. By sharing the present study's experiences, we attempt to contribute to the further improvement of quantitative studies among refugee groups. Which methods are applied (and combined) depends on the studies' research questions and circumstances. After all, the choice of the right strategy is always a balancing act between the necessities of complying with scientific standards and the opportunities and limitations of the field. Overall, our study demonstrates that it is feasible to conduct a "representative enough" (Vearey, 2013: 161) with a limited amount of time and resources.

## ACKNOWLEDGEMENT

Open Access funding enabled and organized by Projekt DEAL.

## PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1111/imig.13130>.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## ORCID

David Schiefer  <https://orcid.org/0000-0001-7710-368X>

## ENDNOTES

- <sup>1</sup> For results of the survey see Düvell et al. (2021). The project also comprised a qualitative interview study, which is, however, not part of the present article.
- <sup>2</sup> An entirely different sampling approach that is not mentioned here makes use of social networks within the target group (snowball sampling, respondent-driven sampling). For reviews on these methods see, among others, Penrod et al. (2003), Marpsat and Razafindratsima (2010) or Reichel and Morales (2017).



- <sup>3</sup> In such cases interviewers asked for the number of groups of people in the apartment who “normally eat together”.
- <sup>4</sup> Similarly high response rates have been reported on other studies on refugees in Turkey (e.g., Müller-Funk et al., 2019: 17).
- <sup>5</sup> 18-year-old individuals (who usually count as adults) were not included because in the DGMM statistic they are grouped together with 15- to 17-year-old persons.

## REFERENCES

- Aljadeeah, S. (2022) Are refugees (really) a hard-to-survey group? Fieldwork experience with Syrian refugees in Germany. *Journal of Refugee Studies*, 35, 1405–1409.
- Baban, F., Ilcan, S. & Rygiel, K. (2017) Syrian refugees in Turkey: pathways to precarity, differential inclusion, and negotiated citizenship rights. *Journal of Ethnic and Migration Studies*, 43(1), 41–57.
- Baio, G., Carlo Blangiardo, G. & Blangiardo, M. (2011) Centre sampling technique in foreign migration surveys: a methodological note. *Journal of Official Statistics*, 27(3), 451–465.
- Bloch, A. (2007) Methodological challenges for national and multi-sited comparative survey research. *Journal of Refugee Studies*, 20(2), 230–247.
- Cibea, A., Kraler, A. & Reichel, D. (2012) *Evaluation of the Immigrant Citizens Survey (ICS)*. Vienna: ICMPD.
- Daily Sabah. (2019) 2,600 unregistered Syrians in Istanbul sent to shelters, governorate says. *Daily Sabah*. Available from: <https://www.dailysabah.com/istanbul/2019/08/01/2600-unregistered-syrians-in-istanbul-sent-to-shelters-governorate-says> [Accessed 11th February 2022].
- DGMM – Directorate General of Migration Management. (2018) *Temporary protection, Republic of Turkey, ministry of interior. Data release as of December 2018*. Available from: <https://en.goc.gov.tr/temporary-protection27> [Accessed 6th December 2018].
- Düvell, F. (2018) From migration transition to integration challenge. *New Perspectives on Turkey*, 58, 179–200.
- Düvell, F. (2019) *Are there really 3.6 million refugees in Turkey or could there be considerably less?* DeZIM Briefing Note, 1. Berlin: German Center for Integration and Migration Research.
- Düvell, F., Schiefer, D., Sağıroğlu, A.Z. & Mann, L. (2021) *How many Syrian refugees in Turkey want to migrate to Europe and can actually do so? Results of a survey among 1,900 Syrians*. DeZIM Research Note, 5. Berlin: German Center for Integration and Migration Research.
- Erdogan, M.M. (2019) *Syrian refugees in Turkey*. Ankara: Konrad Adenauer Stiftung.
- Erdogan, M.M. (2020) *Syrians barometer 2019: a framework for achieving social cohesion with Syrians in Turkey*. Ankara: Orion Kitabevi.
- Erens, B. (2013) Designing high-quality surveys of ethnic minority groups in the United Kingdom. In: Font, J. & Méndez, M. (Eds.) *Surveying ethnic minorities and immigrant populations*. Amsterdam: Amsterdam University Press, pp. 45–65.
- Ersanilli, E., Carling, J. & De Haas, H. (2011) *Methodology for quantitative data collection*. EUMAGINE Project Paper 6A. Oxford/Oslo: IMI/PRIO.
- European Commission. (2019) *Key findings of the 2019 report on Turkey*. Brussels: European Commission.
- Fafchamps, M. & Shilpi, F. (2013) Determinants of the choice of migration destination. *Oxford Bulletin of Economics and Statistics*, 75(3), 388–409.
- Harvey, C., Garwood, R. & El-Masri, R. (2013) *Shifting sands: changing gender roles among refugees in Lebanon*. Joint research report. Oxford: Oxfam and ABAAD-Resource Center for Gender Equality.
- Hubbard, F., Lin, Y., Zahr, D. & Hu, M. (2016) Sampling design. *Guidelines for best practice in cross-cultural surveys*. Ann Arbor, MI: Survey Research Center, Institute for Social Research, University of Michigan. Available from: <https://ccsg.isr.umich.edu/chapters/sample-design/> [Accessed 11th February 2022].
- IOM – International Organisation for Migration. (2018a) *Analysis: baseline assessment in Turkey. April–July 2018*. Geneva: International Organization for Migration.
- IOM – International Organisation for Migration. (2018b) *Analysis: flow monitoring surveys in Turkey. December 2017–February 2018*. Geneva: International Organization for Migration.
- IOM – International Organisation for Migration. (2018c) *Analysis: flow monitoring surveys in Turkey. Field observation report December 2017–February 2018*. Geneva: International Organization for Migration.
- IOM – International Organisation for Migration. (2018d) *MPM Turkey. Migrant presence monitoring. Overview of the situation with migrants*. Annual Report. Geneva: International Organization for Migration.
- Jacobsen, K. & Landau, L.B. (2003) The dual imperative in refugee research: some methodological and ethical considerations in social science research on forced migration. *Disasters*, 27(3), 185–206.
- Kalton, G. (2009) Methods for oversampling rare subpopulations in social surveys. *Survey Methodology*, 35(2), 125–141.
- Kiniklioglu, S. (2020) *Syrian refugees in Turkey: changing attitudes and fortunes*. SWP Comment 2020/C, 05. Berlin: Stiftung Wissenschaft und Politik (SWP).
- Krause, U. (2017) *Researching forced migration: critical reflections on research ethics during fieldwork*. Refugee Studies Centre Working Paper Series, 123. Oxford: University of Oxford.

- Levy, P.S. & Lemeshow, S. (1999) *Sampling of populations: methods and applications*, Third edition. New York: Wiley.
- Lohr, S.L. (2019) *Sampling: design and analysis: design and analysis*. Boca Raton, FL: CRC Press (Taylor & Francis Group).
- Marpasat, M. & Razafindratsima, N. (2010) Survey methods for hard-to-reach populations: introduction to the special issue. *Methodological Innovations Online*, 5(2), 3–16.
- Mckenzie, D.J. & Mistiaen, J. (2007) *Surveying migrant households: a comparison of census-based, snowball, and intercept point surveys*. IZA Discussion Papers, 3173. Bonn: Institute for the Study of Labor (IZA).
- Misago, J.P. & Landau, L.B. (2013) Gutters, gates, and gangs: collaborative sampling in 'post-violence' Johannesburg. *Journal of Refugee Studies*, 26(1), 116–125.
- Muhib, F.B., Lin, L.S., Stueve, A., Miller, R.L., Ford, W.L., Johnson, W.D. et al. (2001) A venue-based method for sampling hard-to-reach populations. *Public Health Reports*, 116(Suppl 1), 216–222.
- Müller-Funk, L. (2021) Research with refugees in fragile political contexts: how ethical reflections impact methodological choices. *Journal of Refugee Studies*, 34(2), 2308–2332.
- Müller-Funk, L. et al. (2019) *Researching urban forced migrants in Turkey and Lebanon: alternative ways to study a vulnerable population in fragile political contexts*. IMIn Working Papers, 151. Oxford: International Migration Institute Network.
- Penrod, J., Preston, D.B., Cain, R.E. & Starks, M.T. (2003) A discussion of chain referral as a method of sampling hard-to-reach populations. *Journal of Transcultural Nursing*, 14(2), 100–107.
- Reichel, D. & Morales, L. (2017) Surveying immigrants without sampling frames – evaluating the success of alternative field methods. *Comparative Migration Studies*, 5(1), 1–22.
- Singh, G. & Clark, B.D. (2013) Creating a frame: a spatial approach to random sampling of immigrant households in inner city Johannesburg. *Journal of Refugee Studies*, 26(1), 126–144.
- Tourangeau, R. (2014) Defining hard-to-survey populations. In: Tourangeau, R., Edwards, B., Johnson, T.P., Wolter, K.M. & Bates, N. (Eds.) *Hard-to-survey populations*. Cambridge: Cambridge University Press, pp. 3–20.
- TRC – Turkish Red Crescent, WFP – World Food Program. (2019) *Refugees in Turkey: livelihoods survey findings 2019*. Ankara: Turkish Red Crescent, World Food Programme.
- Turkstat – Turkish Statistical Institute. (2022) *Central dissemination system*. Available from: <https://biruni.tuik.gov.tr/medas/?kn=95&locale=en> [Accessed 30th March 2022].
- UNHCR. (2014) *Woman alone: the fight for survival by Syria's refugee women*. Geneva: UNHCR.
- UNHCR. (2021) *Global report 2020*. Geneva: UNHCR.
- UNHCR. (2022) *Provincial breakdown of Syrian refugees in Turkey (as of 2 August 2018)*. Available from: <https://reliefweb.int/map/turkey/provincial-breakdown-syrian-refugees-turkey-2-august-2018> [Accessed 29th April 2022].
- United Nations. (2018) *The world's cities in 2018—data booklet (ST/ESA/SER.A/417)*. New York: United Nations.
- Vearey, J. (2013) Sampling in an urban environment: overcoming complexities and capturing differences. *Journal of Refugee Studies*, 26(1), 155–162.
- Vigneswaran, D. & Quirk, J. (2013) Quantitative methodological dilemmas in urban refugee research: a case study of Johannesburg. *Journal of Refugee Studies*, 26(1), 110–116.
- World Bank. (2019) *The mobility of displaced Syrians: an economic and social analysis*. Washington, DC: World Bank.

**How to cite this article:** Schiefer, D., Düvell, F., Sağıroğlu, A.Z. & Mann, L. (2023) Sampling refugee groups: The example of Syrians in Turkey. *International Migration*, 61, 231–247. Available from: <https://doi.org/10.1111/imig.13130>