Briefs on Methodological, Ethical and Epistemological Issues

No.13

Challenges and possible solutions to conducting quantitative surveys with displaced populations

Jana Kuhnt, Charles Martin-Shields, Ruben Wedel

Abstract

Research with and about vulnerable populations, including refugees and migrants, raises several methodological concerns. Due to their experiences, potential respondents might mistrust the motives and independence of researchers and how the information they share might be used. The necessity felt to keep a low profile leads to challenges for researchers to access relevant population groups and difficulties in establishing trust to share information. This article gives examples from across the globe, including Jordan, Turkey and Kenya, on how quantitative researchers try to access these populations and what approaches are used to establish trust, including the involvement of local 'gatekeepers', non-governmental organizations, community leaders, and the employment of refugee enumerators.

Furthermore, conducting quantitative panel surveys with refugee populations that feature a high rate of mobility comes with another set of challenges for researchers. These challenges often evolve around finding the right balance between personal data collection and the protection of personal data. This article illustrates these challenges with examples of panel surveys conducted in Jordan and Turkey.

Another big challenge is the establishment of a representative population sample. Due to a variety of reasons, including that the precise size and distribution of the migrant and refugee population is usually unknown, this has been difficult for researchers. These populations are often settled informally and not included in national surveys and largely not captured in official registries. There are large heterogeneities within these populations and some groups can be described as being rare or hidden systematically. As a result, self-selection is a prominent bias in migrant and refugee samples. This brief sheds light on random sampling strategies that have been used by quantitative researchers to improve the representativeness of the population sample, including innovative mapping strategies using google maps. Also, it discusses the use of social media to reach migrant populations.

Introduction

Survey research with vulnerable populations, specifically refugees and migrants, comes with inherent methodological challenges. These include developing representative samples, accessing hidden populations, and implementing data collection processes. Understanding the contexts that shape many refugees' and migrants' lives can help guide researchers in developing innovative approaches to doing survey research with these communities. The foundation of survey methods is based on population data and willingness to participate. In many cases both are lacking when working with displaced communities. Refugees and migrants are often excluded or missed by national censuses, particularly when settled informally in urban areas. Refugees in camp settings are generally a better accounted for population, but even these settings come with limitations regarding population data. The other side of the equation are the preferences of refugees and migrants themselves. For good reason many of them make efforts to remain 'invisible', including problems with legal status in their host country, surveillance by the country they fled, or the desire to be left alone after a traumatic experience.

We synthesize the current state-of-the art literature on conducting surveys in these contexts in combination with recent field experiences in Jordan, Kenya, Uganda and Turkey to highlight statistically sound, ethical approaches to doing survey work with refugee and migrant communities. We cover this in three sections: developing a population sample, gaining access to the communities, and implementing the survey. By identifying best practices for better data collection, we aim at supporting the development of a sound scientific knowledge base.

Challenges and Potential Solutions

Finding and choosing the target group

The lack of sampling frames for displaced populations is a widely acknowledged problem in forced migration research (e.g., Jacobsen & Landau, 2003; Bloch 2004). In most contexts, particularly in low- and middle-income countries, refugees and displaced people are not accurately represented in census data or other country statistics. National data is often aggregated such that researchers cannot identify migrant populations, i.e. by using the country of birth or ethnicity. For subgroups of the target population sampling frames sometimes exist through registration lists, such as those currently in resettlement programs or registered refugees and asylum seekers. However, even in camps not all refugees are formally registered by UNHCR, NGOs, or relevant host-government agencies. For those settled informally outside of camps or those who arrived a long time ago and have since relocated, the population size and distribution is generally unknown. For example, in 2018 Jordan hosted an estimated number of 1.3 million Syrians of

which around 655,000 were registered refugees. While relatively detailed information on registered refugees is available through UNHCR, information on the specifics of the other Syrians in Jordan remains elusive (MOPIC, 2018). Governments may also be keen to hide or exaggerate the presence of certain groups for political or economic reasons (Jacobsen & Landau, 2003). The high mobility of displaced populations poses an additional challenge to the sampling as their numbers may change quickly and frequently (MacDonald, 2015). These issues complicate the creation of statistically sound sampling frames making it necessary to employ alternative, creative techniques.

Probability Sampling Techniques

In order to establish representative surveys that allow the generalization of findings one has to use probability sampling techniques. These can include random, stratified or cluster samples but the key is that there is a complete and accurate sampling frame and hence, that all members of a population have the same probability of being selected into the survey.

Due to the lack of traditional information on displaced populations, particularly for those in non-camp settings, using a combination of data sources is advisable. Researchers have employed different methods to do so. These mostly include creating lists of names and details for every member of the target group or by obtaining non-identifiable data describing demographics and areas of residence. As a starting point key informants, such as UN agencies, community and aid organizations, clan leaders or service providers can be of great help to identify geographic areas where displaced people reside (Jacobsen & Landau, 2003; Suleiman-Hill & Thompson, 2007). Studies have often used multi-stage cluster sampling where clusters (e.g., neighborhoods or other kinds of enumeration areas) were randomly chosen. In Nairobi, Eppler et al. (2019) worked jointly with local university partners to identify high-density refugee neighborhoods. In these a complete listing of houses inhabiting refugees was conducted with the support of several community leaders. Some studies have used satellite imagery to identify households, including algorithms for automated counting of residential structures (Checchi et al. 2013; Wang et al., 2015). In Uganda, GoogleMaps images were used to identify housing units within refugee camps (Betts et al., 2019). In the selected areas households were sampled randomly. Some studies have assumed that non-registered and registered refugees have comparable residence patterns. They used UNHCR lists to randomly select registered refugees as starting points in the selected geographical areas to enquire about the nearest unregistered household (Doocy et al., 2015).

Further, the lack of baseline data complicates assessing the representativeness of the selected sample. Some studies have used anecdotal information from within communities or relied on consultations with local

professionals to judge whether their sample accurately reflects the target population (Suleiman-Hill & Thompson, 2011). Others have used weighting methods to increase the representativeness of their sample (Enticott et al., 2017).

Non-Probability Sampling Techniques

Due to the paucity of reliable sampling frames and difficulties in identifying the hidden target population, several studies have chosen non-probability sampling techniques that come with the shortcomings of selection bias.

Researchers often use snowballing sampling techniques where the identification of the study population is accomplished through personal endorsement. Here, researchers generally start off by contacting a local organization that knows the respective community and geographical area, such as a refugee, aid or religious organization. These then help in approaching potential research subjects. The major shortcoming of this method is the bias that comes from selecting only certain members of the population that share specific characteristics such as belonging to a common social or ethnic group.

There are several suggestions to reduce the selection bias of this method, such as using multiple entry points into communities to enlarge the contact base (with different demographic characteristics), increasing the sample size, using quotas for key variables, and considering only a limited number of links per contact chain (Suleiman-Hill & Thompson, 2011; Bloch, 2004; Atkinson & Flint, 2001).

Online social media can be another interesting tool for large-scale and cost-effective surveys. They can cover wide geographical areas. Kuhnt et al. (2017) used a social media survey to reach host and refugee youth in Jordan. However, selection bias is large as it only covers technologically literate groups, those able to access the Internet, and is limited to certain age groups. Generally, detailed knowledge about the target group is crucial to increase the generalization of knowledge in the absence of sampling frames.

Applying multiple non-probability sampling methods, such as combining snowball, expert interview and social networking techniques, might be another interesting methodology to establish a more representative sample (Enticott et al., 2017). This often requires additional resources but helps to access more parts of the target population.

Accessing the target group

Once a sample has been identified, researchers have to gain access to the households and physical spaces that displaced people live in and use. In many cases this includes slums and informal settlements, and satellite neighborhoods where people live in large housing blocks. Thus, researchers must develop networks of trust and plan for challenges in actually getting to these areas, which may not have paved roads or are well outside city centers.

Building Trust

The vulnerability of displaced populations may affect their trust towards researchers and make them reluctant to participate in surveys (Bloch, 2007). As they try to remain hidden to stay safe, openness towards strangers, including researchers, is often limited. This is particularly relevant for groups that have ambiguous legal status. Collaboration and continuous active engagement with trusted community members or organizations, clan leaders, or other 'gatekeepers' helps to establish access and trust (Enticott et al., 2017). In Nairobi, community leaders introduced the research team to the neighborhoods and were called in case enumerators were mistrusted by respondents (Eppler et al., 2019). While working with local enumerators helps establishing trust it is important to consider potential bias steaming from political divisions or personal affiliations between enumerator and respondent, particularly in highly sectarian countries, like DRC or Burundi (Jacobsen & Landau, 2003). Researchers should try to employ enumerators who have an understanding for the respective community but do not belong to this group. In Nairobi, Kenyan-Somali enumerators interviewed refugees from Somalia: While they are culturally close, they do not belong to the same refugee community (Eppler et al., 2019).

Logistics

Another dimension stems from logistics as oftentimes displaced populations live in remote (urban and rural) places with bad roads and security problems (Jacobsen & Landau, 2003). Additional budget and time needs to be allocated for this. Particularly in urban settings accessing displaced populations involves entering large apartment blocks, where often building managers control access. In Nairobi, community leaders helped establishing relations to the respective building caretakers and to negotiate access before survey activities started (Eppler et al., 2019).

Surveying the target group

Ethical and human subjects research guidelines only roughly guide how to implement a survey in displaced communities including the types of questions asked and data security protocols to be followed. Therefore, researchers have to develop customised strategies to ensure ethical standards are met in varying contexts while still achieving a high response rate.

Questionnaire design

Unfortunately, there are few standard questionnaires designed for use with displaced populations complicating the comparison across surveys (MacDonald, 2015). Recently, there have been some advances by international organizations proposing harmonized questionnaires.[i] There is also a paucity of databases on this issue from which best practices could be drawn. Though evaluation hubs with a focus on displaced populations like the Syria Evaluation Portal for Coordinated Accountability and Lessons Learning (CALL) exist, they often lack a particular focus on best methodological practices[ii]. As characteristics of the displaced population are often unknown, questionnaires should be using a simple design with easily understandable questions to accommodate also lower educational levels and disabilities. The application of common quality assurance measures such as pre-testing and translation into local language by experts are crucial for the accuracy of the questionnaire.

Confidentiality

Protecting the identity and traceability of the respondent is particularly important in the context of vulnerable populations, such as urban refugees with no or an unclear legal status. In Nairobi, respondents were well aware of the risks associated with participation, and requested proof that a legal aid NGO, had signed off on the data collection process (Eppler et al., 2019). Project implementers are trusted to protect their privacy by participants. Hence, researchers should first consult with implementers to allow them to fulfil their protection obligations. In Turkey, the contact details of project participants were provided by project implementers to researchers only after the participants had given their explicit consent (Roxin et al., 2019). Thus, well designed anonymity procedures not only protect people from the data being used against them, but also from enumerators and researchers accidently divulging identifying details.

Follow-ups

Displaced populations are highly mobile and frequently change their residence. When applying a longitudinal design in research with displaced populations, a high mobility of respondents can reduce the success of contact attempts in the follow-up phase of the research. To avoid a high non-response rate

researcher can apply various techniques, such as using incentives or recording a second telephone number. In Jordan, respondents that were reached in the second phase were asked for contact details for those that could not be reached, given that respondents had often participated in the same project. In Turkey, specific announcements on social media sites were posted, informing respondents about the scheduled survey and thereby increasing the response rate (Roxin et al., 2019). Collecting social media contact information of respondents that are not as frequently changed as mobile phone numbers might be an interesting future path to take.

Conclusion

Global displacement is at a record high and it is increasingly important to overcome inherent methodological challenges when conducting research with displaced populations. Researchers should be open to using creative sampling methodologies while also clearly understanding and stating limitations. This type of transparency is often lacking (Vigneswaran & Quirk, 2013; Jacobsen & Landau, 2003; Sulaiman-Hill & Thompson, 2011) and can mean that recommended policies might not be true for the entire group if representativeness if falsely assumed. When conducting research in the context of displacement researchers should aim for a representative sample and should apply methods to limit selection bias when using non-probability sampling techniques if a representative sample proves unfeasible. Furthermore researchers should invest in thorough exploratory work and extensive pre-testing to ensure smooth data collection, especially in the context of displacement. However, many research projects have to comply with tight timelines and budgets specified by funding agencies and commissioning bodies. It is hence necessary to foster the understanding of the particularities and challenges in the context of displacement research among these institutions.

References

Atkinson, R. and J. Flint (2001): Accessing Hidden and Hard-to-reach Populations: Snowball Research Strategies. *Social Research Update*, 33.

Betts, A., Chaara, I., Omata, N. and O. Sterck (2019): Refugee Economies in Uganda: What difference does the self-reliance model make? https://www.rsc.ox.ac.uk/publications/refugee-economies-in-uganda-what-difference-does-the-self-reliance-model-make

Bloch, A. (2004): Survey research with refugees. Policy Studies, 25 (2).

Bloch, A. (2007): Methodological Challenges National and Multi-sited Comparative Survey Research. Journal of Refugee Studies, 20 (2), pp. 230-247.

Checchi, F., Stewart, B., Palmer, J. and C. Grunder (2013): Validity and feasibility of a satellite imagery-based method for rapid estimation of displaced populations. International Journal of Health Geographics, 12 (4).

Doocy, S., Lyles, E., Roberton, T., Akhu-Zaheya, L., Oweis, A. and G. Burnham (2015): Prevalence and care-seeking for chronic diseases among Syrian refugees in Jordan. BMC Public Health, 15 (1097).

Enticott, J., Shawyer, F., Vasi, S., Buck, K., Cheng, I., Russell, G. Kakuma, R., Minas, H. and G. Meadows (2017): A systematic review of studies with a representative sample of refugees and asylum seekers living in the community for participation in mental health research. BMC Medical Research Methodology, 17 (37).

Eppler, M., Gaetani, S., Köllner, F., Mebrahtu, N., Peters, A., Preiß, C., Martin-Shields, C. and J. Kuhnt (2019): Information and Communication Technology in the Lives of Forcibly Displaced Persons in Kenya. forthcoming

Jacobsen, K. and L. Landau (2003): The Dual Imperative in Refugee Research: Some Methodological and Ethical Considerations in Social Science Research on Forced Migration. Disasters, 27 (3), pp. 185-206.

Kuhnt, J., Rischke, R., David, A. and T. Lechtenfeld (2017): Social cohesion in times of forced displacement – the case of young people in Jordan. CRC Discussion Paper Series. No. 243.

MacDonald, A. (2015): Review of Selected Surveys of Refugee Populations 2000-2014. https://www.efta.int/sites/default/files/documents/statistics/training/Review%20of%20surveys%20of%20refugee%20populations.pdf

Ministry of Planning and International Cooperation (MOPIC), Jordan Response Plan for the Syria Crisis 2018 – 2020, 2018

Roxin et al. (2019): Effectiveness of German Development Cooperation in dealing with conflict-induced forced migration crises, German Institute for Development Evaluation (DEval), Bonn, forthcoming.

Sulaiman-Hill, C. and S. Thompson (2011): Sampling challenges in a study examining refugee resettlement. *BMC International Health and Human Rights*, 11 (2).

UNHCR, Forced Displacement in 2017, Global Trends, 2018. https://www.unhcr.org/dach/wp-content/uploads/sites/27/2018/06/GlobalTrends2017.pdf

Vigneswaran, D. and J. Quirk (2013): Quantitative Methodological Dilemmas in Urban Refugee Research: A Case Study of Johannesburg. Journal of Refugee Studies, 26 (1), pp. 110-116.

Wang, S, So, E. and P. Smith (2015): Detecting tents to estimate the displaced populations for post-disaster relief using high resolution satellite imagery. International Journal of Applied Earth Observation and Geoinformation, 36, pp. 87.93.

[i] Examples are the Vulnerability Assessment Framework Baseline Surveys by UNHCR or the MED-HIMS surveys funded by the EC.

[ii] CALL contains various types of resources. Among them are tools, guidelines and methodologies (e.g. ALNAP "Evaluation of Humanitarian Action Guide" or the ACAPS guidance note "Confidence in Needs Assessment Data: The Use of Confidence Ratings in the Syria Multi-Sectoral Needs Assessment"). However, the vast majority of documents focus on program and project findings rather than on best methodological practices (e.g. questionnaires) that would be particularly useful to researchers and evaluators.