



Establishing a probability sample in a crisis context: the example of Ukrainian refugees in Germany in 2022

Hans Walter Steinhauer  · Jean Philippe Décieux · Manuel Siegert · Andreas Ette · Sabine Zinn

Received: 6 July 2023 / Accepted: 19 February 2024 / Published online: 4 March 2024
© The Author(s) 2024

Abstract Following Russia’s invasion of Ukraine in early 2022, more than one million refugees have arrived in Germany. These Ukrainian refugees differ in many aspects from Germany’s past forced migration experiences and there exists an urgent need for sound data and information for politics, practitioners, and academics. In response, the IAB-BiB/FReDA-BAMF-SOEP study was established to provide high-quality longitudinal data following a register-based probability sample. We detail on an approach for sampling refugees in brief time, making use of two different registers—the German population register and the central register of foreigners—and discuss the quality of the final sample with respect to potential selectivity of participation in the panel. Overall, we demonstrate the benefits and feasibility of establishing register-based samples even in the context of a geopolitical crisis and the necessity of sound data within brief time horizons. We provide guidance that can be followed for similar events in the future.

Keywords Ukrainian Refugees · Probability Sampling · Central Register of Foreigners · Population Registers

✉ Hans Walter Steinhauer · Sabine Zinn

Socio-Economic Panel, German Institute for Economic Research, Mohrenstraße 58, 10117 Berlin, Germany

E-Mail: hwsteinhauer@diw.de

Jean Philippe Décieux · Andreas Ette

Migration and Mobility, Federal Institute for Population Research, Friedrich-Ebert-Allee 4, 65185 Wiesbaden, Germany

Manuel Siegert

Research Centre, Federal Office for Migration and Refugees, Neumeyerstraße 22–26, 90411 Nürnberg, Germany

Sabine Zinn

Humboldt University Berlin, Unter den Linden 6, 10099 Berlin, Germany

1 Introduction

On 24 February 2022, Russia started its military invasion of Ukraine; the escalation of the armed conflict continues to affect large parts of the country. In response to the loss of security and protection, millions of Ukrainians fled the country into neighboring countries and member states of the European Union (EU). Refugees mostly fled to Poland as neighboring country. Germany is hosting the second largest community of Ukrainian refugees within the EU. Before the Russian invasion in February 2022, immigrants from Ukraine constituted a comparatively small group in Germany. At the end of 2021, about 155,000 Ukrainian citizens were living in Germany; this reflects a long-term, decade-long increase—of about 2.6% per year—in the number of Ukrainians. Their migration volume was comparatively small, with an average yearly immigration of around 13,000 Ukrainian citizens arriving in Germany between 2012 and 2021 (see Fig. 1).

This pattern changed fundamentally in 2022. February 2022 saw 14,000 Ukrainian citizens fleeing the war (most of them around the beginning of the war end of February). In March 2022, about 417,000 Ukrainian citizens arrived. Although numbers of arriving Ukrainian refugees quickly decreased thereafter, almost 64,000 refugees arrived in Germany during June 2022. In less than five months, the stock of Ukrainian citizens registered increased almost seven times, reaching 1.02 million Ukrainian citizens registered in Germany at the end of June 2022.

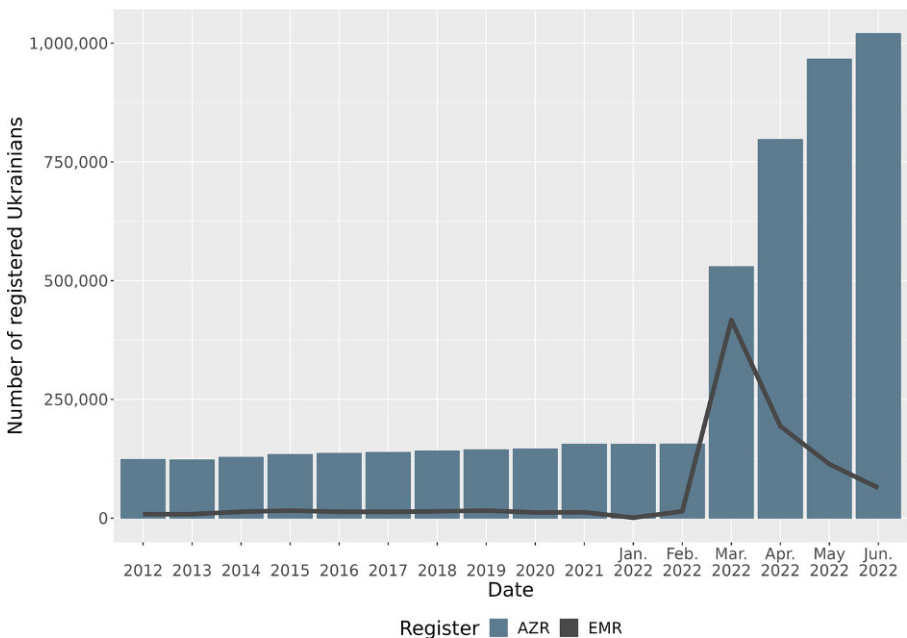


Fig. 1 Development of immigrant flows (EMR) of Ukrainian citizens to Germany and stocks (AZR) of Ukrainian citizens in Germany, 2012–2022 (2012–2021: annual figures, 2022: monthly figures). (Source: Special analysis from the German Central Register of Foreigners (AZR) (reporting date 30 November 2022) and German Population Register (EMR))

The reception of Ukrainian refugees and the provision of options for integration pose major challenges for policymakers, administration, and society. While EU member states, including not least Germany, have learned a lot from previous large-scale arrivals of refugees, the influx of Ukrainians differs from the past in at least five key characteristics (Brücker et al. 2023): First, the demographic composition differs with women, children, and elderly people dominating recent forced migration flows from Ukraine, because adult men must remain in Ukraine for military service. Second, refugees from Ukraine are granted a special legal status by European law (Article 5(1) of European Council Directive 2001/55/EC adopted in July 2001), which enables them to obtain a residence title (residence permit according to § 24 of the Residence Act (temporary protection)) in Germany without an asylum procedure (Federal Office for Migration and Refugees 2022). Third, the acceptance of Ukrainian refugees in the host society seems to be higher than in previous large-scale arrivals of refugees (Dražanová and Geddes 2022). Fourth, refugees arriving in the past were usually allocated (Steinhauer et al. 2019). Ukrainian refugees, on the other hand, had the opportunity to choose their own place of residence, provided they were able to find their own accommodation—for example with relatives, friends, or acquaintances. Only refugees who were unable to provide themselves with accommodation were distributed geographically (Adam et al. 2021). Fifth, the short distance between Ukraine and Germany—compared to origin countries in the Middle East and Central Asia, for example—reduces the costs and threats of travel routes. The greater ease of traveling is also supported by waiving train fares between Ukraine and Germany for refugees. Additionally, geographical proximity makes circular migration between origin and host countries more likely to happen frequently.

The volume and speed of forced migration from Ukraine, together with the substantial differences existing between recent Ukrainian refugees compared to the arrival of refugees in the past, require comprehensive knowledge and sound (longitudinal) data to understand its individual and societal consequences. Whereas several initiatives quickly responded to those data requirements with the implementation of ad-hoc surveys of Ukrainian refugees based on readily available non-probability samples, the IAB-BiB/FReDA-BAMF-SOEP survey was launched with the aim of generating high-quality probability-based longitudinal data on Ukrainian refugees recently arriving in Germany. The project is a joint work between the German Institute for Employment Research (*Institut für Arbeitsmarkt und Berufsforschung*, IAB), the German Federal Institute for Population Research (*Bundesinstitut für Bevölkerungsforschung*, BiB), the Research Center of the German Federal Office for Migration and Refugees (*Forschungszentrum des Bundesamtes für Migration und Flüchtlinge*, BAMF-FZ), and the German Socio-Economic Panel (SOEP). The success of this study depends on the quick and effective creation of a probability sample of Ukrainian refugees in Germany. Only with a probability-based approach it is possible to generalize the results of the study to Ukrainian refugees in Germany. This paper provides details on how we create a random sample using two different administrative registers: the German population register (*Einwohnermelderegister*, EMR) and the German Central Register of Foreigners (*Ausländerzentralregister*, AZR). Using both registers in combination allowed for benefitting from their ad-

vantages while balancing their disadvantages. Specifically, centrally available basic information from the AZR about newly registered Ukrainian citizens from 24 February 2022, onwards provided the basis for sampling municipalities in Germany hosting Ukrainian refugees in a first phase. Here, the AZR provides timely information on the overall number of Ukrainian refugees registered with reception facilities, the police, or foreigners' authorities at the municipal level. In contrast, these numbers are not centrally available from the EMRs, because EMRs are maintained decentral at the level of municipalities. The advantage of the EMR, however, is that it contains individual address data for people registered within the municipality, which was not yet the case in the AZR at that time. For this reason, we draw a sample of municipalities in a first phase using information provided by the AZR on the number of Ukrainian refugees. Within the sampled municipalities, we ask the EMRs to list all Ukrainian nationals aged 18 to 70 who registered after 24 February 2022 together with their addresses. This procedure builds on the example of the "Refugees in the German Educational System (ReGES)" study (Steinhauer et al. 2019), but extends it to a sample covering all German federal states and responding to immediate migration flows.

The paper presents the sampling approach then discusses its strengths and weaknesses with particular emphasis on potential bias due to consent to panel participation. Section 2 provides an overview of general sampling techniques for refugee populations, while Sect. 3 offers an overview of recent surveys of Ukrainian refugees. Details of our approach on sampling refugees are discussed in the following two sections: Sect. 4 provides information on the registration and allocation of Ukrainian refugees in Germany and the sampling of municipalities from the AZR. In Sect. 5, the design of the IAB-BiB/FReDA-BAMF-SOEP is introduced with respect to sampling Ukrainian refugees. Section 7 details the fieldwork and the response rates of the study before Sect. 7 concludes. The paper shows, first, that Germany has by now implemented an efficient system of administrative registration of refugees that can be successfully used for sampling in the context of geopolitical crises and resulting large-scale refugee or migration flows. Second, it shows that the combination of both registers—the EMR and the AZR—allows for establishing high-quality probability samples despite complicated (mostly data protection related) regulations for accessing those registers and even in contexts when information is urgently needed.

2 Sampling techniques for migrant and refugee populations

Refugees are forced migrants. This distinguishes them in essential aspects from other migrants, such as labor migrants or migrants due to family reunification. Voluntary migration usually happens after a long decision-making process (Kley 2017). People leaving their home country to live elsewhere represent only a very small proportion of the home population (worldwide, about 3% in 2015; see Willekens et al. 2016). Refugees, however, flee in a hurry and do not follow a (purely) rational plan regarding their escape route, place of refuge, or about their further life course (Hunkler et al. 2022). Refugees often flee to neighbouring or nearby countries, where they constitute a major proportion of the refugee population. The reason is that war, ex-

tensive persecution, or displacement makes more people fear for life and limb, thus driving them to act. Focusing on recent examples, 30% of Syrian citizens and almost 20% of Venezuelan citizens have left their home country to flee violent unrest and destruction (United Nations High Commissioner for Refugees 2022a).

In the hosting country, nevertheless, refugees are usually still a small group compared to the native population. Moreover, they are highly likely to change their residence frequently (Bloch 1999); although local legal residence requirements often play a key role in refugees' freedom of movement (see El-Kayed and Hamann (2018) for the German case). Hence, according to the definition of Tourangeau (2014), they are a hard-to-reach population (see also Massey 2014; Wenzel et al. 2022). Tourangeau (2014) defines a population as being hard to reach if it is either hard to *sample*, hard to *identify*, hard to find or *contact*, hard to *persuade*, hard to *interview*, or a combination of these aspects. Specifically, a group is hard to sample when there is either no sampling frame available for the group or the group is small with respect to their fraction in the population. Another reason for a population to be hard to sample is mobility. Highly mobile people cannot be easily located at a certain place of residence. A population is hard to identify when it is stigmatized, sensitive, or if screening questions miss members of the population resulting in under-coverage. Groups are hard to find or contact when their members are mobile, not willing to be identified as part of a group, or simply protected by gatekeepers. Given contact is established, some individuals are hard to persuade to participate in the survey. This is often related to busyness, alternatives to spend their time, the survey topic, or the authority issuing the survey. Finally, some individuals are hard to interview because of language problems or because of their cognitive or physical abilities.

In the literature, several methods are proposed to draw samples of hard-to-reach populations (see Andreß and Careja (2018) for migrant populations). The most prominent strategies are location sampling, snowball sampling, respondent driven sampling, convenience sampling, and (screened) register samples. The first two approaches provide non-probability samples. The latter lead to probability samples. In location sampling respondents are recruited in places where they spend a notable amount of time. In snowball sampling a target person gives access to the survey questionnaire to other target persons he or she can contact. Both approaches are commonly used to recruit migrant groups, including refugees (e.g., Agadjanian and Zotova 2012; McKenzie and Mistiaen 2009). However, they yield non-probability samples whose survey statistics require a model-based framework to be extrapolated to the population level. Compared to the design-based approach the methodology for estimation becomes more complex. Because information on the target population is usually not available, it is also not possible to compensate for selection biases (see Groves 2006 and Kalton 2014), e.g., regarding groups that are commonly not well covered by migrant surveys such as uneducated migrants (Amior 2020). Furthermore, research shows that the quality of non-probability samples is significantly worse and less robust compared to random-based samples (see Cornesse et al. 2020; MacInnis et al. 2018). Respondent driven sampling (RDS) is applied in various migrant studies (e.g., Lattof 2018). The basic idea of RDS is to sample people from a hard-to-reach population and make use of their social networks, thus relying on the sampled person to be well connected to other people of the population. Compared

to snowball sampling, RDS does allow for drawing a random sample, when certain assumptions are met, e.g., the referral chains become long enough. However, the conditions for generating a random sample of migrants are difficult to achieve with this technique (especially passing on survey questions through long chains of respondents); see also (Tyldum 2021). Switching RDS to a web mode does not really overcome its obstacles. On the contrary, it makes it vulnerable to misuse and poses a threat to data quality (Sosenko and Bramley 2022). In convenience sampling, interviewers select participants for a survey, for example based on their proximity (e.g., being at a reception center) or certain characteristics (e.g., being user of Facebook). But also, respondents can select themselves into the survey, e.g., an online survey promoted on social media.

To overcome these drawbacks, random samples based on registers are, not only for migrants or refugees, seen as a superior method of generating a sample. This requires, first, register data and, secondly, access to register data. Whether such data exists and is also accessible for scientific purposes varies across countries. In general, the data situation is better in Scandinavian countries than in other European countries or around the world (see Weber and Saarela 2019; Bell et al. 2015). In Germany, the two most comprehensive registers for sampling migrants and refugees are the German population register and the central register of foreigners. In general, the population register constitutes the most comprehensive sampling frame with the legal obligation to register in the local registration office within two weeks after changing address in Germany. However, a major obstacle for sampling individuals at the federal level is that the German population register is not organized centrally. The register, maintained at the level of the municipality, contains addresses and basic personal demographic information (e.g., gender, date of birth, nationality, date of migration, see § 3 Bundesmeldegesetz for the full list) of almost all persons who are officially residing in Germany; thus, also of all officially registered refugees and migrants. The register can be used and accessed for scientific purposes based on § 34 and § 46 Bundesmeldegesetz. Here, the information accessible is limited and each information must be substantiated. Each registration office can decide whether to provide the desired information. Recent examples of using the population register for establishing migrant samples include the project “Socio-cultural integration processes among New Immigrants in Europe” (Diehl et al. 2016), the “German Emigration and Remigration Panel Study” (Ette et al. 2021), and the panel of the German Centre for Integration and Migration Research “DeZIM-Panel” (Dollmann et al. 2022). However, because of the decentralized structure of Germany’s population register, its use must always rely on two-stage sampling. In this sampling technique, a random sample of municipalities (serving as primary sampling units, PSUs) with enough migrants is selected at the first stage and individual addresses (representing the secondary sampling units, SSUs) are drawn at the second stage.

The second register used for sampling migrants and refugees in Germany is the central register of foreigners, which documents all persons who are not German nationals and stay in Germany for more than 90 days (Babka von Gostomski and Pupeter 2008). It receives its information mostly from (local) immigration offices in Germany (“Ausländerbehörden”), whose area of responsibility coincides with that of the German municipalities (“Gemeinde”) and districts (“Kreise”). Thus, two-stage

samples can be selected from this register. Recent examples for using the AZR include the study “Forced Migration and Transnational Family Arrangements” (Sauer et al. 2022) as well as the establishment and refreshing of the IAB-BAMF-SOEP study (Brücker et al. 2016; Kühne et al. 2019 and Steinhauer et al. 2022). Studies using the AZR as a sampling frame also follow a two-stage sampling technique, usually sampling local immigration offices (PSUs) before sampling the individuals registered at those offices (SSUs). When drawing our sample, however, addresses were only available in the AZR for refugees who were undergoing asylum proceedings. Because Ukrainian refugees do not have to undergo an asylum procedure, no addresses were available for them. Moreover, obtaining registered addresses from the immigration authorities would have taken a comparatively long time.

3 Review of existing surveys of Ukrainian refugees

Currently, most of the few existing surveys about Ukrainian refugees use non-probability sampling to create their sample—with all the drawbacks mentioned earlier. A brief overview of studies on Ukrainian refugees is presented in Table 1 providing details on the host country, the field period the survey was conducted, the number of respondents as well as the sampling design applied by the study. With respect to comparative samples of Ukrainian refugees across different hosting countries, the United Nations High Commissioner for Refugees (UNHCR) and partners collected data in Czech Republic, Hungary, Moldova, Poland, Romania, and Slovakia from 4871 Ukrainian refugees using a location sampling approach between 16 May and 15 June 2022. Here, interviews were mostly conducted at locations such as border areas and transit zones or information and assistance points (United Nations High Commissioner for Refugees 2022b). Similarly, the European Union Agency

Table 1 Brief overview on existing studies on Ukrainian Refugees

Study/Conductors	Country	Field period	Respondents	Sampling design
UNHCR & partners	Czech Republic, Hungary, Moldova, Poland, Romania, and Slovakia	16 May to 15 June 2022	4871	Location sampling
EUAA & OECD	European Union	11 April to 7 June 2022	2369	Convenience sampling
BMI	Germany	March 2022	Approx. 2000	Location sampling, convenience sampling
GESIS	Germany, Poland	April to May 2022	Approx. 1300	Convenience sampling
UkrAiA	Austria	March to June 2022	Approx. 1000	Convenience sampling
UkrAiA	Kraków, Poland	March to June 2022	500	Location sampling
Ukrainian Refugees in Poland Survey 2022	Poland	June to August 2022	Approx. 1800	Stratified two-stage random sampling

for Asylum (EUAA) together with the Organization for Economic Co-operation and Development (OECD) collected information on 2369 Ukrainian refugees via convenience sampling using an online mode of data collection during the time between 11 April and 7 June 2022 (European Union Agency for Asylum 2022).

In Germany, the Federal Ministry of the Interior and Community (BMI) launched an early survey in March 2022 at the registration offices of three central hubs in Berlin, Hamburg, and Munich, where they interviewed refugees. Additionally, the survey was advertised on the homepages of the BMI, BAMF, and by the mobile phone app *Germany4Ukraine.de* (Federal Ministry of the Interior and Community 2022). This multisource convenience sampling approach resulted in almost 2000 interviews. Additionally, a web-based survey on Ukrainians staying in Germany or Poland was run by the Leibniz Institute for the Social Sciences (GESIS) in April and May 2022 (Pöttschke et al. 2022). The study recruited around 1300 refugees through adverts on social media. Both studies provide an initial picture of the fates, living situations, attitudes, problems, and needs of the Ukrainian refugees in Germany. However, as both studies do not use random samples, neither can make general statements about the situation of Ukrainian refugees in Germany. Further, both studies consist of only one wave. Such designs cannot provide any insights concerning the situation of Ukrainian refugees and how it is changing over time.

With respect to other major host countries of Ukrainian refugees, Austria launched a rapid response survey (*Ukrainian arrivals in Austria, UkrAiA*) using convenience sampling to quickly learn about the socio-demographics of Ukrainian refugees as well as their educational resources and intentions to stay in Austria or return. The survey was conducted between March and June 2022 using both pen and paper interviews (PAPI) and computer assisted web interviews (CAWI). The survey conducted more than 1000 interviews with adult respondents, also collecting information on their partners and children (Kohlenberger et al. 2022). Parallel to the study conducted in Austria, 500 Ukrainian refugees were surveyed in Kraków, Poland, also using location sampling at different registration spots (Pędziwiatr et al. 2022). Moreover, Poland together with the World Health Organization (WHO) implemented the *Ukrainian Refugees in Poland Survey 2022*. They use a stratified two-stage random sampling design to sample Ukrainian refugees. At the first stage, locations (i.e., PSUs) were stratified by border crossing regions and PSUs were randomly sampled. Persons (i.e., SSUs) aged 18 years and older were sampled using systematic sampling within the PSU, if they had already stayed in Poland for at least two weeks. Roughly 1800 sampled persons also provided information third persons they were traveling with, thus yielding information on about 5000 Ukrainian refugees (Beqiri and Cierpiak-Wolan 2022).

4 Registration procedure of Ukrainian refugees in Germany

Knowing that comprehensive registers are the best choice for creating a random sample of Ukrainian refugees is one side of the coin, but finding and accessing such a register within a reasonable amount of time and creating an appropriate sample in a timely manner is the other. A fundamental issue here is how official authorities

manage and document refugee registration. Due to the organizational confusion of Germany's registers following the large-scale arrival of refugees in 2015 (Bogumil et al. 2018), the immigration administration implemented reforms. These reforms achieved timely registration within the AZR and rapid provision of data.

Foreigners require a so-called residence permit for a legal stay in Germany that goes beyond tourist purposes, i.e., lasts longer than 90 days. These residence permits are issued by the immigration authorities, which is accompanied by registration in the Central Register of Foreigners. Refugees in particular either must apply for asylum upon arrival with the border authorities when entering Germany or, after crossing the border, inside the country. As part of their application, they are registered in the AZR by the relevant authorities. During the registration process, only the addresses of persons who were in the asylum procedure were recorded; the addresses of all other foreigners listed in the AZR had to be obtained from the responsible immigration authorities, which is often comparatively time-consuming.

Due to the special situation in Ukraine, war refugees from there were temporarily exempted from the requirement of a residence title for a period of up to 90 days (Federal Office for Migration and Refugees 2022). At the end of these 90 days, they also had to apply for a residence permit at a foreigner's authority and then they were registered in the AZR. However, Ukrainian refugees are officially registered in the AZR immediately when they apply for state support, e.g., if they need accommodation or access to the social, health or education system. This could occur directly upon arrival.¹ If refugees claimed state support on accommodation, they were also distributed spatially in Germany during this registration. After arriving at their assigned location, they had to report to the local residents' registration office as soon as possible. At this point, they were registered in the EMR.

However, it should be noted that according to the Federal Registration Act (§ 17 Bundesmeldegesetz), every person who settles in a German municipality must register with the local resident's registration office within two weeks of moving in, which in turn results in registration in the EMR. Since a substantial proportion of Ukrainian refugees were able to provide themselves with housing and therefore did not have to seek state support immediately, it can be assumed that a correspondingly considerable proportion of refugees were first registered in the EMR and only then in the AZR. A look at the number of persons entered in the individual registers accordingly indicates that entry in the EMR actually often took place somewhat earlier than entry in the AZR.² Fig. 2 shows the number of refugees (registered as of 31 May 2022, in 100 selected municipalities in Germany with a substantial number of Ukrainian refugees) compared by EMR and AZR data. A short delay between both registration events is obvious.

Each dot in Fig. 2 represents one of 100 municipalities, and the angle bisector (line) indicates equality of the municipality-specific registration numbers in the EMR

¹ Special analysis of the AZR during early 2022 showed that different authorities were used to register. Here, Police registration and reception facilities registered the minority of refugees. The foreigners' registration office and their branches cover more than 86% of the Ukrainian refugee population.

² In this context, it should be noted that the data from the AZR and the EMR is not synchronized by default.

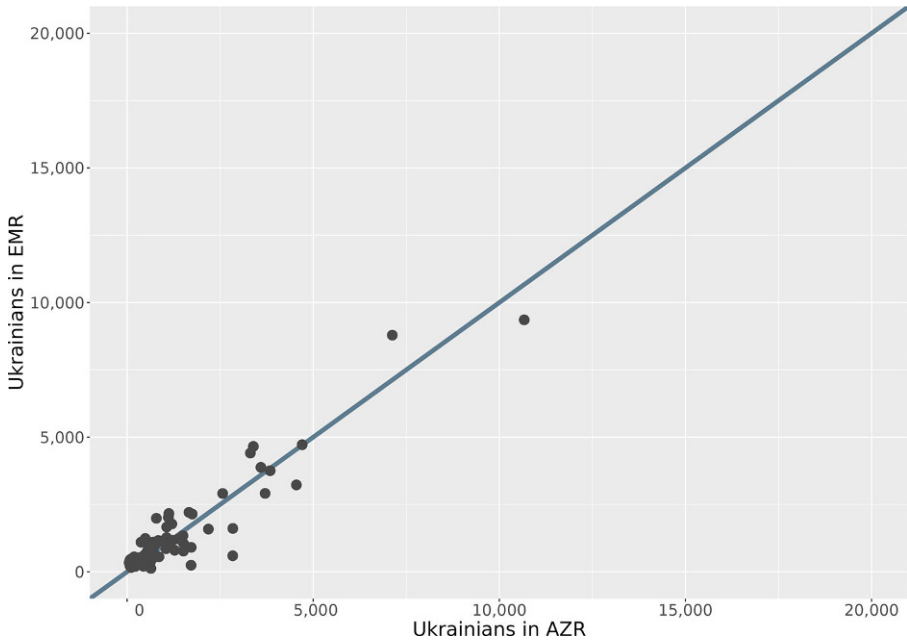


Fig. 2 Number of Ukrainians aged 18 to 70 years registered in the AZR and EMR. (Source: Special analysis from the AZR and EMR, reporting date 31 May 2022. Authors' own calculation)

and the AZR. Points above the line represent municipalities where more inflows of Ukrainian nationals have been registered within the EMR compared to the AZR. For municipalities below the line, the numbers of registrations in the EMR were below the comparable number of registrations within the AZR. Focusing on the differences between AZR and the EMR, approximately 27% of municipalities had more Ukrainians registered in the AZR, whereas 73% of all municipalities had more Ukrainians registered in EMR. There were larger discrepancies, particularly in municipalities with initial reception facilities, like Berlin or Hamburg. This is mainly because people registered in the initial reception facilities were redistributed geographically, which is why they were only registered in the EMR at their ultimate destination. However, in half of the municipalities the number of persons registered with the EMR and AZR differ by 184 people or less. Moreover, the number of refugees reported by EMR and AZR in these municipalities showed a remarkably high correlation of 0.91, indicating a high degree of congruence. This level of congruence is also an indicator for decent quality of both registers.

5 The sample of the IAB-BiB/FReDA-BAMF-SOEP-study

The target population for the IAB-BiB/FReDA-BAMF-SOEP study are refugees aged 18 to 70 with a Ukrainian nationality who immigrated to Germany after 24 February 2022. To draw random samples from this population, there are two

distinct registers are available in Germany: the EMR and the AZR. We made use of both at separate phases of a two-phase sampling (Särndal et al. 2003, Chap. 9). In a first phase we select municipalities throughout Germany based on the number of Ukrainian refugees aged 18 to 70 who were registered in the AZR after 24 February 2022. In the second phase we sampled individuals from the EMR (of the sampled municipalities) because it allows for accessing their address information.

More precisely, we selected 100 municipalities in Germany in the first phase using systematic probability proportional to size sampling (Tillé 2006, Chap. 7). The measure of size used is the number of Ukrainian refugees registered at the foreigners' registration offices in the AZR. For municipalities under the authority of a single foreigners' registration office, we used the number reported by the AZR. In cases where an office covered multiple municipalities, we assigned the municipalities the average number of Ukrainians reported by the office. In doing so, the sampling scheme favors urban municipalities. For sampled municipalities, we contacted the local residents' registration office and asked them to provide a list of all persons aged 18 to 70 years with a Ukrainian nationality who registered there after 24 February 2022. Thus, we assumed all persons listed by the EMRs under these conditions to be Ukrainian refugees. To be able to contact the municipalities and collect the lists of addresses from them within eight weeks, the maximum number of municipalities across Germany we could process was 100. Harmonizing the collected lists yielded 135,575 addresses of Ukrainian refugees meeting the study criteria, slightly surpassing the AZR count of 120,279 as of 31 May 2022.

These addresses underwent validation by the survey research institute infas (Institut für angewandte Sozialforschung), resulting in 132,120 valid addresses available for sampling. Rigorous checks based on family and first name, date of birth, and sex excluded 253 duplicate entries, 2081 addresses with age-related discrepancies, and 1121 addresses that could not be contacted, pre-dated the study period, or people lacked Ukrainian nationality.

From previous SOEP surveys of refugees, we know that refugees have a higher initial response rate compared to general population surveys (see Kühne et al. 2019; Steinhauer et al. 2022 and Jacobsen and Siegert 2023). In contrast to their initial response behavior, we have little information on their onward mobility. To provide a net sample consisting of at least 8000 individuals, we need a gross sample size that is sufficiently higher to compensate for non-participation, invalid addresses, or refugees already having moved on. Therefore, building upon experiences of previous studies, we decided to draw a gross sample of size $n=48,000$.

In the second phase we draw the sample from a list of $N=132,120$ individuals. Based on data provided by the residents' registration offices, it is not possible to identify household compositions at a specific address. Because it is not desirable to survey multiple people of the same household, resulting in duplicate information about the household composition as well as about third persons, this had to be avoided. Classically, households are identified or at least approximated via persons living at the same address having the same family name. In the case of Ukrainian refugees, however, this was not a feasible option because the family names of men and women often differ, e.g., Ukrainian President Volodymyr Zelenskyy and his wife Olena Zelenska. To minimize the risk of sampling multiple people from the

Table 2 Artificial example of data delivered by residents' registration offices

Federal State	City	Postal code	Street	Nr.	Family Name	First Name	Individual	Household
Saarland	Saarbrücken	66111	Blumenstr.	129	Basysta	Maria	1	1
Saarland	Saarbrücken	66111	Blumenstr.	129	Basysta	Anastasia	2	1
Saarland	Saarbrücken	66111	Blumenstr.	129	Basystyi	Andriy	3	1
Saarland	Saarbrücken	66111	Blumenstr.	129	Melnyk	Dmytro	4	2
Saarland	Saarbrücken	66111	Blumenstr.	129	Melnyk	Danylo	5	2
Saarland	Saarbrücken	66111	Blumenstr.	131	Boyko	Yaroslav	6	3
Saarland	Saarbrücken	66113	Dresdener Str.	237	Selenska	Sophia	7	4
Saarland	Saarbrücken	66113	Dresdener Str.	238	Olyinyk	Nazar	8	5
Saarland	Saarbrücken	66113	Dresdener Str.	241	Ivanova	Alina	9	6
Saarland	Saarbrücken	66113	Dresdener Str.	241	Ivanova	Daria	10	6

same household, we sort the harmonized list by address and family name and use systematic sampling (Särndal et al. 2003, p. 73 ff.) to select a sample from this list. Table 2 provides artificial data illustrating the rationale behind this chosen design.

The presented table exhibits five distinct addresses (indicated by the first five columns from Federal State to Nr.) corresponding to ten individuals identified by their first and family names in the respective columns. An identification number for each person is provided in the penultimate column, while the last column indicates a household identifier (an information not available to us). In the absence of household identifiers, a simple random sample of four individuals from this list may inadvertently include multiple members of the same household. To mitigate this risk, a systematic sampling approach was employed, wherein individuals were selected in a systematic manner from the list sorted by address and family name. This method reduces the likelihood of surveying multiple individuals from the same household.

The sample, drawn according to the outlined procedure, resulted in a gross sample of 37,904 women and 10,086 men. For 10 individuals in the sample their sex is unknown. Additionally, information on the year of birth is unavailable for 1604 people. When comparing our sample to the population of Ukrainian refugees aged 18–70 registered in Germany's Central Register of Foreigners from 24 February to 31 May 2022, and considering individuals with available birth year information, we observe the following disparities: Our sample has a mean age of 41 years, compared to 44 years in the population. The median age in our sample is 39 years, in contrast to 44 years in the population. The standard deviation of age in our sample is 13.5, while it is 15.4 in the population. Figure 3 displays the joint distribution by age and sex. The left panel illustrates the proportions of Ukrainians aged 18 to 70 registered in the AZR between February 24, 2022, and May 31, 2022, categorized by sex and age. The right panel depicts the distribution within our gross sample. Notably, except for 18-year-olds, the distributions in the sample and the population closely align.

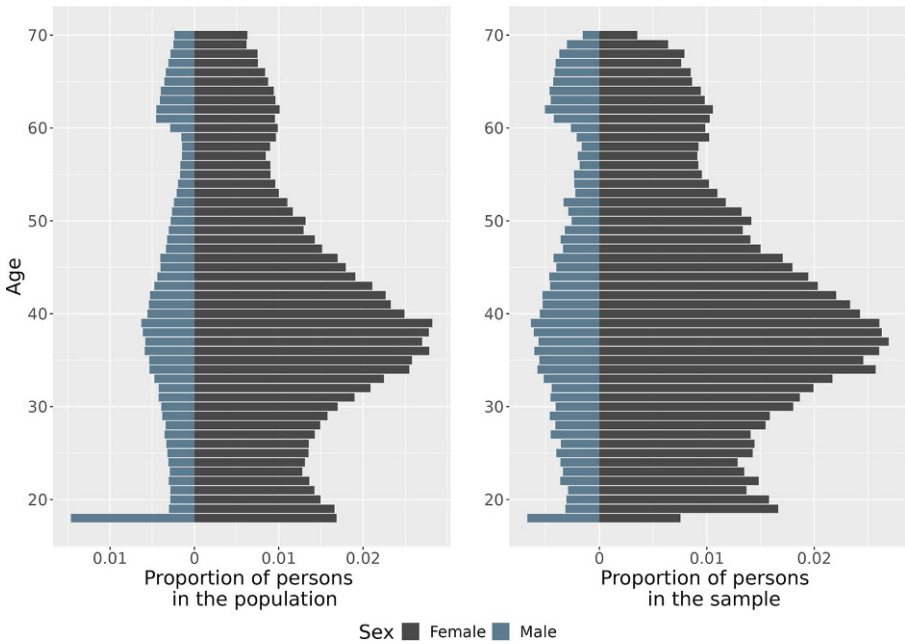


Fig. 3 Proportions of Ukrainian refugees aged 18–70 registered in the Germany’s Central Register of Foreigners (left panel) from 24 February 2022 to 31 May 2022, and the gross sample used for the IAB-BiB/FReDA-BAMF-SOEP-Survey (right panel) by age and sex. (Source: Special analysis from the AZR (reporting date 30 November 2022) and EMR; Authors’ own calculation)

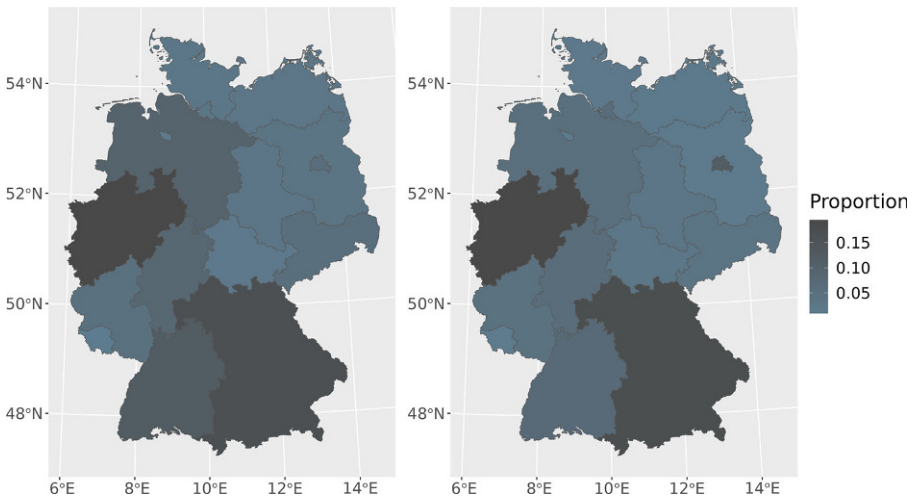


Fig. 4 Comparison of all Ukrainian refugees aged 18–70 and registered in Germany’s Central Register of Foreigners between 24 February 2022 and 31 May 2022, with the gross sample used for the IAB-BiB/FReDA-BAMF-SOEP-Survey by federal state of current residence, in percent. (Source: Special analysis from the AZR (reporting date 30 November 2022) and EMR as well as © GeoBasis-DE/BKG 2022; Authors’ own calculation)

Figure 4 contrasting the regional coverage of the population (left panel) with that of the sample (right panel), reveals a slight over-representation of refugees in Berlin and a marginal under-representation in Baden-Württemberg. Despite these cases, the distributions across federal states for both the population and the sample exhibit congruence.

6 Fieldwork and response rates

The fieldwork for the study started on 24 August by sending out postal invitations to participate in the survey online. A reminder, accompanied by a supplementary PAPI questionnaire, was dispatched on 9 September. The questionnaire included questions on various aspects, such as events related to fleeing (Fluchtgeschichte), arrival in Germany, intentions to stay, housing situation, guidance and support needs, family dynamics (including partners and children), education and qualifications, income situations in both Ukraine and Germany, social networks, health, origins, language proficiency, life satisfaction, and concerns. Out of the 48,000 refugees targeted, 10,395 actively participated in the survey and provided consent to engage in the panel (code 1.1). The remaining disposition codes can be categorized as follows: 10,076 individuals did not receive invitations at the EMR-provided addresses (code 3.1), 24,992 received invitations but opted not to participate (code 2.2), and 2008 either terminated the survey prematurely or did not consent to participate in the panel (code 2.12). As per the standards defined by AAPOR (The American Association for Public Opinion Research 2016), these numbers translate to a response rate of $RR1=0.278$, a refusal rate of $REF1=0.045$, and a cooperation rate of $COOP1=0.832$.

The challenge of being unable to establish contact with 10,076 refugees poses potential complexities in interpretation. It remains indistinct whether these individuals relocated or were simply not listed at the provided postbox addresses. This ambiguity arises from the possibility that refugees have moved to a new address or abroad without filing a forwarding order. Notably, the latter scenario does not pose concerns about bias since such refugees fall outside the scope of our desired population. In contrast, refugees relocating within Germany present a more intricate scenario. While these individuals might introduce concerns regarding possible bias, the sample likely includes refugees who have also moved to new locations. Another plausible explanation is that refugees, leveraging networks in Germany, found accommodation through friends, family, or acquaintances. Although the postal service provides the “c/o” addition for addresses, used by EMRs, it is uncertain whether it was consistently employed in the registration process. Lack of visibility of sampled refugees’ names on postboxes could result in undeliverable letters, particularly for those who recently arrived, potentially leading to an underrepresentation of this subgroup in our sample. In defining the population to sample from the EMR, the assumption was made that Ukrainian nationals registered from 24 February onward will be refugees. However, during the screening section of the survey, 97 individuals were screened out (code 4.1) as they did not meet the criteria of being a refugee. Additionally, feedback from 77 individuals indicated they had moved abroad (code 4.2)

Table 3 Final disposition codes for the gross sample according to AAPOR. (Source: IAB-BiB/FReDA-BAMF-SOEP-Data; Authors' own calculation)

Final disposition code	Number	
1.1	Complete interview	10,395
2.11	Refusal	63
2.12	Break-off or partial	2008
2.2	Non-Contact	24,992
2.31	Death	7
2.32	Physically or mentally unable/ incompetent	26
2.4	New address after field period	259
3.1	Nothing known about address	10,076
4.1	Screen out	97
4.2	Moved abroad	77

in the interim. Further details on final disposition codes are available in Table 3. A total of 8695 interviews were submitted online, with an additional 1700 submitted as PAPI.

The following Table 4 provides the number of refugees in the gross sample as well as for the first wave in 2022 by federal state, sex, age group, and marital status. The information displayed for the gross sample and the first wave in 2022 is the information as provided by the EMR. Certainly, the data highlights a notable concentration of refugees in North Rhine-Westphalia and Bavaria, particularly attributed to significant populations in Düsseldorf and Munich, respectively. These urban centers, being major destinations for Ukrainian refugees, contribute substantially to the overall numbers in their respective federal states. For a huge portion of the sample the marital status is unknown. This is mainly because EMRs did not provide this information. Among known statuses, there is diversity, with a majority of the sample being single or married.

For the gross sample, information provided by the EMR (see Table 4) is the only information available to compare refugees who decided to participate in the panel and those who refused to do so. We model the decision to participate in the panel using a logit model. The dependent variable (y) for the model is the decision to participate in the panel ($y=1$ if final disposition code is 1.1 and $y=0$ if final disposition code is one of 2.11, 2.12, 2.2, 2.32, 2.4, and 3.1). We exclude refugees who died (code 2.31), who were screened out of the population (code 4.1), and who moved abroad (code 4.2) from the analysis because they do not belong to the desired population. In the model we estimate, we control for federal state, age, sex, and marital status by inserting dummy-variables. Figure 5 displays the coefficient plot for the model including the dummy variables which significantly influence the decision to participate in the panel study. The model finds male refugees and refugees sampled in Berlin and Hamburg to be less likely to participate in the panel. Refugees being sampled in Saarland and refugees being married are more likely to participate in the panel study.

The weights accompanying the data are generated through a three-step process. Initially, design weights for the two-phase design are calculated following the methodology outlined by Särndal et al. (2003, Chap. 9). The design weights are

Table 4 Total number of refugees in the gross sample and the first wave of the IAB-BiB/FReDA-BAMF-SOEP-Survey. (Source: IAB-BiB/FReDA-BAMF-SOEP-Data; Authors' own calculation)

	Gross sample	Refusals	Panel
<i>Federal state</i>			
Schleswig Holstein	971	767	204
Hamburg	3397	2743	654
Lower Saxony	2625	2054	571
Bremen	341	267	74
North Rhine-Westphalia	9801	7560	2241
Hesse	3404	2692	712
Rhineland-Palatinate	2036	1559	477
Baden-Württemberg	3473	2669	804
Bavaria	8584	6613	1971
Saarland	289	195	94
Berlin	7421	6059	1362
Brandenburg	642	517	125
Mecklenburg-Western Pomerania	734	559	175
Saxony	1785	1434	351
Saxony-Anhalt	1354	1046	308
Thüringen	1143	871	272
<i>Sex</i>			
Unknown	10	9	1
Male	10,086	8110	1976
Female	37,904	29,486	8418
<i>Age group</i>			
Unknown	1604	1233	371
Older than 17 up to 20	2459	1875	584
Older than 20 up to 25	4025	3127	898
Older than 25 up to 30	4586	3681	905
Older than 30 up to 35	6318	4930	1388
Older than 35 up to 40	7387	5758	1629
Older than 40 up to 45	5886	4588	1298
Older than 45 up to 50	4203	3251	952
Older than 50 up to 55	3196	2519	677
Older than 55 up to 60	2681	2137	544
Older than 60 up to 65	3318	2656	662
Older than 65 up to 70	2337	1850	487
Older than 70	0	0	0
<i>Marital status</i>			
Unknown	28,323	22,346	5977
Single	8445	6609	1836
Civil partnership	6	4	2
Married	8548	6515	2033
Divorced	1949	1534	415
Widowed	729	597	132

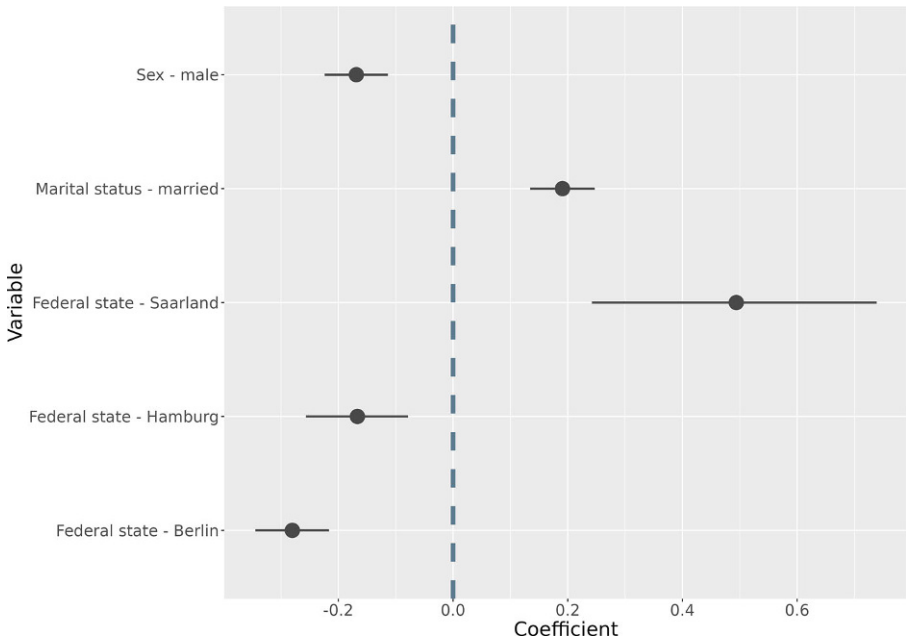


Fig. 5 Coefficient Plot for the model estimating the decision to participate in the panel of the IAB-BiB/FReDA-BAMF-SOEP-Survey. (Source: IAB-BiB/FReDA-BAMF-SOEP-Data; Authors' own calculation)

intended to account for the complexities introduced by the two-phase sampling as well as the systematic (pps) selection. Design weights are then adjusted to account for non-response. This adjustment considers characteristics provided by the EMR (referenced in Table 4). This step addresses potential biases introduced by non-response. Subsequently, raking is applied to margins for sex, age, federal state, and month of immigration, to ensure that the sample distribution aligns with the population distribution, as provided by the AZR effective November 2022. The resulting weights, computed through this process, yield an effective sample size of 8231 and a reasonable design effect of 1.26, computed according to Kish (1992).

7 Conclusions

After Russia's invasion of Ukraine in early 2022, more than one million refugees arrived in Germany. These Ukrainian refugees differ in many aspects from Germany's past forced migration experiences including their demographics, legal status, perceived acceptance, allocation, and the continuity of circular migration flows, not least because of greater geographical proximity. To learn about these recently arriving refugees, their needs, resources, as well as challenges ahead, a survey allowing for generalization to this population was urgently needed. To meet this need quickly, four institutions joined their expertise and resources to create a probability sample for the population of Ukrainian refugees using two different registers: the German

population register and the central register of foreigners. The approach we presented in this paper can be used to draw a sample in the same way for any third country nationals.

The sampling design encompasses a two-phase methodology involving systematic (pps) sampling. This approach may be deemed intricate due to certain drawbacks associated with systematic (pps) sampling, notably the inability to compute second-order inclusion probabilities, thus impeding classical variance estimation (Wolter 2007). Nonetheless, alternative methodologies, such as jackknife or bootstrap methods, emerge as viable options for variance estimation within the context of complex designs.

While utilizing the AZR and the EMR in a two-phase sampling approach, our study is confined to the population of Ukrainian refugees aged 18–70, registered between 24 February 2022 and the time of sampling in the EMR. A comparison between the gross sample drawn from the EMR and the corresponding population registered in the AZR (effective November 2022) reveals a large congruence. The presented analyses show that, even within a geopolitical crisis resulting in large inflows of refugees, the existing registers in Germany constitute a comprehensive sampling frame. Following time and resource constraints, a deliberate decision was taken to concentrate on more populated areas instead of rural areas, aiming towards a larger gross sample.

Comparing the gross and the net sample of Ukrainian refugees with the target population of Ukrainian refugees aged 18–70 that were registered by the end of May 2022 in Germany provides evidence of the overall success of this sampling approach and the high-quality of this probability sample. The paper shows the benefits and feasibility of establishing register-based samples even in contexts of geopolitical crisis and providing sound data within brief time horizons. For politics and practitioners, the data provides comprehensive information for evidence-based decision-making much earlier than in the past. For migration scholars the resulting survey data is highly valuable because this surveying of the target population already starts before selective return and onward movements by the forced migrants have taken place. Moreover, the data can be used in future research and compare results of our study with those of others for comparable measurements. This will be of particular interest to researchers comparing findings from probability and non-probability samples.

Acknowledgements The authors thank all other members of the IAB-BiB/FReDA-BAMF-SOEP project team (in alphabetical order): Herbert Brücker, Martin Bujard, Adriana Cardozo, Markus M. Grabka, Yuliya Kosyakova, Amrei Maddox, Nadja Milewski, Robert Naderi, Wenke Niehues, Nina Rother, Lenore Sauer, Sophia Schmitz, Silvia Schwanhäuser, C. Katharina Spieß, and Kerstin Tanis as well as the members of the project team at infas consisting of Doris Hess, Michael Ruland, and Sabrina Torregroza. Additional thanks to Jan Eberle from the Federal Statistical Office and many colleagues from local registration offices providing invaluable insights into the details as well as their local specifics of the registration procedures in Germany. Finally, the authors thank two anonymous reviewers for their valuable comments and suggestions improving the paper considerably.

Funding Open Access funding enabled and organized by Projekt DEAL.

Conflict of interest H.W. Steinhauer, J.P. Décieux, M. Siegert, A. Ette and S. Zinn declare that they have no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Adam F, Föbker S, Imani D, Pfaffenbach C, Weiss G, Wiegandt C-C (2021) "Lost in transition"? Integration of refugees into the local housing market in Germany. *J Urban Aff* 43(6):831–850. <https://doi.org/10.1080/07352166.2018.1562302>
- Agadjanian V, Zotova N (2012) Sampling and surveying hard-to-reach populations for demographic research: a study of female labor migrants in moscow, russia. *Demogr Res* 26(5):131–150. <https://doi.org/10.4054/DemRes.2012.26.5>
- Amior M (2020) Immigration, local crowd-out and undercoverage bias. (CEP discussion paper 1669). London, UK: Centre for Economic Performance, LSE. <https://cep.lse.ac.uk/pubs/download/dp1669.pdf>. Accessed 1 Dec 2022
- Andreß H-J, Careja R (2018) Sampling migrants in six European countries: how to develop a comparative design? *Comp Migr Stud*. <https://doi.org/10.1186/s40878-018-0099-x>
- Babka von Gostomski C, Pupeter M (2008) Zufallsbefragung von Ausländern auf Basis des Ausländerzentralregisters: Erfahrungen bei der Repräsentativbefragung „Ausgewählte Migrantengruppen in Deutschland 2006/2007“. *Methoden Daten Anal* 2(2):149–177
- Bell M, Charles-Edwards E, Kupiszewska D, Kupiszewski M, Stillwell J, Zhu Y (2015) Internal migration data around the world: Assessing contemporary practice. *Popul Space Place* 21(1):1–17. <https://doi.org/10.1002/psp.1848>
- Beqiri M, Cierpiał-Wolan M (2022) Ukrainian refugees in Poland survey 2022—Preliminary findings. Poland. <https://data.unhcr.org/en/documents/download/97052>. Accessed 2 Dec 2022
- Bloch A (1999) Carrying out a survey of refugees: some methodological considerations and guidelines. *J Refug Stud* 12(4):367–383. <https://doi.org/10.1093/jrs/12.4.367>
- Bogumil J, Burgi M, Kuhlmann S, Hafner J, Heuberger M, Krönke C (2018) Bessere Verwaltung in der Migrations- und Integrationspolitik: Handlungsempfehlungen für Verwaltungen und Gesetzgebung im föderalen System. *Nomos, Baden-Baden* <https://doi.org/10.5771/9783845295862>
- Brücker H, Rother N, Schupp J (2016) IAB-BAMF-SOEP-Befragung von Geflüchteten: Überblick und erste Ergebnisse. *Politikberatung kompakt* 116. Berlin. https://www.diw.de/documents/publikationen/73/diw_01.c.547162.de/diwwkompakt_2016-116.pdf. Accessed 19 Sept 2022
- Brücker H, Ette A, Grabka M, Kosyakova Y, Niehues W, Rother N, Tanis K (2023) Ukrainian refugees in Germany: evidence from a large representative survey. *Comp Popul Stud* 48:395–424. <https://doi.org/10.12765/CPoS-2023-16>
- Cornesse C, Blom A, Dutwin D, Krosnick J, De Leeuw E, Legleye S, Wenz A (2020) A review of conceptual approaches and empirical evidence on probability and nonprobability sample survey research. *J Surv Stat Methodol* 8(1):4–36. <https://doi.org/10.1093/jssam/smz041>
- Diehl C, Lubbers M, Mühlau P, Platt L (2016) Starting out: New migrants' socio-cultural integration trajectories in four European destinations. *Ethnicities* 16(2):157–179. <https://doi.org/10.1177/1468796815616158>
- Dollmann J, Mayer S, Lietz A, Siegel M, Köhler J (2022) Setting up an offline recruited online access panel with an oversampling of immigrants and their descendants: the German deZIM.panel <https://doi.org/10.31235/osf.io/mdpnx>
- Dražanová L, Geddes A (2022) Attitudes towards Ukrainian refugees and governmental responses in 8 European countries. <https://www.asileproject.eu/attitudes-towards-ukrainian-refugees-and-governmental-responses-in-8-european-countries/> (Created September). Accessed 3 Jan 2023
- El-Kayed N, Hamann U (2018) Refugees' access to housing and residency in German cities: internal border regimes and their local variations. *Soc Incl* 6(1):135–146. <https://doi.org/10.17645/si.v6i1.1334>

- Ette A, Décieux J, Erlinghagen M, Auditor J, Sander N, Schneider N, Witte N (2021) Surveying across borders: the experiences of the German emigration and remigration panel study. In: Erlinghagen M, Ette A, Schneider N, Witte N (eds) *The global lives of German migrants*. Springer, Cham, pp 21–39 https://doi.org/10.1007/978-3-030-67498-4_2
- European Union Agency for Asylum (2022) Surveys of arriving migrants from Ukraine. https://euaa.europa.eu/sites/default/files/publications/2022-06/2022_06_14_EUAA_SAM_UKR_Factsheet.pdf. Accessed 13 Oct 2022
- Federal Ministry of the Interior and Community (2022) Geflüchtete aus der Ukraine. https://www.bmi.bund.de/SharedDocs/downloads/DE/veroeffentlichungen/nachrichten/2022/umfrage-ukraine-fluechtlinge.pdf?__blob=publicationFile&v=3. Accessed 13 Oct 2022
- Federal Office for Migration and Refugees (2022) Questions and answers on entering Ukraine and staying in Germany. https://www.bamf.de/SharedDocs/Anlagen/DE/AsylFluechtlingsschutz/faq-ukraine-en.pdf?__blob=publicationFile&v=16. Accessed 3 Jan 2023
- Groves R (2006) Nonresponse rates and nonresponse bias in household surveys. *Public Opin Q* 70(5): 646–675. <https://doi.org/10.1093/poq/nfl033>
- Hunkler C, Scharrer T, Suerbaum M, Yanasmayan Z (2022) Spatial and social im/mobility in forced migration: revisiting class. *J Ethn Migr Stud*. <https://doi.org/10.1080/1369183X.2022.2123431>
- Jacobsen J, Siegert M (2023) Establishing a panel study of refugees in Germany: first wave response and panel attrition from a comparative perspective. *Field Methods*. <https://doi.org/10.1177/1525822X231204817>
- Kalton G (2014) Probability sampling methods for hard-to-sample populations. In: Tourangeau R, Edwards B, Johnson T, Wolter K, Bates N (eds) *Hard-to-survey populations*. Cambridge University Press, Cambridge, pp 401–423 <https://doi.org/10.1017/CBO9781139381635.024>
- Kish L (1992) Weighting for unequal Pi. *J Off Stat* 8(2):183–200
- Kley S (2017) Facilitators and constraints at each stage of the migration decision process. *Popul Stud* 71(Sup1):35–49. <https://doi.org/10.1080/00324728.2017.1359328>
- Kohlenberger J, Buber-Ennser I, Rengs B, Setz I, Riederer B (2022) *UkrAiA Abschlussbericht Stadt Wien. Austria*. https://www.ukraia.at/wp-content/uploads/2022/08/ukraia_final_report_city_of_vienna.pdf. Accessed 2 Dec 2022
- Kühne S, Jacobsen J, Kroh M (2019) Sampling in times of high immigration: the survey process of the IAB-BAMF-SOEP survey of refugees. *Survey methods: insights from the field*. <https://doi.org/10.13094/SMIF-2019-00005>
- Lattof S (2018) Collecting data from migrants in Ghana: Lessons learned using respondent-driven sampling. *Demogr Res* 38(36):1017–1058. <https://doi.org/10.4054/DemRes.2018.38.36>
- MacInnis B, Krosnick J, Ho A, Cho M-J (2018) The accuracy of measurements with probability and Nonprobability survey samples: replication and extension. *Public Opin Q* 82(4):707–744. <https://doi.org/10.1093/poq/nfy038>
- Massey D (2014) Challenges to surveying immigrants. In: Tourangeau R, Edwards B, Johnson T, Wolter K, Bates N (eds) *Hard-to-survey populations*. Cambridge University Press, Cambridge, pp 270–292 <https://doi.org/10.1017/CBO9781139381635.017>
- McKenzie D, Mistiaen J (2009) Surveying migrant households: a comparison of census-based, snowball and intercept point surveys. *J R Stat Soc Ser A* 172(2):339–360
- Pędziwiatr K, Brzozowski J, Nahorniuk O (2022) *Refugees from Ukraine in Kraków*. <https://owim.uek.krakow.pl/wp-content/uploads/user-files/reports/OWIMrefugees.pdf>. Accessed 2 Dec 2022
- Pötzschke S, Weiß B, Hebel A, Piepenburg J, Poppek O (2022) Geflüchtete aus der Ukraine – Erste deskriptive Ergebnisse einer Onlinebefragung in Deutschland und Polen <https://doi.org/10.34879/gesisblog.2022.60>
- Särndal C-E, Swensson B, Wretman J (2003) *Model assisted survey sampling*. Springer, New York
- Sauer L, Kraus E, Kassam K, Schühler S, Pupeter M, Wolfert S, Schneekloth U (2022) *Forced migration and transnational family arrangements—Eritrean and Syrian refugees in Germany (TransFAR): methodology report. BiB data and technical reports 1/2022*, Wiesbaden. <https://d-nb.info/1256734772/34>. Accessed 12 Jan 2023
- Sosenko F, Bramley G (2022) Smartphone-based Respondent Driven Sampling (RDS): A methodological advance in surveying small or ‘hard-to-reach’ populations. *PLoS One* 17(7):e270673. <https://doi.org/10.1371/journal.pone.0270673>
- Steinhauer H, Zinn S, Will G (2019) Sampling refugees for an educational longitudinal survey. *Survey methods: insights from the field*. <https://doi.org/10.13094/SMIF-2019-00007>

- Steinhauer H, Siegers R, Siegert M, Jacobsen J, Zinn S (2022) Sampling, nonresponse, and weighting of the 2020 refreshment sample (M6) of the IAB-BAMF-SOEP refugee panel. SOEP survey paper 1104. Berlin. <https://www.econstor.eu/handle/10419/261440>. Accessed 1 Dec 2022
- The American Association for Public Opinion Research (2016) Final dispositions of case codes and outcome rates for surveys
- Tillé Y (2006) Sampling algorithms. Springer, New York
- Tourangeau R (2014) Defining hard-to-survey population. In: Tourangeau R, Edwards B, Johnson T, Wolter K, Bates N (eds) Hard-to-survey populations. Cambridge University Press, Cambridge, pp 2–20 <https://doi.org/10.1017/CBO9781139381635.003>
- Tyldum G (2021) Surveying migrant populations with respondent-driven sampling. Experiences from surveys of east-west migration in Europe. *Int J Soc Res Methodol* 24(3):341–353. <https://doi.org/10.1080/13645579.2020.1786239>
- United Nations High Commissioner for Refugees (2022a) Global trends—Forced displacement in 2021. Copenhagen. <https://www.unhcr.org/media/global-trends-report-2021>. Accessed 2 Jan 2023
- United Nations High Commissioner for Refugees (2022b) Lives on hold: profiles and intentions of refugees from Ukraine. <https://data.unhcr.org/en/documents/download/94176>. Accessed 13 Oct 2022
- Weber R, Saarela J (2019) Circular migration in a context of free mobility: Evidence from linked population register data from Finland and Sweden. *Popul Space Place* 25(4):e2230. <https://doi.org/10.1002/psp.2230>
- Wenzel L, Husen O, Sandermann P (2022) Surveying diverse Subpopulations in refugee studies: reflections on sampling, implementation, and translation strategies drawn from experiences with a regional quantitative survey on refugee parents in Germany. *J Refug Studie*. <https://doi.org/10.1093/jrs/feac043>
- Willekens F, Massey D, Raymer J, Beauchemin C (2016) International migration under the microscope. *Science* 352(6288):897–899. <https://doi.org/10.1126/science.aaf6545>
- Wolter KM (2007) Introduction to variance estimation. Springer, New York

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.