

Realizing a Global Survey of Emigrants through Facebook and Instagram

Steffen Pöttschke* & Bernd Weiß

GESIS – Leibniz Institute for the Social Sciences
(Mannheim, Germany)

October 11, 2021

Abstract: Research on international migrants has seen a sharp increase during the last decades, yet sampling them remains a major challenge, especially in a cross-national setting and on a global scale. While various sampling methods are established in the field, most of them cannot easily be implemented globally due to their dependence on specific administrative or infrastructure elements or simply their costs. Since Social Networking Sites (SNS) operate on a global scale, they provide a sampling frame that can be utilized for the targeted recruitment of migrants worldwide. Increasingly used for research purposes and among the largest and most popular SNSs are Facebook and Instagram. In our project GEOOS (*German Emigrants Overseas Online Survey*), we utilize paid advertisements on these networks to target German emigrants, particularly Germans living outside of Europe. Our research aims to ascertain whether such ads could be used to recruit a nonprobability (migrant) sample on a global scale. More specifically, we are interested in the success of this approach concerning three performance indicators: Cost efficiency, coverage, and sample size. Our advertisement campaign ran for 18 days and resulted in total costs of about 2,223 Euro. This investment led a total of 3,895 individuals to complete the survey; of those, 98 percent belonged to the target population, meaning they were (a) either born in Germany or held German citizenship and (b) did not live in Germany. GEOOS participants lived in a total of 148 countries and territories around the globe. Similar to findings reported in previous studies on this target population, the largest sub-groups resided in predominantly Anglo-phone countries; however, taken together, participants in these countries only constitute 38 percent of our overall sample, with nearly a quarter of GEOOS participants ($n = 867$) living in Middle and South America, 862 residing in Asian countries, and 476 in Africa. Furthermore, a considerable share of our sample is constituted by individuals who would either not have been included in a sampling frame based on German population registers or who would have been unlikely to be reached through this method due to incomplete or outdated information.

Keywords: SNS sampling, non-probability sampling, migration research, emigrants, Facebook, Instagram, global survey

* Corresponding author (steffen.poetschke@gesis.org)

1 Introduction

The mobility of people across shorter and longer distances is as old as humankind itself. Since the end of the 20th century, and in the context of rapid globalization of almost all aspects of social and cultural life, mobility and migration are also increasingly in the focus of public and academic interest. Accordingly, research on international migrants has seen a sharp increase during the last decades (Pisarevskaya et al. 2020). However, migrants can, in most settings, be considered a hard-to-reach population (Tourangeau 2014). Therefore the sampling of respondents remains a significant challenge, especially in a cross-national setting where the research design necessitates the use of identical or at least highly comparable sampling methods in different countries (Careja and Andreß 2018; Font and Méndez 2013). Furthermore, quantitative migration research usually focuses on immigrants in specific countries. This means that researchers generally limit their analysis to a single country and investigate, for instance, how migrants fare in them or what consequences their arrival has for the respective societies (Mieriņa 2019). Alternatively, some surveys look at migrant populations in a predetermined (small) set of countries to achieve specific insights from a comparative perspective (cf. Ersanilli and Koopmans 2013; Recchi et al. 2019; Schoumaker and Beauchemin 2015).

Various sampling methods are applied in migration research, such as population register-based sampling (Careja and Bevelander 2018), name-based telephone sampling (Humpert and Schneiderheinze 2000; Prandner and Weichbold 2019), random route, and location sampling (Reichel and Morales 2017); yet they usually presuppose that scholars focus their research efforts on a limited number of selected countries or even smaller administrative units. Consequently, these methods could not easily be implemented globally for cross-national research purposes due to their dependence on specific administrative or infrastructure elements (e.g., population registers), because of the sheer costs a global implementation would occur (e.g., training and deployment of local enumerators) or since the selection of countries is not known to researchers prior to the start of the fieldwork. In contrast, Social Networking Sites (SNS) such as Facebook, Instagram or Twitter (with Instagram being a part of Facebook, Inc.)¹, which are amongst the largest and most popular (Datareportal 2021), operate on an almost global scale. These SNS are increasingly used for research purposes and provide a sampling frame that can be utilized for the recruitment of emigrants worldwide.

A major limitation of SNS-based recruitment, though, is that these networks usually do not offer full access to all individual users in order to invite them to the survey, meaning that there is no register or databank listing all users of Facebook and Instagram available to scholars that could be used directly to draw a sample. While earlier studies (cf. Baltar and Brunet 2012; Brickman Bhutta 2012) employed mostly personal messages and survey invitations posted on boards and in user groups, more recently,

¹ In the following, we use Facebook Inc. to refer to the company of the same name and Facebook, as well as Instagram, when referencing the social networking sites which that company operates.

scholars engaged in survey-based research have increasingly turned to paid ads to target user groups with specific characteristics and to increase the number of potential respondents.

It is important to note that, when employing SNS for sampling purposes, the researcher does not have control over any nonresponse mechanisms and respondents self-select into the sample, which results in a nonprobability sample (Cornesse et al. 2020). Nevertheless, SNSs also offer major benefits such as their nearly worldwide accessibility, low sampling costs, and the possibility to recruit a high number of participants within a short timeframe (Grow et al. 2020; Kühne and Zindel 2020).

Overall, an important aim of our project, the *German Emigrants Overseas Online Survey* (GEOOS) was to ascertain whether advertisements in social networking sites, more specifically Facebook and Instagram, could be used to recruit migrants of a specific origin in a multitude of countries, ideally on a global scale. More specifically, in the present study, we are interested in three performance indicators which result in the following research questions: (1) *Relative cost efficiency*: How does the method perform at a global level? Are there relevant differences in costs etc., between regions? (2) *Coverage*: (2a) Are we able to recruit participants on all continents? (2b) How does the method perform on continents and in regions where previous studies using established methods were only able to recruit few participants? However, given the comparatively small budget of our project (see below), we can only look at this question in relative terms. (2c) Were we able to recruit respondents who would have been missed by other recruitment methods? Could this approach, therefore, complement more established recruitment methods in a meaningful way? (3) *Sample size*: Did we succeed in reaching sample sizes comparable to projects that used other sampling approaches? More precisely: How effective was our approach in comparison to those employed by previous studies?

Our project adds to the literature as it constitutes the first study that embarks to sample *emigrants* via advertisements on Facebook and Instagram on a (nearly) global level. More specifically, while there has been a small number of cross-national studies using advertisements for sampling purposes, they only targeted a comparative small number of countries each. Finally, we contribute to the small but growing body of research that expands the use of SNS sampling beyond countries in the global north (cf. Olamijuwon 2021; Rosenzweig et al. 2020; Samuels and Zucco 2014; Thomson and Ito 2014; UNESCO and Metropolis Canada 2020).

Even though there have by now been a number of publications describing the recruitment of survey participants through Facebook and Instagram, we go into some details as – to the best of our knowledge – none of the previous studies attempted to sample respondents from a single origin country on a global level, or any other target group in a cross-national survey of comparable geographic scope, for this matter. The remainder of the article is structured as follows: First, we provide a brief overview regarding previous quantitative studies of emigrants, particularly but not exclusively from Germany, and the use of SNS in survey research. Secondly, we will describe our sampling approach in some detail. This will

be followed by an account of the results of our targeting efforts and a short description of our sample. The paper closes with some conclusions.

2 State of the art

2.1 Surveying emigrants

Emigrants are a particularly hard-to-reach target population as they are potentially scattered over a large number of countries, which are, in many cases, unknown to researchers before the actual fieldwork starts. Consequently, the small number of projects that previously tried to survey emigrants irrespective of their countries of residence had to develop specific sampling approaches. In their study of skilled German workers, academics, and executives abroad, Pfeiffer and Heimer (2007), for example, contacted individuals and institutions which could function as multipliers to inform emigrants about their survey and strategically placed information in publications of business organizations. While this approach resulted in a sample sufficiently large for quantitative analysis ($n = 1,410$), its diversity in terms of country of residence was limited as 79 percent of the respondents lived in Europe or North America, and four countries (USA, Great Britain, Switzerland, and Canada) alone accounted in sum for more than half of the sample (Pfeiffer and Heimer 2007:19). Another project, *The Emigrant Communities of Latvia* (Kaša and Mieriņa 2019), used a partially similar yet even more complex strategy, combining a wide array of methods to reach and recruit emigrants. These included private and governmental multipliers, a dedicated media strategy, and engagement with diaspora groups on a number of social networking sites, for example, Odnoklassniki, Vkontakte, and Facebook (Mieriņa 2019). The resulting dataset included information on 14,068 respondents (including 4,784 partially completed surveys) residing in 118 countries, with the majority participating from EU member states (including the UK), Norway, Russia, Canada, and the US (Mieriņa 2019:16–17). While the strategies followed by these projects certainly have their merits, it is not clear in how far they could easily be replicated for other target populations. The approach used by the above-mentioned study targeting German emigrants worked particularly well because the target population was defined by its professional status. Furthermore, this and the Latvian study both heavily relied on the existence and cooperation of suitable multipliers which effectively worked as gatekeepers. Consequently, the success of these recruitment strategies depends, at least in part, also on researchers' affiliation, networks, and institutional support for the corresponding projects.

Another set of studies that previously tried to globally survey emigrants of specific nationalities used an “origin-based sampling approach”, meaning that they relied on infrastructures or administrative records in the country of origin as sampling frames. The first example in this regard consists of studies that used existing surveys to identify and recruit members of their target population. Schupp and colleagues (2008), for example, investigated whether former participants of the *German Socio-Economic Panel* (GSEOP) who left Germany between 2002 and 2005 could be recontacted in the context of an emigrant study. While this approach did work in general, it yielded only 23 completed interviews, partially due

to the small number of eligible dropouts from the GSOEP ($n = 228$) and an even smaller number of individuals for whom valid addresses could be identified ($n = 52$). Following a similar approach, Ghimire et al. (2019) asked participants of the Nepalese *Chitwan Valley Family Study* (CVFS) to identify individuals whom they considered members of their households but who lived abroad at the time of the interview. If respondents reported having such family members, they were asked to provide their contact information. Following this approach, the research team was able to identify 1,009 eligible individuals in 37 countries, of whom 915 eventually participated in their survey. This success can, in part, be explained by the extraordinary degree of flexibility employed during the data collection process (see Ghimire et al. 2019:1194–95 for details). Blaskó (2015) presents the results of a study that replicated the aforementioned approach, building on the Hungarian sample of the *Labour Force Survey*. However, in this case, the first stage of the sampling process provided the researchers only with contact details of 546 individuals, of which 125 eventually participated in the survey. Hence, similar to the GSOEP study, this project provided important methodological insights but did not yield a sample big enough for many statistical analyses.

Another origin-based strategy has recently been employed by the *German Emigration and Remigration Panel Study* (GERPS) (Erlinghagen et al. 2021), building on an earlier pilot study (Ette et al. 2015) conducted largely by the same scholars. In this case, German population registers were used to draw a probability sample of German nationals who deregistered with the intention to move abroad between July 2017 and June 2018. Data collection for the first wave of the GERPS project took place from November 2018 to February 2019 meaning that the sampling frame covered German nationals who deregistered their address, with the purpose of moving abroad no more than 16 months before the fieldwork started. To increase participation rates, different monetary and non-monetary incentives were used (Ette et al. 2020). The use of population registers as a sampling frame for research on German emigrants is possible because any resident of Germany is obliged by law to register changes of residence with the authorities within 14 days, including relocation to other countries. When doing so, individuals can provide a new address abroad on a voluntary basis. In a multistage probability sampling procedure, GERPS selected German nationals marked as emigrants from the population registers, resulting in a gross sample of 26,273 individuals. However, of those, only slightly more than one-third (35 percent) provided a new address abroad to which an invitation to participate in the web survey could be sent by postal mail. An equivalent invitation was sent to the last known German address of the remaining sample members; the hope being that they might either have placed a forwarding order with postal services or arranged for someone to check for or receive mail in their name (see Ette et al. 2020 for detailed descriptions of the sampling procedure and contact strategy). This sampling strategy resulted in a total of 4,525 interviews completed by German nationals living abroad. The described sampling approach was most successful in Europe, where nearly three-quarters of the interviewed emigrants resided ($n = 3,344$; own calculation using Erlinghagen and Schneider 2020) (also see Section 5.2). It worked less effectively in other world regions, though. While still a comparatively large combined subsample was

achieved in the USA and Canada ($n = 456$), the subsamples in the Near and Middle East (incl. Turkey; $n = 111$), and Africa ($n = 77$) were much smaller. The project's methods report furthermore states that emigrants who did not provide a (valid) address abroad had a 22 percent lower likelihood of participating in the survey (Ette et al. 2020:66). Additionally, its authors point out that the probability of sampled emigrants to participate declined the longer before the survey their emigration had been registered, reaffirming similar findings from the project's pilot study (Ette et al. 2015). As the authors hypothesize, this is, at least in part, likely due to continuing mobility of emigrants during the early stages of their migration.

Hence, while the register-based sampling approach is indeed innovative for research focusing on emigrants, its success depends on various factors. Firstly, thinking beyond the German example, it can only be implemented for emigrants from countries that pose equivalent obligations on their citizens to register their residence and any change thereof, even in the case of emigration. Furthermore, the method works best in cases in which emigrants provide a valid address in their new country of residence, which apparently only a fraction of German emigrants does. Regarding provided addresses, the GERPS project and the related pilot study clearly showed that the probability of successful contact is highest shortly after the deregistration of the respective individuals. This effect was visible in the both the samples of GERPS and its pilot study (Ette et al. 2015, 2020) even though the used addresses were fairly recent in both cases. These findings are not that surprising as it might be assumed that a relevant portion of emigrants, including those who are willing to share their new addresses with German authorities, might have arranged only for short or mid-term accommodation in their target country that would allow them to search for a more permanent home once they had a chance to familiarize themselves with their new hometowns. Hence, it can be concluded that his method is somewhat less suited to sample emigrants who left their country of origin several years before a planned research project. Additionally, there might be a risk of sampling bias due to differences in cultural and administrative norms as, especially, some non-European countries might use address formats that emigrants are not fully familiar with, leading to invalid records and possibly a higher non-contact rate even for those individuals who registered a new address. While Japanese addressing system (Universal Postal Union 2021) differs, for example, greatly from the German one, other countries might not even have fully implemented addressing systems (Universal Postal Union 2012). Finally, and on a minor note, it is worth mentioning that using German population registers for sampling is generally a rather complex exercise binding substantial resources (Salentin and Schmeets 2017); consequently, it might not always be a feasible option for research projects with limited funds.

Despite these limitations, GERPS showed that German population registers can, especially with regard to European target countries, be used to achieve high-quality probability-based samples of individuals who registered their emigration with the authorities in recent months prior to the sampling for a given project.

2.2 Advertisements on social networking sites as a sampling tool

Facebook and other social networking sites have increasingly been used to recruit participants for surveys of hard-to-reach populations during the last decade. Initially, researchers mainly employed snowball strategies, most often posting survey invitations in specific groups, directly contacting group members, or asking gatekeepers to forward corresponding messages (Baltar and Brunet 2012; Brickman Bhutta 2012). However, in recent years scholars increasingly turned to the use of advertisements as a recruitment tool. A systematic literature review by Whitaker et al. (2017) lists, for example, a total of 35 studies that have been published between 2012 and 2017 on medical and health research-related topics using Facebook advertisements to recruit participants. Other studies used ads on Facebook, and to a lesser degree Instagram, to survey, for example, voluntary and forced international migrants (Carlini et al. 2015; Elçi, Kirişoğlu, and Üstübcü 2021; Ersanilli and van der Gaag 2020; Pötzschke and Braun 2017), members of LGBTQ communities (Guillory et al. 2018; Kühne and Zindel 2020), supporters of conspiracy myths (Iannelli et al. 2018), and employees in specific sectors of the labor market (Schneider and Harknett 2019). Surveys targeting the general population in selected countries have, for instance, investigated opinions on climate change (Zhang et al. 2020) and collected timely cross-national data during the COVID-19 pandemic (Grow et al. 2020). A small but growing number of studies has, furthermore, started to expand this approach to other SNS such as V Kontakte and Odnoklassniki (Ersanilli and van der Gaag 2020; Rocheva, Varshaver, and Ivanova forthcoming) targeting Eastern-European populations.

To the best of our knowledge, our study is the first that embarks to sample emigrants via advertisements on Facebook and Instagram on a (nearly) global level, i.e., without preselecting a small number of countries of residence. While McAleese et al. (2016) also used Facebook to reach potential participants for their study of Irish emigrant health professionals, they did not employ advertisements but contacted users through messages and posts in specific Facebook groups. As mentioned above, the emigrant communities of Latvia survey employed Facebook and other SNS in a similar way as part of a complex sampling strategy to reach as many Latvian emigrants as possible (Mieriņa 2019).

In general, previous cross-national studies that used advertisements on SNS to recruit survey participants usually targeted only a comparatively small set of countries each (cf. Grow et al. 2020; Olamijuwon 2021; Pötzschke and Braun 2017; UNESCO and Metropolis Canada 2020). In an early project that was somewhat more ambitious with regard to its geographic scope, Thomson and Ito (2014) tried to sample Facebook users in 20 countries for a survey on privacy concerns. They collected valid responses from 17 countries ($n = 399$). However, the authors were not satisfied with their response rates as their survey registered 15 observations or less in 12 of these 17 countries.

3 Sampling strategy, advertisement design, and campaign settings

The following describes the GEOOS project's sampling strategy and the main features of its advertisement campaign. While key elements of advertisement campaigns in Facebook's SNS are

discussed as far as necessary, earlier publications provide more in-depth descriptions of such campaigns' logic and structure (cf. Kühne and Zindel 2020; Pöttschke forthcoming).

3.1 Study design

The present study's aim was to recruit Germans living *outside* of Europe for a web survey by employing targeted advertisement campaigns on Facebook and Instagram. The decision to focus the sampling efforts on other continents was a strategic one, as – in comparison – German Emigrants in European countries are (a) easier to reach and (b) knowledge regarding German emigrants in overseas is particularly scarce.² For sampling purposes, German emigrants were defined as individuals born in Germany and/or holding German citizenship who live outside Germany. The survey questionnaire included several additional instruments that could be used to refine this definition for analytical purposes, such as year of settlement in the country of residence and origin of both parents. None of the survey participants were promised or did receive any monetary or material incentives for their participation.

We used Facebook Inc.'s advertisement manager (FAM, <https://www.facebook.com/adsmanager>) to place ads and manage advertisement campaigns on both Facebook and Instagram. The advertisements included links that led users to an externally hosted web survey. Additionally, a dedicated Facebook page was established to provide SNS users with information about the project, research team, and our research institute. This page also included a link that would guide visitors to the survey's landing page. A third avenue of recruitment was constituted by a snowball element. More specifically, towards the end of the survey, participants were invited to pass on a link to other German emigrants they know, encouraging them to participate. Such a snowball component can be used to identify and mitigate biases associated with samples recruited solely through SNS because participants recruited by Facebook or Instagram might themselves encourage other emigrants to participate in the survey who are no (active) users of either network. However, it was anticipated that this additional element would only yield a small number of additional observations as incentives were neither offered for the initial participants nor for any additional respondents they might recruit. The survey was setup in a way that allows retracing each participant's path of recruitment.

3.2 Design of an advertisement campaign

Through the FAM, Facebook Inc. offers the possibility to target specific user sub-populations on Facebook and Instagram. Advertisers can select specific targeting variables to define the desired audience for their advertisements. GEOOS employed four indicators: (1) a variable specifying users' country of residence, (2) the 'Lived in Germany' indicator that belongs to the target variable sub-category 'expats' (in the following: expat variable), (3) age, and (4) gender. Given the above-stated definition of the survey's target population, the first two variables were most central to our sampling

² The Russian Federation and Turkey were both included in the group of Asian countries as large portions of their respective national territory lie on this continent.

efforts. Both have been employed successfully in previous research targeting migrants in selected countries (Ersanilli and van der Gaag 2020; Pötzschke and Braun 2017).

The structure of advertisement campaigns and the basic logic applied by the targeting algorithm need to be taken into account in order to use advertisements on Facebook and Instagram effectively as a sampling tool. In short, these advertisement campaigns consist of three levels: (1) a *campaign* that includes (2) one or more *ad sets* which (3) feature one or more *advertisements* that comprise of, for instance, varying visual (e.g., pictures or videos) or textual stimuli. All targeting variables are defined at the ad set level. Importantly, over the lifetime of a campaign, the algorithm used to deliver advertisements will increasingly favor those advertisements that perform best *within a given ad set*, i.e., that generate the most clicks to an external website. Arcia (2014) pointed out that this might introduce additional bias as users who share certain traits might be more inclined to click on advertisements than others. If several ads with different pictures are used to appeal to a target group consisting of users of varying backgrounds this can be problematic. More precisely this is the case if subgroups that share specific socio-demographic traits are assumed to be more numerous than others and to prefer certain pictures, that might not appeal strongly to members of smaller subgroups. Furthermore, previous research has indicated that the same principle is likely to apply to ad sets targeting more than one region, meaning that the algorithm seems not to deliver ads evenly distributed across geographic units, such as countries, but to increase ad delivery in those locations in which they perform well (Pötzschke and Braun 2017). These examples highlight that it is advisable that researchers identify variables that could introduce significant bias in their sample and stratify campaigns at the ad set level accordingly. At the same time, it should be kept in mind that this substantially increases the complexity of the sampling procedure and potentially its costs, as smaller target populations might be more expensive to reach than larger ones. Hence, a project should use as many ad sets as necessary but as few as possible.

A central goal of this study was to ascertain whether the method could be used to achieve a diverse sample in terms of respondent's country of residence. Consequently, the geographic distribution of reached Facebook and Instagram users is important. However, for logistical and budget reasons, it was not possible to target each non-European country individually. Therefore, we decided to differentiate between larger regions (i.e., continents) in a first step. However, many of those still contained large sets of countries for which it could be assumed that the size of the target population in them varied greatly. Therefore, Facebook's Graph API (Application Programming Interface) was used to programmatically extract the number of monthly active users (MAU) for each country and territory³. Based on these figures, countries, and territories were grouped into two to three sub-groups per continent. The first group is constituted in each case of those countries for which the highest MAU per continent was reported. Apart from North America and Oceania, the remaining countries and territories were then

³ The geo-political country level differentiation used in FAM lists selected subnational units and dependent territories as separate units alongside the countries they belong to (e.g., Puerto Rico separate from USA, Hong Kong and Macau in addition to China; see Appendix 1).

again split into two groups, based on their geographic distribution and the MAU values. This resulted in 13 geographic clusters shown in Figure 1. A full list of the targeted countries and territories, including their MAU values, is provided in Appendix 1.⁴ It should be noted that Facebook and Instagram are not (officially) available in all countries; consequently, some countries (namely Cuba, Iran, Sudan, Syria, and North Korea) could not be set as target locations of advertisements at the time of the survey.

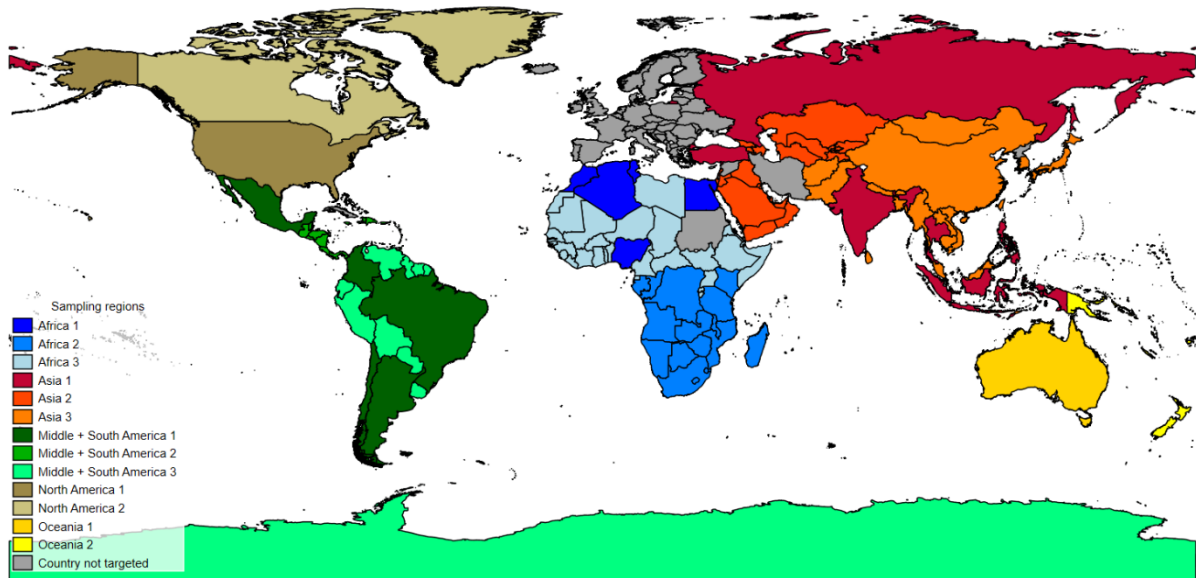


Figure 1: Overview of targeted countries and targeting regions

To reduce gender and age bias introduced by Facebook Inc.'s advertisement algorithm, distinct ad sets were used to targeted male and female users in two age cohorts separately. Gender is only offered as a binary variable in FAM. However, for the target population of this study, both individual figures summed up to the total displayed when the option 'all' was selected. Consequently, it seemed that no members of the target population were excluded by using the two available variables. Additionally, the option to select a third gender was included in the survey questionnaire. The mean age of the target population within each of the 13 regions (values provided in FAM) was used to determine the cut-point between age cohorts, aiming in each case at targeting two cohorts of comparable size. Therefore, the structure of the age cohorts slightly differs between sampling regions. The combination of thirteen sampling regions, two gender groups, and two age cohorts resulted in a total of 52 ad sets (see Table 1).

We decided to place the advertisements for this survey on Facebook and Instagram and to include users of mobile devices as well as desktop PCs. This decision was taken to reach as many users of varying sociodemographic backgrounds in as many locations as possible. Earlier research had shown that the user base of SNS and user's behavior on them differ between socio-demographic groups (Beisch and Schäfer 2020; Gambo and Özad 2020). Furthermore, it had to be taken into account that the use of

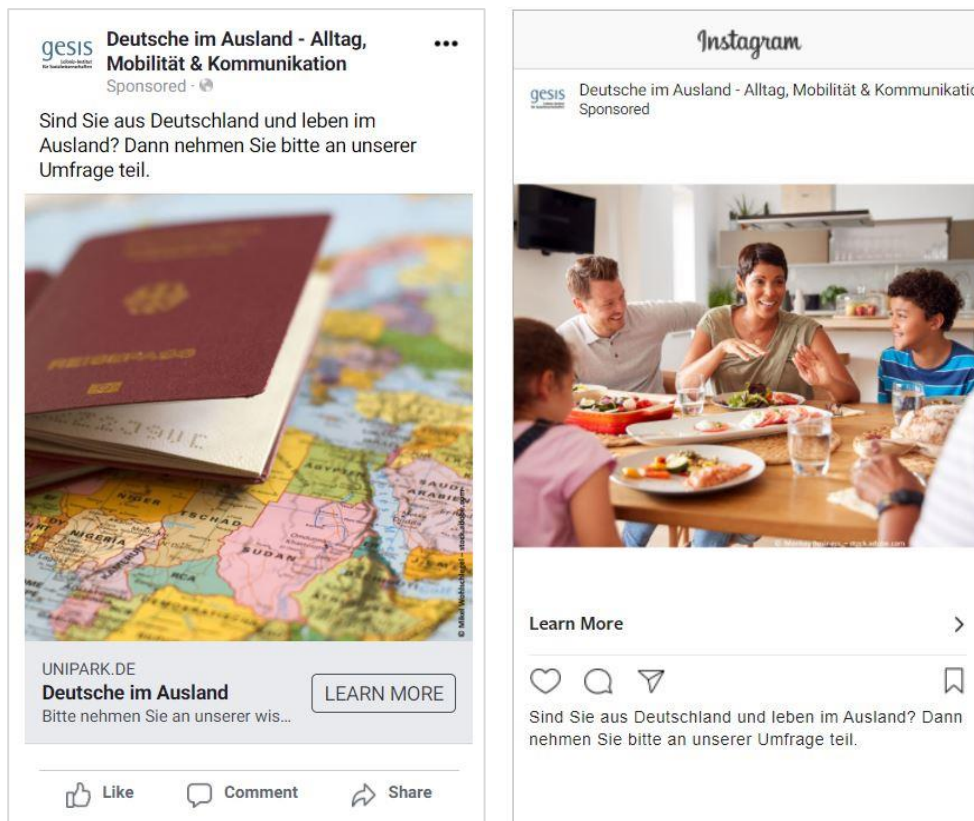
⁴ Taking the geographic location of the larger part of their national territories as reference, we included both Turkey and the Russian Federation in the group of Asian countries for our sampling purposes. Due to its small population, Antarctica was included in the sampling region 'Middle and South America 3'.

computational devices in general (ARD/ZDF Forschungskommission 2021) and while responding to web surveys, in particular, varies between genders and age groups and may depend on general preference for certain devices (Haan, Lugtig, and Toepoel 2019; Keusch and Yan 2017). Finally, the usability of devices might also differ between specific locations depending on the infrastructure and living situation of potential respondents.

3.3 Design of advertisements

The advertisements themselves were designed to appear as similar as possible on both SNS. With this goal in mind, we opted to have them displayed in user's *Facebook news feed* or their *Instagram feed*. For more information on these and other placement options, see Facebook (2021c). This placement meant that the ads would have a very similar optic on both networks and across the different device types. Additionally, we used the placement option *Facebook right column* for users of personal computers. All advertisement texts and the accompanying Facebook page were completely in German. This means that language served as an indirect screening element, even though we did not use it as a targeting variable.

Our advertisements were identical in all regards except for the picture they used. We selected a total of 13 images so that the advertisements would appeal to individuals of different backgrounds, with diverging interests, and tastes. Consequently, the motives featured different content reaching from objects related to the survey's topic (e.g., passports), over pictures of German landmarks to such showing individual persons, individuals holding German flags, couples, or families. Figure 2 shows two examples of advertisements used in the GEOOS project.



Copyright of the pictures: Mikel Wohlschlegel / stock.adobe.com (passport, left); Monkey Business / stock.adobe.com (family, right).

Figure 2: Examples of advertisements as shown on mobile devices on Facebook (left) and Instagram (right)

The header of the advertisements contained the survey’s title “Germans abroad – Daily live, mobility and communication”, while the text of the actual advertisement reads “Are you from Germany and live abroad? If so, please participate in our survey”. Furthermore, on Facebook, the ad contained a line (in bold) saying “Germans abroad” and the additional encouragement “Please participate in our scientific survey”. However, the latter phrase was only partially visible, cutting off after the first syllable of the word scientific (German: *wissenschaftlich*), but showing enough of the phrase to serve as additional encouragement to participate. Advertisements on Instagram suppress these last two text elements.

Our initial campaign design used the same 13 advertisements in all ad sets. Hence, in the beginning, our sampling strategy used a campaign consisting of 52 ad sets, featuring a total of 676 advertisements targeting Germans in 194 non-European countries and territories.

Since several of the above-listed design elements have the potential to impact both the performance of advertisements and the composition of the sample recruited through them, an effort was made to register in the dataset as much information as possible regarding the specific ads survey participants reacted to. To this end, we specified six URL parameters (cf. Facebook 2021d) that were assigned to each

observation indicating which (1) sampling region, (2) age cohort, and (3) gender the ad set targeted, on which (4) SNS and in which (5) placement the ad was displayed, as well as the (6) picture it showed.

3.4 Budgeting and running the campaign

Facebook Inc. does not provide fixed tariffs or a set price list for running advertisement campaigns. Instead, the costs are determined during the campaign's lifetime through an automated bidding process, mainly depending on the defined target group and the interest of other advertisers in this group (cf. Facebook 2021a). Therefore, advertisers can beforehand only define the budget they want to use and the way it should be spent. For the latter, there are two basic options, namely *lifetime budget*, and *daily budget*. The former means that a budget is defined for the whole lifetime of an ad set, permitting the algorithm to optimize its use during this time and across various days. This means that the advertisements in a given ad set could be delivered more often during periods for which the algorithm determines a higher chance of reaching the intended goal (e.g., accumulating link clicks). This might result in a very uneven use of the budget across days. In contrast, with the *daily budget* option, a specific amount can be set that the algorithm will try to spend each day allowing for a more even spread across days (cf. Facebook 2021b).

Furthermore, advertisers need to specify whether they want to pay for advertisements per click or per impression. While the former means that Facebook Inc. charges for every link-click that users perform on the advertisement, the latter results in costs occurring every time the advertisements are delivered to a user's Facebook or Instagram account. Since this, however, does not mean that said user would even see or read the advertisement, the cost per click (CPC) system is generally considered preferable from a cost-benefit point of view in survey research (Kühne and Zindel 2020; Pötzschke and Braun 2017). While this option was available to all advertisers in the past, Facebook Inc., unfortunately, changed its policy and currently offers this option only for advertisement accounts that have already spent a certain amount of money. When the project was planned, a note in the Facebook Ad Manager stated that the option would become available once the new account we used for this survey spent at least 8 Euro. Therefore, we conducted a pretest several days before setting up the actual survey hoping that the option would become available. However, even though 9 Euro were spent during the pretest, the CPC option remained unavailable in the short term. Consequently, we decided to run the campaign using the cost per impression setting.

The project had a total advertisement budget of 2,391 Euro at its disposal to sample respondents for the survey. Using the cost per impression setting, we initially planned for an advertisement campaign field period of 10 days and distributed the budget across all 52 ad sets using the *lifetime* budget option. A budget of 50 Euro was allocated to most ad sets. The exceptions being those targeting migrants in sampling regions Oceania 1 and 2, North America 2 (using 37.5 Euro each), and North America 1 (using 32.25 Euro, see Table 1). We took this decision for budgetary reasons considering that established approaches used in previous studies were comparatively successful in sampling German emigrants

overseas in those regions. Based on this setup, the budget available for advertisements on a single day totaled 239 Euro.

The advertisement campaign was launched on August 13, 2020, at 0h00 (CEST). While the advertisements immediately resulted in survey participation, the overall performance of the campaign remained far below our expectations. Most striking was the fact that ads were only delivered in the first 6-10 hours (CEST) of a day even though only a fraction of the available budget had been spent. More precisely, the costs amounted on average only to 30.30 Euro at each of the first three days of the campaign. Even considering that fluctuations were to be expected using the lifetime option (see above), we judged this to be an underperformance and took different measures intended to enhance the daily outcome. This included the reduction of used pictures from 13 to 7, as available sources suggested that a high number of ads in an ad set combined with a small budget might result in the corresponding ads not being delivered (OAS, 2016). However, since these and other measures did not bear fruit, we decided to stop the initial campaign and start a new one using daily budgets on August 21, 0h00 (CEST). For this new campaign, the remaining funds per ad set were used to calculate the available daily budget for each ad set. Campaign 2 used the same 52 ad sets as campaign 1 but continued using the reduced number of pictures for the advertisements. This second campaign ended on August 30, 1h59 p.m. (CEST), as the budget of all ad sets was near depletion. Taken together, the two advertisement campaigns ran 18 days and resulted in total costs of 2,222.90 Euro.

3.5 Questionnaire and fielding

GEOOS used a rather comprehensive questionnaire. In line with the overall aims of the study it focused on demographic aspects and information related to respondents' migration and experience abroad. Additionally, the project collected information on social media use, integration in the countries of residence, and social cross-border relations. Furthermore, the instrument included a small number of questions investigating impacts of the COVID-19 pandemic.

The survey data collection started on August 13 and concluded on September 12, 2020. Hence, in contrast to the advertisement campaign, the field period of the survey was exactly one month. This approach was chosen to allow for the participation of respondents who might, for example, have received the survey link from other participants or learned about the project through the like of contacts on Facebook or Instagram.

Table 1 provides an overview of the campaign structure, costs that occurred per ad set and campaign, as well as the resulting unique clicks.

Table 1: Ad sets and campaign structure

Sampling region	Variables defining ad sets		Campaign 1 (August 13-20, 8 days)		Campaign 2 (August 21-30, 10 days) ¹	
	Gender	Age range	Unique clicks	Amount spent (EUR)	Unique clicks	Amount spent (EUR)
Africa 1 - Biggest 5	female	18-28	18	0.92	221	41.79
Africa 1 - Biggest 5	female	29-65+	41	1.04	465	42.53
Africa 1 - Biggest 5	male	18-28	35	1.29	598	42.52
Africa 1 - Biggest 5	male	29-65+	71	1.33	867	43.15
Africa 2 - Southern Africa	female	18-33	1	0.07	122	43.09
Africa 2 - Southern Africa	female	34-65+	8	0.18	367	43.40
Africa 2 - Southern Africa	male	18-33	2	0.09	110	43.18
Africa 2 - Southern Africa	male	34-65+	4	0.20	312	43.26
Africa 3 - Remaining countries	female	18-28	4	0.35	77	42.78
Africa 3 - Remaining countries	female	29-65+	7	0.59	144	43.35
Africa 3 - Remaining countries	male	18-28	12	0.58	176	42.98
Africa 3 - Remaining countries	male	29-65+	18	0.68	315	43.03
Asia 1 - Biggest 6	female	18-29	6	0.59	417	46.21
Asia 1 - Biggest 6	female	30-65+	28	0.54	1113	46.21
Asia 1 - Biggest 6	male	18-29	14	0.95	612	47.29
Asia 1 - Biggest 6	male	30-65+	63	0.93	1168	46.88
Asia 2 - W + Cen. Asia	female	18-30	10	0.32	268	42.81
Asia 2 - W + Cen. Asia	female	31-65+	14	0.37	454	42.67
Asia 2 - W + Cen. Asia	male	18-30	9	0.42	450	42.38
Asia 2 - W + Cen. Asia	male	31-65+	20	0.44	627	42.52
Asia 3 - S. + E. + SE Asia	female	18-30	16	1.38	214	48.30
Asia 3 - S. + E. + SE Asia	female	31-65+	28	1.34	307	48.29
Asia 3 - S. + E. + SE Asia ²	male	18-30	26	1.28	408	48.85
Asia 3 - S. + E. + SE Asia	male	31-65+	51	2.30	619	54.96
Middle + S. America 1 - Biggest 5	female	18-30	36	3.67	284	43.03
Middle + S. America 1 - Biggest 5	female	31-65+	154	3.68	559	42.04
Middle + S. America 1 - Biggest 5	male	18-30	36	2.57	413	43.36
Middle + S. America 1 - Biggest 5	male	31-65+	120	3.35	734	42.81
Middle + S. America 2 - Cent. A. + Caribbean	female	18-31	18	3.00	90	44.23
Middle + S. America 2 - Cent. A. + Caribbean	female	32-65+	94	3.39	270	43.71
Middle + S. America 2 - Cent. A. + Caribbean	male	18-31	26	3.11	131	44.29
Middle + S. America 2 - Cent. A. + Caribbean	male	32-65+	92	3.29	258	44.29
Middle + S. America 3 - South America	female	18-31	26	3.48	82	44.21
Middle + S. America 3 - South America	female	32-65+	106	3.69	254	43.92
Middle + S. America 3 - South America	male	18-31	33	3.46	145	44.33
Middle + S. America 3 - South America	male	32-65+	108	3.80	279	44.02

Sampling region	Variables defining ad sets		Campaign 1 (August 13-20, 8 days)		Campaign 2 (August 21-30, 10 days) ¹	
	Gender	Age range	Unique clicks	Amount spent (EUR)	Unique clicks	Amount spent (EUR)
North America 1 - Biggest (USA)	female	18-42	27	2.20	329	30.65
North America 1 - Biggest (USA)	female	43-65+	55	2.61	594	29.74
North America 1 - Biggest (USA)	male	18-42	17	2.03	219	30.96
North America 1 - Biggest (USA)	male	43-65+	34	2.19	349	30.45
North America 2 - Remaining countries ²	female	18-42	26	2.59	285	33.25
North America 2 - Remaining countries	female	43-65+	58	2.64	458	32.91
North America 2 - Remaining countries	male	18-42	31	2.67	253	33.08
North America 2 - Remaining countries	male	43-65+	52	2.82	387	33.03
Oceania 1 - Biggest (Australia)	female	18-38	15	1.62	237	35.87
Oceania 1 - Biggest (Australia)	female	39-65+	25	1.82	356	35.61
Oceania 1 - Biggest (Australia)	male	18-38	8	1.60	165	35.79
Oceania 1 - Biggest (Australia)	male	39-65+	25	1.69	240	35.71
Oceania 2 - Remaining countries ¹	female	18-33	23	1.37	162	34.87
Oceania 2 - Remaining countries ¹	female	34-65+	16	1.58	231	34.82
Oceania 2 - Remaining countries ¹	male	18-33	12	1.38	91	34.32
Oceania 2 - Remaining countries ¹	male	34-65+	16	1.52	139	34.17

Note: All statistics provided by the Facebook Ad Manager.

¹ The delivery of the four ad sets in region Oceania 2 was already halted on August 27, as the corresponding budgets were nearing depletion at this date.

² Two ad sets were started on a lifetime budget in campaign 2 by mistake. They were switched to daily budget on August 24.

4 Recruitment results

A total of 3,895 participants completed our project's survey. In the following, we will first describe the results of the recruitment process and then turn to the survey itself. Depending on the respective recruitment stage, we might refer to samples of different sizes.

4.1 Performance of advertisements

The second column in Table 2 presents the sampling frame of our survey as provided by FAM before the start of the ad campaign. Of the approx. 720,000 Facebook and Instagram users who complied with our target criteria, nearly 30 percent lived in target region Asia 1, constituted by India, Indonesia, the Philippines, Thailand, Turkey, and the Russian Federation (see Appendix 1). With 20 percent of the overall sampling frame, users in the United States formed the biggest sub-group in a single country. Hence, taken together, these two target regions alone accounted for half of the reach estimated by the FAM before the start of our campaign. Table 2 provides key indicators of the advertisement campaigns across the different target regions.

Table 2: Facebook performance indicators of ads across sampling regions (cumulative values of all ad sets by sampling region)

	Estimated reach	Users reached by ads (reach)	Impressions	Unique clicks	Amount spent	Costs per unique click
Africa 1 - Biggest 5	78,100	54,208	318,313	2,316	174.57 €	0.08 €
Africa 2 - Southern Africa	17,700	10,736	72,527	926	173.47 €	0.19 €
Africa 3 - Remaining countries	30,300	18,345	102,801	753	174.34 €	0.23 €
Asia 1 - Biggest 6	213,000	87,620	360,779	3,421	189.60 €	0.06 €
Asia 2 - W + Cen. Asia	42,700	28,351	183,654	1,852	171.93 €	0.09 €
Asia 3 - S. + E. + SE Asia	57,000	35,218	192,861	1,669	206.70 €	0.12 €
Middle + S. America 1 - Biggest 5	61,900	50,846	256,344	2,336	184.51 €	0.08 €
Middle + S. America 2 - Cent. A. + Caribbean	10,400	12,790	106,774	979	189.31 €	0.19 €
Middle + S. America 3 - South America	10,600	12,617	107,384	1,033	190.91 €	0.18 €
North America 1 - Biggest (USA)	146,000	15,423	23,358	1,624	130.83 €	0.08 €
North America 2 - Remaining countries	24,200	12,466	26,147	1,550	142.99 €	0.09 €
Oceania 1 - Biggest (Australia)	21,100	10,203	24,127	1,071	149.71 €	0.14 €
Oceania 2 - Remaining countries	6,300	5,781	30,148	690	144.03 €	0.21 €
Total	719,300	354,604	1,805,217	20,220	2,222.90 €	0.11 €

Note: The shown statistics were provided by the Facebook Ad Manager.

Our data show that there are clear differences in the targeting costs for our population of interest in different world regions. At first sight, it might seem that there would be a direct link between the costs per click and the number of eligible users (“Estimated reach”) in a given region. Costs per click were, for example, lowest in Asia 1 and North America 1 (USA). However, the table also shows that a single click was, on average, most expensive in region Africa 3, even though a lower reach had been estimated for six of the project's other targeting regions beforehand. Hence, the naïve assumption of a linear relation between the size of the target group in a given region (estimated reach) and the cost per click does not hold. The reason for this lies, at least in part, with the fact that the project had to use the payment per impression option instead of directly paying for each click (see section 3.4). Consequently, the costs per click depend, first, on the costs of advertisement impressions in a given region and second on the average number of impressions that were needed to generate a click on the survey link. As Table 2 shows, the latter value varied greatly between regions. In other words, the probability of users clicking on our ads was not constant across the targeted regions. This means that in a region for which the FAM indicated a lower potential reach, clicks could still be cheaper than in a region hosting more individuals of the target group if the product of costs per impressions and number of impressions needed to generate a click were lower in the former than in the latter.

The question as to whether the recruitment procedure targets the intended geographic region is of utmost importance in cross-national migration research. As mentioned above, we employed a URL parameter

that clearly identified the ad set, and hence the geographic region, in which the link was included that each respondent used to reach the survey. Additionally, the survey also contained a question asking respondents for their country of residence. Consequently, we were able to compare the (self-reported) country of residence and the region, with the information provided by Facebook about a user's country of residence and region, respectively. A large overlap would indicate that the targeting mechanism worked correctly. Naturally, this comparison can only be applied to the 3,632 surveys that were completed by participants who reached the survey through one of the advertisement links (see Table 3, below). Indeed, our analysis shows a match between the geographic URL parameter and the answer to the survey question for the country of residence in 92.7 percent of the respective cases. Detailed analysis reveals a complete match for one targeting region (Oceania 1, i.e., Australia). In 11 of the remaining 12 targeting regions, the information matches for most respondents, showing discrepancies only in one to nine observations each, corresponding to 0.4 to 13.6 percent of the respective subsamples. However, there is one subsample, namely North America 2 (targeting Bermuda, Canada, Greenland, and Saint Pierre and Miquelon, see Appendix 1), for which our data show a mismatch for 200 out of 597 respondents (33.5 percent). Given the mentioned high matching rate in the remaining regions, this finding is puzzling at first. Naturally, a first step is to double-check whether the geographic targeting used in the corresponding ad sets had been specified correctly, which did not reveal any errors in this case. Fortunately, during the survey, respondents were also asked how they first learned of our project. While 83 percent of all participants, for whom both regional information match, reported that they first became aware of the survey through an advertisement on Facebook or Instagram, the same holds true for only 33 percent of cases where our data show a mismatch. Furthermore, in the case of the ad sets targeting North America 2, only 24 percent of the respondent for whom the regional information did not match learned about the project through an advertisement, while a majority received word through other channels, such as Facebook and Instagram messages of friends (26 percent), "likes" by their contacts (15 percent), or Facebook groups (13 percent). These findings suggest that the reported mismatch is likely due to situations in which respondents living outside the targeted region reached the survey after having received the link assigned to North America 2 from third parties. For example, it is conceivable that Facebook users living in New Zealand visited a Facebook group for German emigrants in which a user who lived in Canada shared the link they received from an ad set targeting this region. Consequently, and considering all available data, we conclude that the geographic targeting appeared to have worked within a reasonable margin of error.

4.2 Completed surveys and sampling costs

The sample sizes shown in Table 3 underline the success of the used sampling approach. A total of 3,895 individuals completed the survey; of those, 98 percent were either born in Germany or held German citizenship but did not live in Germany; we consider these as "emigrants in a broadly defined sense". Furthermore, 89 percent of the sample even fulfilled all three criteria at the same time. That is, they were born in Germany, held German citizenship, and did not live in Germany at the time of the survey

(“emigrants narrowly defined”). The same applies if we focus only on those 3,632 respondents who were directly or indirectly recruited through the advertisements (see above). Hence, concurring with the findings of Pöttschke and Braun (2017), these results underline the precision of a targeting approach using the combination of users’ country of residence and the variable “lived in [Germany]” (formerly “expat Germany”).

Table 3: Completed surveys

	Completed surveys (total)	Emigrants broadly defined	Emigrants narrowly defined
Advertisements	3,632	3,558	3,222
Facebook page	74	74	67
Snowball	189	184	173
Total	3,895	3,816	3,462

The values in Table 3 also hint at a future line of research worth exploring: While we did not offer any incentives, as it is usually done when employing snowball or respondent-driven sampling (Heckathorn 1997), we collected 184 additional completed surveys from German emigrants through our snowball add-on (five percent of the overall sample). This is encouraging because it might constitute a possibility to extend the sample to respondents who are themselves not users of the employed SNS, thereby allowing to counterbalance biases associated with this method.

Putting the total of 3,816 surveys that were completed by eligible respondents in relation to the advertisement budget of 2,222.90 Euro means that the sampling costs per valid observation amounted to an average of 0.58 Euro. If we only considered those respondents who used an advertisement link to reach the survey, the costs would still be at a modest 0.62 Euro. Given these low costs, the method can only be described as extremely costs efficient.

While there are differences regarding the costs per completed questionnaire across sampling regions, we are not able to provide exact figures for them. This is due to the fact that the advertisement links employed in some sampling regions have also been used by participants living in other parts of the world. Hence, especially if the budget used in sampling region ‘North America 2’ would simply be divided by the number of completed surveys for which the corresponding link has been registered in the dataset, this would lead to a serious underestimation of the costs. In order to provide, nevertheless, some information regarding the cost differences, we take a more cautious approach considering only completed surveys of those respondents in the category ‘emigrants, broadly defined’ for whom the geographic information registered through the URL parameter match their answer to our survey question ($n = 3,308$). However, it should be kept in mind that, by doing so, we are likely to slightly overestimate the corresponding costs. In 10 of 13 sampling regions, these potentially too high estimates are still under

one Euro per completed questionnaire, ranging from 36 to 91 Eurocent (see Appendix 2). The sampling costs per completed questionnaire were highest in the sampling regions ‘Middle and South America 2’ (consisting of countries in Central America and the Caribbean, 1.11 Euro), ‘Africa 2’ (mainly Northern African countries, 1.28 Euro), and ‘Africa 3’ (mainly countries in Western Africa and the adjoining northern part of the sub-Saharan region, 3.11 Euro). It is noteworthy that four of the five regions in which the sampling costs were the lowest consist mainly of predominantly English-speaking countries in Northern America and Oceania. Furthermore, in the one African sampling region in which the costs per completed questionnaire were decisively lower than in the other two (i.e., ‘Africa 2’, 0.73 Euro), 56 percent of the respondents resided in South Africa and another 18 percent in Namibia. The latter not only being a country in which English is one of the most frequently spoken lingua franca but also a former German colony.

5 Sample description

5.1 Socio-demographic and geographic composition of the sample

In the following, we will briefly cover the socio-demographic composition as well as the geographic distribution of the sample. The corresponding information stems from our survey data, *not* from participants’ user profiles or advertisement campaign related statistics. The socio-demographic composition of the overall sample with regard to the avenue of recruitment (via advertisements, Facebook page, or snowball sampling) reveals interesting details (see Appendix 3). Overall, more females (57 percent) than males (43 percent) participated in the survey. With regard to advertisement-based recruitment, this is true for both SNS; though, respondents recruited via Instagram were much more likely to be female (69 percent) than male (31 percent). The snowball sampling, compared to the advertisement-based recruitment, produced an even less diverse sample with respect to sex, i.e., here, about three-quarters of the respondents are female. A second socio-demographic variable is age in years. While the mean age of the total sample is about 47 years, we see considerable differences between the Facebook- (48 years) and Instagram-recruited (38 years) subsamples again. Furthermore, and this is no surprise (AAPOR et al. 2010; Antoun 2015), the overall level of educational attainment of our sample is ‘high’⁵ (71 percent). The respondents from the Instagram sample are slightly higher educated, i.e., about 75 percent of the Instagram sample report high educational attainment vs. 69 percent in the Facebook sample. The share of respondents that report a high level of educational attainment is particularly large for the snowball sample (81 percent). Employment status is another important socio-demographic characteristic. Most of our participants report being employed (69 percent). There are only minor differences between employed and unemployed respondents among the subgroups. However, respondents that came through Instagram are overall more likely to be in education, which corresponds

⁵ Educational attainment has been categorized as follows: “Low” (currently attending school, no degree, Polytechnic Secondary School, G 8/9, Secondary General School-leaving Certificate), “Intermediate” (Polytechnic Secondary School, G 10, Intermediate School-leaving Certificate), “High” (Applied University Entrance Qualification (FH), University Entrance Qualification).

with the lower mean age mentioned above. Finally, we have a look at the duration of stay in the respective country of residence. Contrary to participants of most of the previous surveys on German emigrants (see Section 2), the majority of respondents in this project had already spent considerable time in their current country of residence. More specifically, about 56 percent stated that they already lived there between six and 10 years. Nearly a quarter had resided for 21 years or longer in the country from which they participated in the survey. Respondents that participated via Instagram, as well as those recruited through the snowball add-on, report slightly lower durations than the other group. So, all in all, respondents that were recruited via Instagram as well as via snowball sampling are more likely to be female, younger, better educated, and report a lower duration of stay in their country of residence than those recruited through Facebook.

The geographic distribution of the sample is shown in Figure 3 and illustrates that the use of advertisement on SNS resulted in a geographically highly diverse sample. All in all, GEOOS participants lived in a total of 148 countries and territories around the globe. Somewhat similar to findings reported in previous studies on German emigrants, the largest sub-groups resided in predominantly Anglo-phone countries; more specifically, in Canada ($n = 416$), New Zealand ($n = 359$), Australia ($n = 344$), and the United States of America ($n = 323$, excluding overseas territories). However, what sets our study apart from previous projects is that taken together, participants in these countries only constitute 38 percent of the overall sample instead of constituting its absolute majority. Indeed, nearly a quarter of GEOOS participants ($n = 866$) lived in Middle and South America, 881 resided in Asian countries, and 477 in Africa, two-thirds of them in Sub-Saharan countries.

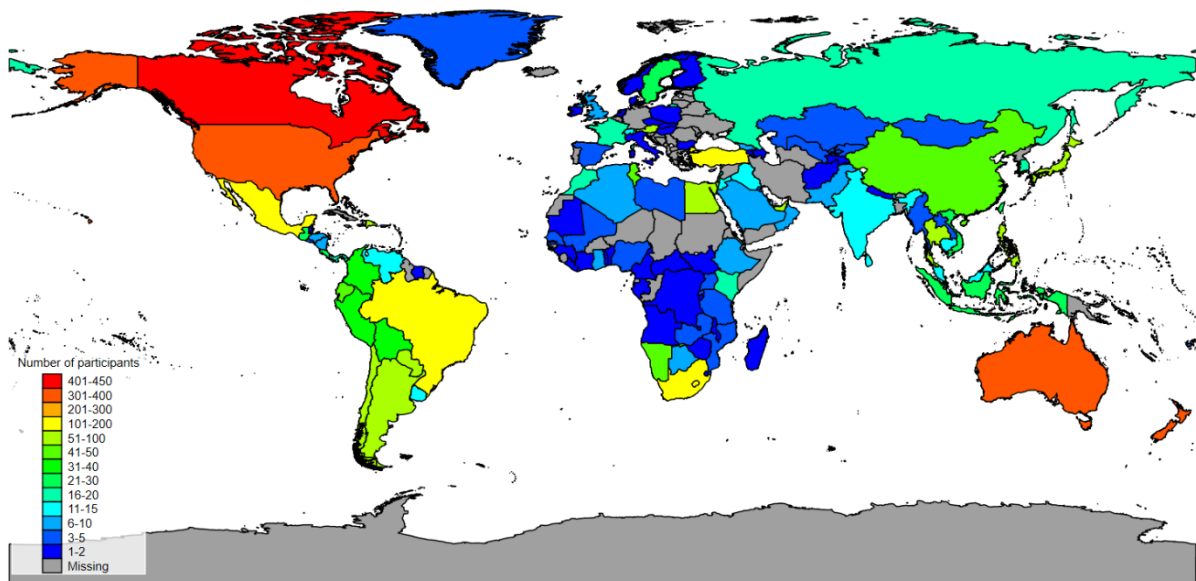


Figure 3: Distribution of GEOOS participants by the size of sub-samples in specific countries (migrants broadly defined)

Furthermore, the data show that the used sampling approach allowed us to survey ten or more members of the target population in 53 countries and territories, and more than 50 in 18 of them. Even though these are comparatively small numbers if seen on their own, considering our small sampling budget of only slightly more than 2,200 Euro we find these results remarkable and encouraging regarding future research.

5.2 Comparison of the geographic distribution of the GERPS and GEOOS sample

To further evaluate the geographic distribution of our sample, we compare it to the distribution of emigrant respondents in the *German Emigration and Remigration Panel Study* (GERPS), which achieved the most comprehensive sample of German emigrants to-date (see Section 2.1 for a detailed description of the GERPS project). For this comparison, we employed this project's public use dataset (Erlinghagen and Schneider 2020), which does not indicate the specific countries in which respondents lived but assigns observations to different country groups. In Table 4, we present the number of observations in these groups alongside those reached in GEOOS.

Table 4: Geographic distribution of GERPS and GEOOS samples

	GERPS		GEOOS comparison group		GEOOS total	
	Freq.	Percent	Freq.	Percent	Freq.	Percent
Europe - <i>not explicitly targeted in GEOOS</i>	3,344	73.9	60	1.8	132	3.5
Near and Middle East (incl. TUR)	111	2.5	336	9.9	371	9.7
Asia (incl. RUS, UKR, BLR)	280	6.2	471	13.9	510	13.4
Africa	77	1.7	429	12.6	477	12.5
North America (CAN + USA)	456	10.1	637	18.8	739	19.4
North America (other) <i>- not identifiable in GERPS public use file</i>	n.a.	n.a.	3	0.1	5	0.1
Latin America (incl. MEX)	129	2.9	810	23.9	866	22.7
Oceania	118	2.6	648	19.1	714	18.7
Unknown	10	0.2	2	0.1	2	0.1
Total	4,525	100	3,396	100	3,816	100

Note: GERPS values are based on own calculations using Erlinghagen and Schneider (2020). In accordance with the definition of the GERPS target population, only individuals holding German nationality are included. While GERPS differentiates between several sub-groups of countries within Europe, these are combined in this table as GEOOS did not directly target emigrants in European countries. North American countries other than Canada, Mexico, and the USA are shown separately as it is unclear whether they are included in the GERPS category 'North America'.

To allow for a direct comparison of the geographic diversity of samples achieved with the recruitment methods employed in GEOOS and GERPS, the columns labeled 'GEOOS comparison group' in Table 4 consider only surveys completed by German nationals⁶ living abroad who reached our survey by clicking on an advertisement link. The table shows that, despite the limited resources and short fielding period of our project, more German emigrants participated in GEOOS from all non-European regions than in GERPS. This even holds true for North America, i.e., the overseas region in which the latter study was able to recruit its highest number of respondents. Furthermore, our sampling method yielded samples size of more than 300 individuals in all directly targeted regions. This also applies to those regions where GERPS reached only few emigrants, namely Africa (n = 77) and the Near and Middle East (n = 111). In fact, the size of the GEOOS comparison sample is more than six times that of GERPS in Latin America and while it is more than five times bigger in Africa and Oceania.

Compared to the sampling method applied in GERPS, the sampling frame used in our study had the big advantage that it enabled us to recruit emigrants who had already spent several years abroad instead of being limited to those who deregistered their German residence during less than two years prior to the survey. However, even if we limit the comparison sample further to those respondents who lived two years or less in their current country of residence, the geographic distribution of this sub-group (n = 723) is comparatively balanced across the non-European regions listed in Table 4⁷, with the sample sizes in two regions still surpassing GERPS, namely in Africa (n = 87) and Oceania (n = 177).

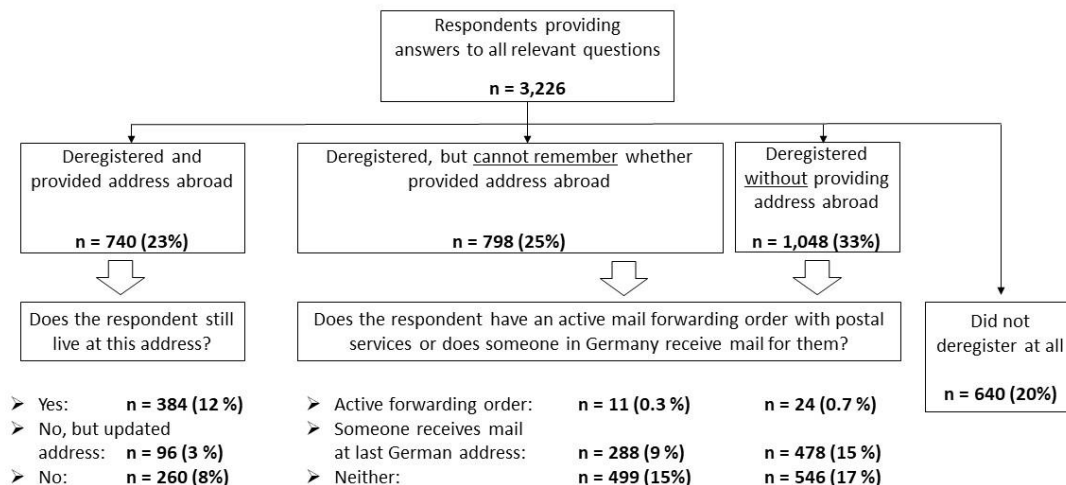
⁶ Nationality, but not place of birth, was one of the parameters used to define the target population in GERPS (see section 2.1).

⁷ Near and Middle East 10 percent, Asia 19 percent, Africa 12 percent, North America 17 percent, Latin America 16 percent, and Oceania 25 percent.

5.3 Reachability of respondents with other methods

We also collected information as to whether respondents could have been reached by established sampling methods for single or cross-national surveys, more specifically via name-based telephone sampling or by using Germany's population registers as a sampling frame. As telephone sampling is not a feasible option for a global survey of emigrants, we will focus on the question as to whether participants met the basic preconditions to be included in a register sample. We, furthermore, limit this discussion to those 3,226 emigrants (i.e., 85 percent of the overall sample) who provided valid answers to all questions related to this issue. More specifically, we asked respondents whether they deregistered with German authorities when moving abroad and provided the registrar's office with their new address. If not, we inquired whether they could be reached via their last known German address. Figure 4 provides an overview of the corresponding distribution of answers.

Figure 4: Deregistration from German population registers and reachability via last German address



As Figure 4 shows, while 23 percent of the participants who provided information on this aspect did deregister with German authorities and provided a new address abroad, only 15 percent still lived at these addresses or had updated their information. Consequently, only these 480 respondents could have been reached by postal invitations send to foreign addresses included in German population registers. This also means that more than one-third of those respondents whose addresses would have been included in the sampling frame had moved to other locations in the meantime and could not have been recruited by this approach.

Following a register-based approach, a possibility to reach those emigrants who deregistered without providing a foreign address would consist in sending a survey invitation to their last known German address (see Section 2.1). Indeed, 33 percent of the respondents included in this analysis stated that they had not provided a new address when deregistering, with another 25 percent not being sure about it. Taken together, only 35 of these respondents had an active forwarding order with postal services in

place (1 percent of the respondents). However, 766 respondents (24 percent of the group overall) said they had arranged for someone to receive mail in their name at their last German address. However, Figure 4 also indicates that 17 percent of respondents could neither have been reached abroad (as they did not provide a corresponding address) nor by sending survey invitations to their last known address in Germany. The same is likely for another 15 percent of the respondents who were not sure whether they had registered a foreign address. Hence, taken together, these individuals constitute nearly one-third of all participants who answered the questions investigating these issues.

Finally, one-fifth ($n = 640$) of the GEOOS respondents for whom we have the needed information did not deregister at all and, hence, would have been missed by a register-based sampling frame.

In a nutshell, this means that of the mentioned 3,226 respondents who provided the required information, 15 percent could be reached at a foreign address registered with German authorities, and 25 percent might be reached by mail sent to their last known German address. However, 45 percent of those respondents in our sample who provided information on their deregistration could certainly not be reached by a survey that used register-based sampling; the same is very likely for an additional 15 percent of these participants. Hence, our results indicate a considerable coverage issue of the register-based approach.

6 Conclusion

In this paper, we presented methodological insights stemming from the *German Emigrants Overseas Online Survey* (GEOOS), which, to the best of our knowledge, constitutes the first survey that used advertisements on Facebook and Instagram to recruit respondents on a (nearly) global level. More specifically, we targeted German emigrants in countries beyond Europe. Our study addresses a significant research gap, as emigrants from most countries constitute hard-to-survey populations due to their high level of global geographic dispersion, which in consequence leads to a lack of suitable and easily accessible sampling frames.

GEOOS employed a geographically clustered strategy to sample members of its target population on a global level, excluding Europe. We decided to focus on non-European regions, as previously tested sampling approaches for German emigrants achieved their best results in European countries while being much less effective overseas, particularly in the Global South. Consequently, both methods to target German emigrants and information on this group particularly lack in the non-European context. However, based on the satisfactory results delivered by our sampling approach in other world regions, we have no reason to assume that it would not produce good results in Europe, too. In terms of the geographic diversity of the resulting sample, the SNS advertisement-based approach used in GEOOS outperformed previously used approaches to recruit emigrants in general and German emigrants in particular. Regarding the targeted non-European regions, this study's sample is not only distributed across more countries but also larger than those achieved by most previous surveys. This holds not only

true for the total sample size but also for the individual sub-samples in different regions. Notably, using this method, we were also able to survey larger groups of respondents in regions in which previous studies had particular problems to recruit respondents and consequently showed undercoverage of the target group, namely in Africa and the Middle East.

Within less than three weeks, and using only a comparatively small budget of just over 2,223 Euro, we were able to collect 3,816 completed surveys from individuals who were either born in Germany or held German citizenship (or both) but lived in another country. Furthermore, even when applying a narrower definition, counting only respondents who were born in Germany and held German citizenship, the total sample amounts to 3,462 participants. This corresponds to average sampling costs of 0.58 Euro and 0.62 Euro, respectively, per completed survey. On a more detailed level, our results show that costs varied between targeting regions, being, for example, lower in North America, Oceania, and parts of Asia than in Africa and Central America. The findings regarding the distribution of these differences can, however, not easily be projected onto other (emigrant) target groups as advertisement costs in Facebook and Instagram are highly dependent on the composition of the target group. On a general note, the presented results show, however, that no direct linear relation can be assumed between the size of the target group in a specific region and associated sampling costs, at least when using the payment per impression option.

The presented data also indicates that a considerable part of our sample could not have been reached by a population register-based sampling approach, which is the most promising established method. As mentioned, many GEOOS respondents did indeed not deregister their German residence with the authorities or did at least not provide a new address abroad and could also not have been reached by postal mail sent to their last known German address. Furthermore, we were able to recruit respondents who had already spent several years abroad which constitutes another group for which other methods usually show undercoverage.

Our research also highlights that URL parameters can (and should) be used to review and judge the precision of the advertisement targeting. While we generally found that respondents recruited through the advertisements lived in the targeted regions, the combination of URL parameters and specific survey questions allowed us to better understand or at least theorize what might have happened in cases in which this was not true.

Notwithstanding its advantages, SNS sampling has obvious limitations. Most important in this regard are its non-probability character and the associated selection biases. Advertisements on networks such as Facebook and Instagram can obviously only reach, in a direct way, individuals who use these SNS (Sen et al. 2021). On top of this, there might be a self-selection bias, meaning that those target group users who react to such advertisements and eventually participate in the survey might differ systematically from those who do not. In this regard, population register-based sampling, as employed by the German Emigration and Remigration Panel Study, has the advantage that it provides scholars

with additional information on the individuals included in the sampling frame (such as age and sex). This information then allows for a thorough analysis of possible biases in the achieved sample. More importantly, the German population registers can be used to draw probability samples, even though not of emigrants in general but “only” of individuals who deregistered their German place of residence stating emigration as reason within a specific timeframe.

We reckon that the approach presented in this paper has a high potential to enable large cross-national migration and particularly emigration surveys. Sampling via Facebook and Instagram advertisements can be realized with a fraction of the resources needed for the implementation of other approaches both in terms of time and money. Its accessibility to the broader scientific community is, furthermore, strengthened by the fact that it builds on Facebook Inc.’s basic business model, meaning that no formal access to and cooperation with this company is needed. Since no method is available that would allow drawing probability samples of emigrants from all countries worldwide without any systematic limitations, the approach presented here can be seen as a useful addition to migration scholar’s methodological toolbox. As a cost-effective stand-alone non-probability method, it can be employed to research many substantive questions on a global level. Furthermore, it could be used in concert with other approaches, such as population register-based or snowball respectively respondent-driven sampling combining the strengths of different methods and thereby counterbalancing their weaknesses.

Regarding the latter point, GEOOS already demonstrated that respondents recruited through advertisements could serve as seeds in a snowball sampling approach that might be used to reach respondents who are not users of SNS. Our study only included a small pilot element to test the feasibility of this idea, and we see this as a promising avenue for future research.

References

- AAPOR, Prepared for the AAPOR Executive Council by a Task Force operating under the auspices of the AAPOR Standards Committee, with members including: R. Baker, S. J. Blumberg, J. M. Brick, M. P. Couper, M. Courtright, J. M. Dennis, D. Dillman, M. R. Frankel, P. Garland, R. M. Groves, C. Kennedy, J. Krosnick, P. J. Lavrakas, S. Lee, M. Link, L. Piekarski, K. Rao, R. K. Thomas, and D. Zahs. 2010. "Research Synthesis: AAPOR Report on Online Panels." *Public Opinion Quarterly* 74(4):711–81. doi: 10.1093/poq/nfq048.
- Antoun, Christopher. 2015. "Who Are the Internet Users, Mobile Internet Users, and Mobile-Mostly Internet Users?: Demographic Differences across Internet-Use Subgroups in the U.S." Pp. 99–117 in *Mobile Research Methods: Opportunities and challenges of mobile research methodologies*, edited by D. Toninelli, R. Pinter, and P. de Pedraza. London: Ubiquity Press.
- Arcia, Adriana. 2014. "Facebook Advertisements for Inexpensive Participant Recruitment Among Women in Early Pregnancy." *Health Education & Behavior* 41(3):237–41. doi: 10.1177/1090198113504414.
- ARD/ZDF Forschungskommission. 2021. "ARD/ZDF-Onlinestudie - Gerätenutzung 2020." Retrieved February 23, 2021 (<https://www.ard-zdf-onlinestudie.de/geraetenutzung/>).
- Baltar, Fabiola, and Ignasi Brunet. 2012. "Social Research 2.0. Virtual Snowball Sampling Method Using Facebook." *Internet Research* 22(1):57–74. doi: 10.1108/10662241211199960.
- Beisch, Natalie, and Carmen Schäfer. 2020. "Internetnutzung Mit Großer Dynamik: Medien, Kommunikation, Social Media - Ergebnisse Der ARD/ZDF-Onlinestudie 2020." *Media Perspektiven* (9):462–81.
- Blaskó, Zsuzsa. 2015. *Studying Emigration by Extending a Large-Scale Household Survey. Methodology, Evaluation and Descriptive Findings*. 21. Budapest: Hungarian Demographic Research Institute.
- Brickman Bhutta, Christine. 2012. "Not by the Book. Facebook as a Sampling Frame." *Sociological Methods & Research* 41(1):57–88. doi: 10.1177/0049124112440795.
- Careja, Romana, and Hans-Jürgen Andreß. 2018. "In Search of a Frame. Challenges and Opportunities for Sampling Immigrant Minorities." *Comparative Migration Studies* 6(37). doi: 10.1186/s40878-018-0103-5.
- Careja, Romana, and Pieter Bevelander. 2018. "Using Population Registers for Migration and Integration Research. Examples from Denmark and Sweden." *Comparative Migration Studies* 6(1):19. doi: 10.1186/s40878-018-0076-4.
- Carlini, Beatriz H., Luciana Safioti, Tessa C. Rue, and Lyndsay Miles. 2015. "Using Internet to Recruit Immigrants with Language and Culture Barriers for Tobacco and Alcohol Use Screening. A Study Among Brazilians." *Journal of Immigrant and Minority Health* 17(2):553–60. doi: 10.1007/s10903-013-9934-1.
- Cornesse, Carina, Annelies G. Blom, David Dutwin, Jon A. Krosnick, Edith D. De Leeuw, Stéphane Legleye, Josh Pasek, Darren Pennay, Benjamin Phillips, Joseph W. Sakshaug, Bella Struminskaya, and Alexander Wenz. 2020. "A Review of Conceptual Approaches and Empirical Evidence on Probability and Nonprobability Sample Survey Research." *Journal of Survey Statistics and Methodology* smz041. doi: 10.1093/jssam/smz041.
- Datareportal. 2021. "Digital 2021 April Global Statshot Report." Retrieved August 10, 2021 (<https://datareportal.com/reports/digital-2021-april-global-statshot>).

- Elçi, Ezgi, Eda Kirişçioglu, and Aysen Üstübcü. 2021. "How COVID-19 Financially Hit Urban Refugees: Evidence from Mixed-method Research with Citizens and Syrian Refugees in Turkey." *Disasters* disa.12498. doi: 10.1111/disa.12498.
- Erlinghagen, Marcel, Andreas Ette, Norbert F. Schneider, and Nils Witte, eds. 2021. *The Global Lives of German Migrants. Consequences of International Migration Across the Life Course*. Cham: Springer International Publishing.
- Erlinghagen, Marcel, and Norbert Schneider. 2020. *German Emigration and Remigration Panel Study (GERPS) 2018 (Data File Version 1.0.0)*. Köln: GESIS Data Archive.
- Ersanilli, Evelyn, and Marieke van der Gaag. 2020. *MOBILISE Data Report. Online Surveys. Wave 1. preprint*. SocArXiv. doi: 10.31235/osf.io/79gca.
- Ersanilli, Evelyn, and Ruud Koopmans. 2013. *The Six Country Immigrant Integration Comparative Survey (SCIICS). Technical Report*. Berlin: Wissenschaftszentrum Berlin für Sozialforschung (WZB).
- Ette, Andreas, Jean Philippe Décieux, Marcel Erlinghagen, Andreas Genoni, Jean Guedes Auditor, Frederik Knirsch, Simon Kühne, Luisa Mörchen, Matthias Sand, Norbert F. Schneider, and Nils Witte. 2020. *German Emigration and Remigration Panel Study (GERPS). Methodology and Data Manual of the Baseline Survey (Wave 1)*. 01/2020. Wiesbaden: Bundesinstitut für Bevölkerungsforschung.
- Ette, Andreas, Lenore Sauer, Friedrich Scheller, Dawid Bekalarczyk, Marcel Erlinghagen, Marcus Engler, Jan Schneider, and Caroline Schultz. 2015. *International Mobil. Dokumentation Der Befragung von Auswanderern Und Rückwanderern Aus Deutschland*. Wiesbaden: Bundesinstitut für Bevölkerungsforschung.
- Facebook. 2021a. "About Ad Auctions." *Business Help Center*. Retrieved February 23, 2021 (<https://www.facebook.com/business/help/430291176997542>).
- Facebook. 2021b. "About Budgets." *Business Help Center*. Retrieved February 23, 2021 (<https://www.facebook.com/business/help/214319341922580>).
- Facebook. 2021c. "About Placements in Ads Manager." *Business Help Center*. Retrieved February 23, 2021 (<https://www.facebook.com/business/help/407108559393196>).
- Facebook. 2021d. "Add URL Parameters to Your Ads." *Business Help Center*. Retrieved February 23, 2021 (<https://www.facebook.com/business/help/1016122818401732>).
- Font, Joan, and Mónica Méndez. 2013. "The Methodological Challenges of Surveying Populations of Immigrant Origin." Pp. 11–41 in *Surveying Ethnic Minorities and Immigrant Populations. Methodological Challenges and Research Strategies*, edited by J. Font and M. Méndez. Amsterdam: Amsterdam University Press.
- Gambo, Sarah, and Bahire Ofe Özad. 2020. "The Demographics of Computer-Mediated Communication: A Review of Social Media Demographic Trends among Social Networking Site Giants." *Computers in Human Behavior Reports* 2:100016. doi: 10.1016/j.chbr.2020.100016.
- Ghimire, Dirgha J., Nathalie E. Williams, Arland Thornton, Linda Young-DeMarco, and Prem Bhandari. 2019. "Strategies for Origin-Based Surveying of International Migrants." *Journal of Ethnic and Migration Studies* 45(7):1185–1206. doi: 10.1080/1369183X.2017.1394178.

- Grow, André, Daniela Perrotta, Emanuele Del Fava, Jorge Cimentada, Francesco Rampazzo, Sofia Gil-Clavel, and Emilio Zagheni. 2020. "Addressing Public Health Emergencies via Facebook Surveys. Advantages, Challenges, and Practical Considerations." *Journal of Medical Internet Research* 22(12):e20653. doi: 10.2196/20653.
- Guillory, Jamie, Kristine F. Wiant, Matthew Farrelly, Leah Fiacco, Ishrat Alam, Leah Hoffman, Erik Crankshaw, Janine Delahanty, and Tesfa N. Alexander. 2018. "Recruiting Hard-to-Reach Populations for Survey Research. Using Facebook and Instagram Advertisements and In-Person Intercept in LGBT Bars and Nightclubs to Recruit LGBT Young Adults." *J Med Internet Res* 20(6):e197. doi: 10.2196/jmir.9461.
- Haan, Marieke, Peter Lugtig, and Vera Toepoel. 2019. "Can We Predict Device Use? An Investigation into Mobile Device Use in Surveys." *International Journal of Social Research Methodology* 1–15. doi: 10.1080/13645579.2019.1593340.
- Heckathorn, Douglas D. 1997. "Respondent-Driven Sampling. A New Approach to the Study of Hidden Populations." *Social Problems* 44(2):174–99.
- Humpert, Andreas, and Klaus Schneiderheinze. 2000. "Stichprobenziehung Für Telefonische Zuwandererumfragen. Einsatzmöglichkeiten Der Namensforschung." *ZUMA Nachrichten* 24(47):36–63.
- Iannelli, Laura, Fabio Giglietto, Luca Rossi, and Elisabetta Zurovac. 2018. "Facebook Digital Traces for Survey Research. Assessing the Efficiency and Effectiveness of a Facebook Ad-Based Procedure for Recruiting Online Survey Respondents in Niche and Difficult-to-Reach Populations." *Social Science Computer Review*. doi: 10.1177/0894439318816638.
- Kaša, Rita, and Inta Mieriņa, eds. 2019. *The Emigrant Communities of Latvia: National Identity, Transnational Belonging, and Diaspora Politics*. Cham: Springer International Publishing.
- Keusch, Florian, and Ting Yan. 2017. "Web Versus Mobile Web: An Experimental Study of Device Effects and Self-Selection Effects." *Social Science Computer Review* 35(6):751–69. doi: 10.1177/0894439316675566.
- Kühne, Simon, and Zaza Zindel. 2020. "Using Facebook and Instagram to Recruit Web Survey Participants. A Step-by-Step Guide and Application." *Survey Methods: Insights from the Field*. doi: 10.13094/SMIF-2020-00017.
- McAleese, Sara, Barbara Clyne, Anne Matthews, Ruairí Brugha, and Niamh Humphries. 2016. "Gone for Good? An Online Survey of Emigrant Health Professionals Using Facebook as a Recruitment Tool." *Human Resources for Health* 14(S1):34. doi: 10.1186/s12960-016-0130-y.
- Mieriņa, Inta. 2019. "An Integrated Approach to Surveying Emigrants Worldwide." Pp. 13–33 in *The Emigrant Communities of Latvia: National Identity, Transnational Belonging, and Diaspora Politics, IMISCOE Research Series*, edited by R. Kaša and I. Mieriņa. Cham: Springer International Publishing.
- Olamijuwon, Emmanuel Olawale. 2021. "Characterizing Low Effort Responding among Young African Adults Recruited via Facebook Advertising" edited by Z. Gao. *PLOS ONE* 16(5):e0250303. doi: 10.1371/journal.pone.0250303.
- Online Advertising School (OAS). 2016. "Facebook Ads Not Delivering? Here's How to Fix Them." Retrieved March 9, 2021 (<https://web.archive.org/web/20161114220245/http://www.onlineadvertisingschool.com/facebook-ads-not-delivering-heres-fix/>).

- Pfeiffer, Iris, and Andreas Heimer. 2007. *Gründe Für Die Auswanderung von Fach- Und Führungskräften Aus Wirtschaft Und Wissenschaft*. Berlin: Prognos.
- Pisarevskaya, Asya, Nathan Levy, Peter Scholten, and Joost Jansen. 2020. "Mapping Migration Studies. An Empirical Analysis of the Coming of Age of a Research Field." *Migration Studies* 8(3):455–81. doi: 10.1093/migration/mnz031.
- Pötzschke, Steffen. forthcoming. "The Sampling of Migrants through Advertisements on Facebook and Instagram." in *Practitioners' Guide on Harnessing Big Data for Migration Policy*, edited by IOM Global Migration Data Analysis Centre. Geneva: International Organization for Migration.
- Pötzschke, Steffen, and Michael Braun. 2017. "Migrant Sampling Using Facebook Advertisements: A Case Study of Polish Migrants in Four European Countries." *Social Science Computer Review* 35(5):633–53. doi: 10.1177/0894439316666262.
- Prandner, Dimitri, and Martin Weichbold. 2019. "Building a Sampling Frame for Migrant Populations via an Onomastic Approach. Lesson Learned from the Austrian Immigrant Survey 2016." *Survey Methods: Insights from the Field (SMIF)*. doi: 10.13094/SMIF-2019-00006.
- Recchi, Ettore, Adrian Favell, Fulya Apaydin, Roxana Barbulescu, Michael Braun, Irina Ciornei, Niall Cunningham, Juan Díez Medrano, Deniz Duru, Laurie Hanquinet, Steffen Pötzschke, David Reimer, Justyna Salamońska, Mike Savage, Janne Solgaard Jensen, and Albert Varela, eds. 2019. *Everyday Europe. Social Transnationalism in an Unsettled Continent*. Bristol: Policy Press.
- Reichel, David, and Laura Morales. 2017. "Surveying Immigrants without Sampling Frames. Evaluating the Success of Alternative Field Methods." *Comparative Migration Studies* 5(1):1. doi: 10.1186/s40878-016-0044-9.
- Rocheva, Anna, Evgeni Varshaver, and Nataliya Ivanova. forthcoming. "Targeting on Social Networking Sites (SNS) as a Sampling Strategy for Online Migrant Surveys. The Challenge of Biases and Search for Possible Solutions." in *Migration Research in a Digitized World. Using Innovative Technology to Tackle Methodological Challenges, IMISCOE Research Series*, edited by Pötzschke, Steffen and Rinken, Sebastian. Cham: Springer.
- Rosenzweig, Leah R., Parrish Bergquist, Katherine Hoffmann Pham, Francesco Rampazzo, and Matto Mildenerger. 2020. *Survey Sampling in the Global South Using Facebook Advertisements*. preprint. SocArXiv.
- Salentin, Kurt, and Hans Schmeets. 2017. "Sampling Immigrants in the Netherlands and Germany." *Comparative Migration Studies* 5(21). doi: 10.1186/s40878-017-0062-2.
- Samuels, David, and Cesar Zucco. 2014. "The Power of Partisanship in Brazil: Evidence from Survey Experiments." *American Journal of Political Science* 58(1):212–25.
- Schneider, Daniel, and Kristen Harknett. 2019. "What's to Like? Facebook as a Tool for Survey Data Collection." *Sociological Methods & Research* 004912411988247. doi: 10.1177/0049124119882477.
- Schoumaker, Bruno, and Cris Beauchemin. 2015. "Reconstructing Trends in International Migration with Three Questions in Household Surveys. Lessons from the MAFE Project." *Demographic Research* 32:983–1030.
- Schupp, Jürgen, Nico A. Siegel, Marcel Erlinghagen, and Tim Stegmann. 2008. *Leben außerhalb Deutschlands. Eine Machbarkeitsstudie zur Realisierung von Auslandsbefragungen auf Basis*

- des Sozio-oekonomischen Panels (SOEP)*. Berlin; DIW: Deutsches Institut für Wirtschaftsforschung.
- Sen, Indira, Fabian Flöck, Katrin Weller, Bernd Weiß, and Claudia Wagner. 2021. “A Total Error Framework for Digital Traces of Human Behavior on Online Platforms.” *Public Opinion Quarterly* online first. doi: 10.1093/poq/nfab018.
- Thomson, Robert, and Naoya Ito. 2014. “Facebook Advertisements for Survey Participant Recruitment. Considerations from a Multi-National Study.” *International Journal of Electronic Commerce Studies* 5(2):199–218. doi: 10.7903/ijecs.1175.
- Tourangeau, Roger. 2014. “Defining Hard-to-Survey Populations.” Pp. 3–20 in *Hard-to-survey populations*, edited by R. Tourangeau, B. Edwards, T. P. Johnson, K. M. Wolter, and N. Bates. New York: Cambridge University Press.
- UNESCO and Metropolis Canada. 2020. *Measuring Attitudes & Perceptions on the Impact of COVID-19 in Select Sub-Saharan African Cities*.
- Universal Postal Union. 2012. *Addressing the World. An Address for Everyone*. Berne: Universal Postal Union.
- Universal Postal Union. 2021. “Postal Addressing Systems in Member Countries and/or Territories. Japan.” *Addressing Solutions*. Retrieved May 15, 2021 (<https://www.upu.int/en/Postal-Solutions/Programmes-Services/Addressing-Solutions#addressing-knowledge-centre>).
- Whitaker, Christopher, Sharon Stevelink, and Nicola Fear. 2017. “The Use of Facebook in Recruiting Participants for Health Research Purposes. A Systematic Review.” *Journal for Medical Internet Research* 19(8):2017/09/05. doi: 10.2196/jmir.7071.
- Zhang, Baobao, Matto Mildenerger, Peter D. Howe, Jennifer Marlon, Seth A. Rosenthal, and Anthony Leiserowitz. 2020. “Quota Sampling Using Facebook Advertisements.” *Political Science Research and Methods* 8(3):558–64. doi: 10.1017/psrm.2018.49.

7 Appendix

Appendix 1: Targeted countries, monthly active users, and sampling regions

Country name	Country code	Monthly active users ¹	Sampling region ²
Algeria	DZ	30,000	Africa 1 - Biggest 5
Egypt	EG	11,000	Africa 1 - Biggest 5
Morocco	MA	17,000	Africa 1 - Biggest 5
Nigeria	NG	24,000	Africa 1 - Biggest 5
Tunisia	TN	11,000	Africa 1 - Biggest 5
Angola	AO	1,000	Africa 2 - Southern Africa
Botswana	BW	1,000	Africa 2 - Southern Africa
Comoros	KM	1,000	Africa 2 - Southern Africa
Democratic Republic of the Congo	CD	1,900	Africa 2 - Southern Africa
Gabon	GA	1,000	Africa 2 - Southern Africa
Kenya	KE	3,900	Africa 2 - Southern Africa
Lesotho	LS	1,000	Africa 2 - Southern Africa
Madagascar	MG	1,000	Africa 2 - Southern Africa
Malawi	MW	1,000	Africa 2 - Southern Africa
Mauritius	MU	1,000	Africa 2 - Southern Africa
Mozambique	MZ	1,000	Africa 2 - Southern Africa
Namibia	NA	1,000	Africa 2 - Southern Africa
Republic of the Congo	CG	1,000	Africa 2 - Southern Africa
Seychelles	SC	1,000	Africa 2 - Southern Africa
South Africa	ZA	10,000	Africa 2 - Southern Africa
Swaziland	SZ	1,000	Africa 2 - Southern Africa
Tanzania	TZ	2,000	Africa 2 - Southern Africa
Zambia	ZM	1,000	Africa 2 - Southern Africa
Zimbabwe	ZW	1,000	Africa 2 - Southern Africa
Benin	BJ	2,900	Africa 3 - Remaining African countries
British Indian Ocean Territory	IO	1,000	Africa 3 - Remaining African countries
Burkina Faso	BF	1,000	Africa 3 - Remaining African countries
Burundi	BI	1,000	Africa 3 - Remaining African countries
Cameroon	CM	1,900	Africa 3 - Remaining African countries
Cape Verde	CV	1,000	Africa 3 - Remaining African countries
Central African Republic	CF	1,000	Africa 3 - Remaining African countries
Chad	TD	1,000	Africa 3 - Remaining African countries
Côte d'Ivoire	CI	2,200	Africa 3 - Remaining African countries
Djibouti	DJ	1,000	Africa 3 - Remaining African countries
Equatorial Guinea	GQ	1,000	Africa 3 - Remaining African countries
Eritrea	ER	1,000	Africa 3 - Remaining African countries
Ethiopia	ET	3,600	Africa 3 - Remaining African countries
French Southern Territories	TF	1,000	Africa 3 - Remaining African countries
Ghana	GH	8,100	Africa 3 - Remaining African countries
Guinea	GN	1,000	Africa 3 - Remaining African countries
Guinea-Bissau	GW	1,000	Africa 3 - Remaining African countries
Liberia	LR	1,000	Africa 3 - Remaining African countries

Country name	Country code	Monthly active users ¹	Sampling region ²
Libya	LY	4,600	Africa 3 - Remaining African countries
Mali	ML	1,000	Africa 3 - Remaining African countries
Mauritania	MR	1,000	Africa 3 - Remaining African countries
Mayotte	YT	1,000	Africa 3 - Remaining African countries
Niger	NE	1,000	Africa 3 - Remaining African countries
Réunion	RE	1,000	Africa 3 - Remaining African countries
Rwanda	RW	1,000	Africa 3 - Remaining African countries
Saint Helena	SH	1,000	Africa 3 - Remaining African countries
Sao Tome and Principe	ST	1,000	Africa 3 - Remaining African countries
Senegal	SN	1,100	Africa 3 - Remaining African countries
Sierra Leone	SL	1,000	Africa 3 - Remaining African countries
Somalia	SO	2,400	Africa 3 - Remaining African countries
South Sudan	SS	1,000	Africa 3 - Remaining African countries
The Gambia	GM	1,000	Africa 3 - Remaining African countries
Togo	TG	1,000	Africa 3 - Remaining African countries
Uganda	UG	1,500	Africa 3 - Remaining African countries
Western Sahara	EH	1,000	Africa 3 - Remaining African countries
India	IN	79,000	Asia 1 - Biggest 6
Indonesia	ID	71,000	Asia 1 - Biggest 6
Philippines	PH	14,000	Asia 1 - Biggest 6
Thailand	TH	14,000	Asia 1 - Biggest 6
Turkey	TR	45,000	Asia 1 - Biggest 6
Russia	RU	29,000	Asia 1 - Biggest 6
Armenia	AM	1,000	Asia 2 - Western + Central Asia
Azerbaijan	AZ	1,300	Asia 2 - Western + Central Asia
Bahrain	BH	1,000	Asia 2 - Western + Central Asia
Georgia	GE	2,000	Asia 2 - Western + Central Asia
Iraq	IQ	14,000	Asia 2 - Western + Central Asia
Israel + Palestine	IL	4,500	Asia 2 - Western + Central Asia
Jordan	JO	4,000	Asia 2 - Western + Central Asia
Kazakhstan	KZ	2,200	Asia 2 - Western + Central Asia
Kuwait	KW	1,500	Asia 2 - Western + Central Asia
Kyrgyzstan	KG	1,000	Asia 2 - Western + Central Asia
Lebanon	LB	3,800	Asia 2 - Western + Central Asia
Oman	OM	1,000	Asia 2 - Western + Central Asia
Palestine	PS	2,100	Asia 2 - Western + Central Asia
Qatar	QA	1,400	Asia 2 - Western + Central Asia
Saudi Arabia	SA	3,900	Asia 2 - Western + Central Asia
Tajikistan	TJ	1,000	Asia 2 - Western + Central Asia
Turkmenistan	TM	1,000	Asia 2 - Western + Central Asia
United Arab Emirates	AE	6,700	Asia 2 - Western + Central Asia
Uzbekistan	UZ	1,000	Asia 2 - Western + Central Asia
Yemen	YE	1,000	Asia 2 - Western + Central Asia
Afghanistan	AF	3,000	Asia 3 - Southern + Eastern + South-eastern Asia
Bangladesh	BD	11,000	Asia 3 - Southern + Eastern + South-eastern Asia

Country name	Country code	Monthly active users ¹	Sampling region ²
Bhutan	BT	1,000	Asia 3 - Southern + Eastern + South-eastern Asia
Brunei	BN	1,000	Asia 3 - Southern + Eastern + South-eastern Asia
Cambodia	KH	2,200	Asia 3 - Southern + Eastern + South-eastern Asia
China	CN	8,300	Asia 3 - Southern + Eastern + South-eastern Asia
Hong Kong	HK	2,600	Asia 3 - Southern + Eastern + South-eastern Asia
Japan	JP	5,600	Asia 3 - Southern + Eastern + South-eastern Asia
Laos	LA	1,200	Asia 3 - Southern + Eastern + South-eastern Asia
Macau	MO	1,000	Asia 3 - Southern + Eastern + South-eastern Asia
Malaysia	MY	5,700	Asia 3 - Southern + Eastern + South-eastern Asia
Maldives	MV	1,000	Asia 3 - Southern + Eastern + South-eastern Asia
Mongolia	MN	1,200	Asia 3 - Southern + Eastern + South-eastern Asia
Myanmar	MM	1,600	Asia 3 - Southern + Eastern + South-eastern Asia
Nepal	NP	1,300	Asia 3 - Southern + Eastern + South-eastern Asia
Pakistan	PK	9,600	Asia 3 - Southern + Eastern + South-eastern Asia
Singapore	SG	2,700	Asia 3 - Southern + Eastern + South-eastern Asia
South Korea	KR	4,100	Asia 3 - Southern + Eastern + South-eastern Asia
Sri Lanka	LK	1,500	Asia 3 - Southern + Eastern + South-eastern Asia
Taiwan	TW	2,500	Asia 3 - Southern + Eastern + South-eastern Asia
Timor-Leste	TL	1,000	Asia 3 - Southern + Eastern + South-eastern Asia
Vietnam	VN	6,400	Asia 3 - Southern + Eastern + South-eastern Asia
Argentina	AR	6,800	Middle + South America 1 - Biggest 5
Brazil	BR	18,000	Middle + South America 1 - Biggest 5
Chile	CL	5,900	Middle + South America 1 - Biggest 5
Colombia	CO	5,400	Middle + South America 1 - Biggest 5
Mexico	MX	35,000	Middle + South America 1 - Biggest 5
Anguilla	AI	1,000	Middle + South America 2 - Central America + Caribbean
Antigua	AG	1,000	Middle + South America 2 - Central America + Caribbean
Barbados	BB	1,000	Middle + South America 2 - Central America + Caribbean
Belize	BZ	1,000	Middle + South America 2 - Central America + Caribbean
Bouvet Island	BV	1,000	Middle + South America 2 - Central America + Caribbean
British Virgin Islands	VG	1,000	Middle + South America 2 - Central America + Caribbean
Cayman Islands	KY	1,000	Middle + South America 2 - Central America + Caribbean
Costa Rica	CR	1,100	Middle + South America 2 - Central America + Caribbean
Dominica	DM	1,000	Middle + South America 2 - Central America + Caribbean
Dominican Republic	DO	3,200	Middle + South America 2 - Central America + Caribbean
El Salvador	SV	1,900	Middle + South America 2 - Central America + Caribbean
Grenada	GD	1,000	Middle + South America 2 - Central America + Caribbean
Guadeloupe	GP	1,000	Middle + South America 2 - Central America + Caribbean
Guatemala	GT	1,400	Middle + South America 2 - Central America + Caribbean
Haiti	HT	1,000	Middle + South America 2 - Central America + Caribbean
Honduras	HN	1,000	Middle + South America 2 - Central America + Caribbean
Jamaica	JM	1,000	Middle + South America 2 - Central America + Caribbean
Martinique	MQ	1,000	Middle + South America 2 - Central America + Caribbean
Montserrat	MS	1,000	Middle + South America 2 - Central America + Caribbean
Netherlands Antilles	AN	1,000	Middle + South America 2 - Central America + Caribbean

Country name	Country code	Monthly active users ¹	Sampling region ²
Nicaragua	NI	1,000	Middle + South America 2 - Central America + Caribbean
Panama	PA	1,000	Middle + South America 2 - Central America + Caribbean
Puerto Rico	PR	1,000	Middle + South America 2 - Central America + Caribbean
Saint Barthélemy	BL	1,000	Middle + South America 2 - Central America + Caribbean
Saint Kitts and Nevis	KN	1,000	Middle + South America 2 - Central America + Caribbean
Saint Martin	MF	1,000	Middle + South America 2 - Central America + Caribbean
Saint Vincent and the Grenadines	VC	1,000	Middle + South America 2 - Central America + Caribbean
Sint Maarten	SX	1,000	Middle + South America 2 - Central America + Caribbean
South Georgia and the South Sandwich Islands	GS	1,000	Middle + South America 2 - Central America + Caribbean
St. Lucia	LC	1,000	Middle + South America 2 - Central America + Caribbean
The Bahamas	BS	1,000	Middle + South America 2 - Central America + Caribbean
Trinidad and Tobago	TT	1,000	Middle + South America 2 - Central America + Caribbean
Turks and Caicos Islands	TC	1,000	Middle + South America 2 - Central America + Caribbean
US Virgin Islands	VI	1,000	Middle + South America 2 - Central America + Caribbean
Aruba	AW	1,000	Middle + South America 3 - South America
Bolivia	BO	1,400	Middle + South America 3 - South America
Bonaire, Sint Eustatius and Saba	BQ	1,000	Middle + South America 3 - South America
Curaçao	CW	1,000	Middle + South America 3 - South America
Ecuador	EC	2,200	Middle + South America 3 - South America
Falkland Islands	FK	1,000	Middle + South America 3 - South America
French Guiana	GF	1,000	Middle + South America 3 - South America
Guyana	GY	1,000	Middle + South America 3 - South America
Paraguay	PY	1,600	Middle + South America 3 - South America
Peru	PE	3,800	Middle + South America 3 - South America
Suriname	SR	1,000	Middle + South America 3 - South America
Uruguay	UY	1,000	Middle + South America 3 - South America
Venezuela	VE	2,100	Middle + South America 3 - South America
Antarctica	AQ	1,000	Middle + South America 3 - South America
United States of America	US	160,000	North America 1 - Biggest
Bermuda	BM	1,000	North America 2 - Remaining North American countries
Canada	CA	27,000	North America 2 - Remaining North American countries
Greenland	GL	1,000	North America 2 - Remaining North American countries
Saint Pierre and Miquelon	PM	1,000	North America 2 - Remaining North American countries
Australia	AU	23,000	Oceania 1 - Biggest
American Samoa	AS	1,000	Oceania 2 - Polynesia + Melanesia + Micronesia
Christmas Island	CX	1,000	Oceania 2 - Polynesia + Melanesia + Micronesia
Cocos (Keeling) Islands	CC	1,000	Oceania 2 - Polynesia + Melanesia + Micronesia
Cook Islands	CK	1,000	Oceania 2 - Polynesia + Melanesia + Micronesia
Federated States of Micronesia	FM	1,000	Oceania 2 - Polynesia + Melanesia + Micronesia
Fiji	FJ	1,000	Oceania 2 - Polynesia + Melanesia + Micronesia
French Polynesia	PF	1,000	Oceania 2 - Polynesia + Melanesia + Micronesia
Guam	GU	1,000	Oceania 2 - Polynesia + Melanesia + Micronesia
Heard Island and McDonald Islands	HM	1,000	Oceania 2 - Polynesia + Melanesia + Micronesia

Country name	Country code	Monthly active users ¹	Sampling region ²
Kiribati	KI	1,000	Oceania 2 - Polynesia + Melanesia + Micronesia
Marshall Islands	MH	1,000	Oceania 2 - Polynesia + Melanesia + Micronesia
Nauru	NR	1,000	Oceania 2 - Polynesia + Melanesia + Micronesia
New Caledonia	NC	1,000	Oceania 2 - Polynesia + Melanesia + Micronesia
New Zealand	NZ	5,900	Oceania 2 - Polynesia + Melanesia + Micronesia
Niue	NU	1,000	Oceania 2 - Polynesia + Melanesia + Micronesia
Norfolk Island	NF	1,000	Oceania 2 - Polynesia + Melanesia + Micronesia
Northern Mariana Islands	MP	1,000	Oceania 2 - Polynesia + Melanesia + Micronesia
Palau	PW	1,000	Oceania 2 - Polynesia + Melanesia + Micronesia
Papua New Guinea	PG	1,000	Oceania 2 - Polynesia + Melanesia + Micronesia
Pitcairn	PN	1,000	Oceania 2 - Polynesia + Melanesia + Micronesia
Samoa	WS	1,000	Oceania 2 - Polynesia + Melanesia + Micronesia
Solomon Islands	SB	1,000	Oceania 2 - Polynesia + Melanesia + Micronesia
Tokelau	TK	1,000	Oceania 2 - Polynesia + Melanesia + Micronesia
Tonga	TO	1,000	Oceania 2 - Polynesia + Melanesia + Micronesia
Tuvalu	TV	1,000	Oceania 2 - Polynesia + Melanesia + Micronesia
United States Minor Outlying Islands	UM	1,000	Oceania 2 - Polynesia + Melanesia + Micronesia
Vanuatu	VU	1,000	Oceania 2 - Polynesia + Melanesia + Micronesia
Wallis and Futuna	WF	1,000	Oceania 2 - Polynesia + Melanesia + Micronesia

Note that the shown statistics were provided by the Facebook Ad Manager (FAM).

¹ Values extracted on August 8, 2020, using the Facebook Graph API. For privacy reasons, FAM does not provide detailed estimates for user groups with less than 1,000 individuals. Consequently, in this table “1,000” should always be read as “1,000 or less”.

² The labels provided in this column give a general description of the corresponding country group, please note that the labels do not necessarily fit all countries in a group (e.g., “Middle + South America 3” also includes several Caribbean countries. We, furthermore, differentiate between North, Middle and South America. Consequently, Mexico is included in the second instead in the first of these groups.

Appendix 2: Approximate sampling cost per completed survey by sampling region (advertisement sample only)

Sampling regions	Completed surveys (with matching geographic indicators)	Amount spent (on ads in sampling regions)	Costs per completed survey (with matching geographic indicators)
Africa 1 - Biggest 5	136	174.57 €	1.28 €
Africa 2 - Southern Africa	239	173.47 €	0.73 €
Africa 3 - Remaining countries	56	174.34 €	3.11 €
Asia - Biggest 6	334	189.60 €	0.57 €
Asia 2 - W + Cen. Asia	209	171.93 €	0.82 €
Asia 3 - S. + E. + SE Asia	278	206.70 €	0.74 €
Middle + S. America 1 - Biggest 5	454	184.51 €	0.41 €
Middle + S. America 2 - Cent. A. + Caribbean	170	189.31 €	1.11 €
Middle + S. America 3 - South America	209	190.91 €	0.91 €
North America 1 - Biggest (USA)	295	130.83 €	0.44 €
North America 2 - Remaining countries	393	142.99 €	0.36 €
Oceania 1 - Biggest (Australia)	277	149.71 €	0.54 €
Oceania 2 - Remaining countries	258	144.03 €	0.56 €

Note: The costs per completed survey for the different regions are only approximations and likely to slightly overestimate the actual costs. Our calculation uses the full budget spent on each region, but we do not include all observations assigned to these regions according to the registered URL parameters. Instead, the amount spent is divided only by the number of surveys completed by German emigrants for which the regional information retrieved from the URL parameter match the location country of residence stated in the survey. Statistics were provided by the Facebook Ad Manager.

Appendix 3: Sample composition among avenues of recruitment

		Total ¹		Advertisements				Facebook page		Snowball	
		Percent	Frequency	Facebook		Instagram		Percent	Frequency	Percent	Frequency
				Percent	Frequency	Percent	Frequency				
Sex	Male	42.82	1623	45.01	1363	30.99	119	42.47	31	27.32	50
	Female	56.91	2157	54.76	1658	68.75	264	56.16	41	72.13	132
	Diverse	0.26	10	0.23	7	0.26	1	1.37	1	0.55	1
Age (mean)		46.64	3760	47.69	3002	37.85	380	48.32	74	46.25	183
Education ²	Low	10.21	385	10.97	331	7.69	29	15.07	11	1.10	2
	Intermediate	19.09	720	19.62	592	17.51	66	10.96	8	18.13	33
	High	70.70	2666	69.41	2094	74.80	282	73.97	54	80.77	147
Employment	Employed	69.23	2590	68.74	2058	69.54	258	68.92	51	77.05	141
	Unemployed	3.69	138	3.61	108	5.12	19	4.05	3	2.19	4
	Other	23.23	869	24.48	733	14.82	55	24.32	18	19.13	35
Duration of stay ⁴	In education	3.85	144	3.17	95	10.51	39	2.70	2	1.64	3
	< 1 year	3.64	138	3.50	106	5.48	21	6.76	5	1.63	3
	1 year	9.07	344	8.65	262	12.27	47	9.46	7	8.70	16
	2 years	7.97	302	7.69	233	10.97	42	9.46	7	6.52	12
	3 years	6.33	240	6.47	196	6.01	23	5.41	4	4.89	9
	4 years	4.99	189	4.69	142	6.79	26	2.70	2	4.35	8
	5 years	5.14	195	4.85	147	5.22	20	1.35	1	11.41	21
	6-10 years	18.04	684	17.90	542	17.49	67	13.51	10	21.74	40
	11-20 years	21.58	818	21.96	665	17.49	67	24.32	18	25.00	46
	21-30 years	10.95	415	11.06	335	9.92	38	16.22	12	9.24	17
	31-40 years	5.94	225	6.21	188	4.44	17	6.76	5	3.26	6
41-50 years	3.14	119	3.37	102	2.09	8	1.35	1	2.17	4	
51-60 years	2.22	84	2.48	75	1.31	5	2.70	2	1.09	2	
> 60 years	1.00	38	1.16	35	0.52	2	0.00	0	0.00	0	

¹ Note that the total frequency (fourth column) cannot be reproduced by summarizing over the subgroups. This is due to missing values in the platform variable (“Advertisements”) which indicates whether a respondent was recruited by an advertisement via Facebook or Instagram.

² Educational attainment levels are constructed as follows: “Low” (still student, no degree, Polytechnic Secondary School, G 8/9, Secondary General School-leaving Certificate), “Intermediate” (Polytechnic Secondary School, G 10, Intermediate School-leaving Certificate), “High” (Applied University Entrance Qualification (FH), University Entrance Qualification)

³ The category “Other” includes retired or chronically ill respondents or homemakers.

⁴ Duration of stay in the country of residence.