

Report No: AUS0000423

Afghanistan

World Bank Phone Survey of Afghan Returnees Methodology and representativeness

July 2018

POV



© 2018 The World Bank
1818 H Street NW, Washington DC 20433
Telephone: 202-473-1000; Internet: www.worldbank.org

Some rights reserved

This work is a product of the staff of The World Bank. The findings, interpretations, and conclusions expressed in this work do not necessarily reflect the views of the Executive Directors of The World Bank or the governments they represent. The World Bank does not guarantee the accuracy of the data included in this work. The boundaries, colors, denominations, and other information shown on any map in this work do not imply any judgment on the part of The World Bank concerning the legal status of any territory or the endorsement or acceptance of such boundaries.

Rights and Permissions

The material in this work is subject to copyright. Because The World Bank encourages dissemination of its knowledge, this work may be reproduced, in whole or in part, for noncommercial purposes as long as full attribution to this work is given.

Attribution—Please cite the work as follows: “World Bank. 2018. World Bank Phone Survey of Afghan Returnees: Methodology and representativeness. © World Bank.”

All queries on rights and licenses, including subsidiary rights, should be addressed to World Bank Publications, The World Bank Group, 1818 H Street NW, Washington, DC 20433, USA; fax: 202-522-2625; e-mail: pubrights@worldbank.org.

World Bank Phone Survey of Afghan Returnees¹

Methodology and representativeness

Table of Contents

1. Introduction	4
2. Identifying the population of interest	5
1. Stage 1: Random Digit Dialing and Interactive Voice Recognition	6
2. Stage 2: Respondent driven sampling	8
3. Survey design and execution	9
1. Design of questionnaire	10
2. Execution of survey	11
3. Response rates of phone interview	12
4. Cost of the survey	12
4. Effects of applied sampling methods	13
1. IVR selection	13
2. RDS selection	14
3. Discussion of sampling methods	16
5. Representativeness	17
1. The true population of Afghan returnees	17
2. Bivariate test of representativeness	20
3. Multivariate test of representativeness	21
6. Adjusting for survey bias: Survey weights	23
7. Conclusion	24
Sources	25
Appendix 1	26

¹ This activity was financed by the DFID Trust Fund for Forced Displacement, and their support is gratefully acknowledged. Authors: Nandini Krishnan, Christina Wieser and Thea Yde-Jensen.

1. Introduction

In early 2017, as Afghanistan witnessed the largest influx of refugee returns in recent years, little was known about the location patterns and needs of the more than 2 million Afghan returnees who had come back since 2015. Separate agencies managed registration databases for different groups, and the lack of access to SIM cards (due to the lack of necessary documentation) severely limited the broad applicability of findings from phone monitoring surveys. At the same time, existing data sources had not been fully exploited to characterize different groups of displaced populations, nor had attempts been made to link them to data collected by humanitarian agencies.

While by no means perfect, a joint World Bank-UNHCR-IOM engagement has evolved significantly to a point where several key elements are now in place to enable a systematic tracking and needs assessment of Afghan returnees. These elements include (i) data sharing agreements between the UNHCR and the WB and a draft agreement between WB and IOM; (ii) implementation of universal SIM card distribution to returnees at registration/encashment centers; (iii) tested protocols and questionnaires for phone surveys to track and monitor returnee movements and socio-economic outcomes; and (iv) analytical inputs based on existing data, revised as new information is incorporated. The WB team's role has largely been in strengthening coordination and collaboration for data sharing and joint analysis, leveraging the comparative advantage of humanitarian and development agencies; technical advice to UNHCR and IOM on sampling, questionnaire design, and interpretation of analytical findings; and in raising trust fund resources for a mobile phone survey to test random-digit-dialing integrated voice recognition software for the identification of a sample of returnees with mobile phones and a follow-up personal phone interview on location decisions, employment and livelihoods and access to basic social services.

This note describes the methodology underpinning (iii) above – protocols for data collection and questionnaires for phone surveys to track and monitor returnee movements and socio-economic outcomes. The objective of this exercise was to shed light on (i) movements and choice of location; (ii) labor market attachment and; (iii) access to services of post-2014 returnees². A related working paper describes the findings from this exercise related to the pre-and post-return lives and livelihoods of this most recent wave of Afghan returnees.

Collecting high-quality household data in conflict-affected countries is challenging because face-to-face interviews and identification of eligible participants are unfeasible. Instead, the increasing availability of mobile phones in low income countries opens for new and innovative ways of collecting data. Although mobile penetration rate is low in Afghanistan compared to other low-income countries, data suggest that the rates are increasing and while just 30% of Afghan household owned one or more mobile phones in 2007, that number had increased to 87% in 2017 (NRVA 2007/08 & The Asia Foundation, 2018). An additional difficulty given the objectives of this exercise was the lack of any comprehensive sample frame for returnees, either through a database of phone numbers for all returnees or through a stratum in the national surveys.

This data collection project relied on mobile technology to collect information on the re-settlement of Afghan returnees and this note describes the methodology applied to identify and survey eligible survey participants. A key challenge in undertaking mobile phone data collection was a clear identification of the

² Orange Door Research and Viamo were contracted to conduct the data collection

surveyed returnee population, with the intention of being able to draw more meaningful inferences on outcomes for the returnee population at large based on the survey design. The note will start by discussing the details of how survey participants were identified using a combination of Random Digit Dialing (RDD) and Interactive voice response (IVR) and respondent driven sampling (RDS). Section 3 will present the survey design, implementation and response rates. Section 4 discusses the implications of the sampling methods and section 5 analyzes the representativeness of the collected data. Section **Error! Reference source not found.** discusses the use of sampling weights to adjust for sampling bias and finally, section 7 summarizes up the lessons learned.

The key insights learned from collecting household data in a conflict country using mobile technology include:

- Relying on mobile technology to collect data in a conflict affected country is safer and less costly than collecting data face-to-face.
- RDD and IVR methods are useful to identify populations that are hard to locate. Relatively low mobile penetration rates in Afghanistan and a narrow target group extended the time (and costs) of using RDD and IVR to identify eligible survey participant. As mobile penetration rates increase, RDD and IVR will become increasingly useful in locating target populations when official records are not available.
- Respondent driven sampling can complement the RDD and IVR methods in identifying and increasing the sampling size. Respondent driven sampling can be both time and cost effective if the target population tend to know other eligible survey respondents. However, to avoid biasing the sample, it is important to limit the number of referrals each respondent can pass on.
- Phone survey interviews may be less personal but offer, at the same time, more anonymity to the respondents. In the context of returnees in Afghanistan, respondents were generally interested in participating in the phone survey.
- Relying on mobile technology to identify and interview respondents affects the representativeness of the survey. Young men from urban areas were more likely to respond to phone calls and hence participate in the interview. However, asking questions related to the household as a whole can compensate for the selection caused by relying on mobile technology for data collection.

2. Identifying the population of interest

The main purpose of collecting information on returnees was to assist the design and targeting of development programs for *recent* and *future* returnees. Afghans have been living in conflict for decades, and are among the largest protracted refugee populations in the world. Displacement due to violence remains a pervasive feature: in the 2013/14 ALCS, 2.4% of the population reported being internally displaced and 6.7% identified as post-2002 returnees. Hence, to ensure that the data reflects the situation of *recent* returnees, the target population was limited to Afghans who returned to Afghanistan in 2014 or later, also referred to as *post-2014 returnees*.

The project intended to fill data gaps on various aspects of post-2014 returnees' re-settlement, including the geographical location of re-settlement. The identification of eligible participants was therefore a crucial part of the project. Existing data registries could not be used to assist the identification as they covered a small proportion of returnees. Instead, the (somewhat) widespread use of mobile phones in Afghanistan

permitted identifying and interviewing eligible respondents (post-2014 returnees) using mobile technology.

In the first stage of the project, eligible survey participants were identified through Random Digit Dialing (RDD) and Interactive Voice Recognition (IVR) and in the second stage of the project, eligible survey participants were interviewed over the phone by trained enumerators and additional eligible survey respondents were identified through respondent driven sampling (RDS). Separating the RDD-IVR identification of eligible participants and the actual interview into two steps not only saved enumerators' time but brief interaction with potential survey participants can increase response rates of the actual (and lengthier) survey. In the absence of systematic non-response or refusal between the first stage (RDD-IVR) and the second stage (personal phone interview), this protocol can generate a representative sample of returnees with mobile phones. In practice, issues of non-response and refusal are common in phone surveys, and this survey is no exception. Efforts to boost the sample through RDS and observable bias in responses will be discussed in detail in section 2.2 and 4.2.

1. Stage 1: Random Digit Dialing and Interactive Voice Recognition

RDD and IVR were used to identify eligible survey participants, in principle, a representative sample of post-2014 returnees with access to mobile phones. The process was fully automated and the identification of phone numbers and screening interviews for eligibility were done using mobile technology and with no use of interviewers.

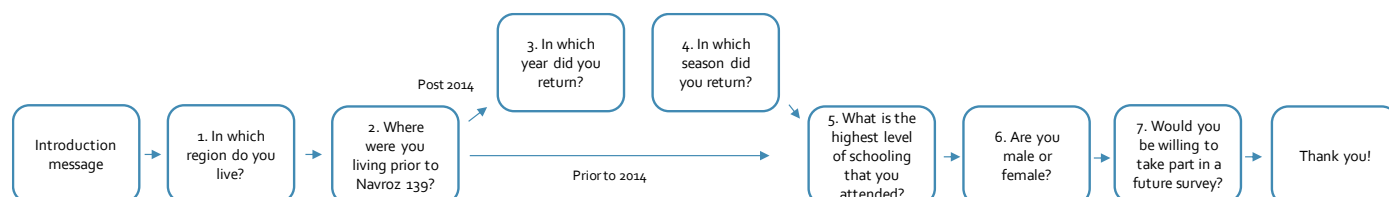
The first step of the process was to call Afghan phone numbers. This was done using an RDD technique that generates and dials random digits resembling the 11-digit structure of Afghan mobile phone numbers. The RDD was set up to only generate mobile phone numbers because landlines are uncommon in Afghanistan (and mostly found in government agencies, large businesses, etc. rather than households). The first two digits of Afghan phone numbers correspond to the international call code for Afghanistan (+93), the third digit for mobile phone numbers is 7 and the RDD process randomly generated the remaining 8-digits. The random digits, resembling Afghan mobile phone numbers, were dialed and calls were carried out every day of the week from 7.30 AM to 9.00 PM local time. As a result of the RDD process, each phone number was dialed one time³. Only dialed numbers that went on to a ringtone were considered actual phone numbers and calls that were answered, were greeted by an IVR.

The second step of the process was to identify post-2014 returnees willing to participate in the survey. This was done using an IVR technology that allows the phone to interact with respondents through the keypad and collects data on the correspondence with the respondents. The IVR uses an automated voice to interview respondents and in this case, a male voice was considered the better choice because mobile ownership is significantly higher among men than women in Afghanistan (The Asia Foundation, 2018). In practice, an automated male voice greeted the respondent and allowed him/her to choose between continuing the survey in one out of three languages. Once the language was selected, the automated voice introduced the survey and posed up to 7 questions mapping core characteristics of the respondent. If the respondent entered an incorrect response (for example the number "9" when the available response options are "1" or "2") the question was repeated. The question was also repeated if no response was

³ Most phone numbers were generated only once using the random digit generator and were called once. 2.5% of the 66,000 legit phone numbers were generated more than one time using the random digit generator and were therefore called more than one time.

entered within a few seconds. Only respondents who stayed on line to the final question were asked to indicate whether they were willing to participate in a future survey (see Figure 1). Being a post-2014 returnee (question 2) was the only requirement to be eligible to participate in the survey and the additional 4-6 questions were included in the IVR to establish confidence in the survey and increase the chances that the respondents would in fact be willing to participate in a lengthier survey when contacted again.

Figure 1 - Line of questions in IVR



Number of dials and response rates of RDD and IVR

In total, 2.6 million random digits were dialed; however, most of these digits were not actual phone numbers (see Table 1). Only 2.5 pct. or 66,000 of the digits dialed went on to a ring tone and were considered legitimate mobile phone numbers. 60% of the calls to legitimate phone numbers were answered but only 8% were answered by post-2014 returnees. This translates into 12.7% of the answered phone numbers being answered by post-2014 returnees – slightly higher than the estimated share of post-2014 returnees of the total Afghan population⁴. The higher share of respondents being post-2014 returnees may reflect that Afghan returnees tend to settle in urban areas where phone signals are good⁵ and are more likely to own mobile phones than the average Afghan⁶ (ALCS, 2013/14).

Table 1: Response to RDD

Response to RDD	Digits dialed
Phone number exists but call not answered	26,703 (40%)
Lived in Afghanistan prior to 2014	34,303 (52%)
Post-2014 returnee	4,991 (8%)
Total	66,000 (100%)

Of the people reached using RDD and IVR, 65% of the post-2014 returnees and 79% of Afghans who were not post-2014 returnees expressed interest in participating in a future survey (see Table 2). The different rates of interest reflect that fewer recent returnees completed the IVR and ultimately, the RDD and IVR

⁴ The population in Afghanistan was 34.66 million in 2016 (World Bank Group Databank) and approximately 3.2 million Afghans returned from the beginning of 2014 to April 2018 (UNHCR and IOM). These numbers suggest a rough estimate of 9% of the Afghan population having returned to Afghanistan since 2014.

⁵ 37% of households who returned from displacement after 2002 lived in Kabul in 2014 compared to just 5% of non-displaced households (ALCS, 2013/2014)

⁶ 82% of households who returned from displacement after 2002 owned at least one mobile phone in 2014 compared to 65% of non-displaced households (ALCS, 2013/2014)

techniques identified 3,254 post-2014 returnees who expressed interest in the screening interview to participate in a more comprehensive survey.

Table 2: Response to IVR

Response to IVR	Non-post-2014 returnees	Post-2014 returnees
Did not complete the IVR	3,992 (12%)	1,206 (24%)
Not interested in participating in the survey	3,290 (10%)	531 (11%)
Interested in participating in the survey	27,021 (79%)	3,254 (65%)
Total	34,303 (100%)	4,991 (100%)

As the numbers suggest, a very large number of phone calls had to be made to identify eligible survey participants. To reach one post-2014 returnee 520 random digits had to be dialed and to reach someone who was also willing to participate in the survey, 800 random digits had to be dialed. In fact, just to reach any respondent 66 random digits had to be dialed and compared to other studies using RDD and IVR to identify survey participants, these numbers suggest that mobile penetration rates in Afghanistan are low. For example, L’Engle et al (2018) collected data using RDD and IVR in Ghana, a country with significantly more mobile subscriptions per capita than Afghanistan, and had to dial 23 random numbers to reach a person above 18 years and willing to participate in their survey. The difference in response rates across countries highlights the increased utility of using mobile technologies to collect data as mobile penetration rates in Afghanistan and other countries increase.

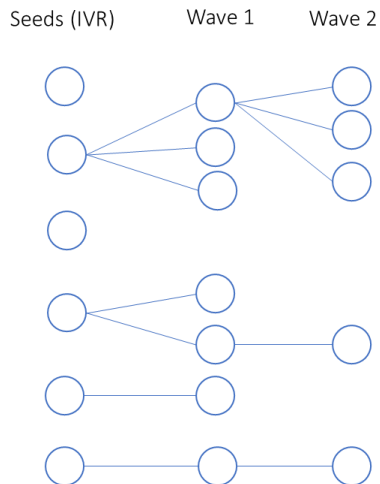
The automatization of the RDD and IVR techniques reduces the complexity of needing to dial large numbers of random digits to identify eligible survey participant and the fact that only 8% of legitimate phone numbers belonged to post-2014 returnees also emphasizes the usefulness of having an automated screening process rather than having enumerators making the screening calls. Identifying eligible survey participants using RDD and IVR was not only more time-effective than having enumerators call random digits or relying on door-to-door interviews, it also avoided placing enumerators in potentially unsafe situations caused by the on-going conflict in Afghanistan and the fact that between 65% and 75% of reached Afghans were willing to participate in a future survey, shows that the depersonalization of the screening interview (due to using IVR rather than real enumerators) did not prevent people from showing interest in the survey. It is noted that the credibility and trustworthiness of the IVR introduction can influence the participation rates. For example, using a male voice for the IVR may have constrained female respondents from participating in IVR survey. The representativeness of the data will be examined in section 5.

2. Stage 2: Respondent driven sampling

Eligible respondents (identified through RDD and IVR) who expressed interest in participating in the survey, were later called by enumerators to complete the actual interview. Prior to conducting the interview, enumerators confirmed that the respondent had in fact returned to Afghanistan after 2014 and was willing to participate in the survey (more detail on the survey design and execution is provided in Section 3). By the end of the interview, the respondents were asked how many post-returnees they knew and whether they would be willing to share phone numbers of eligible participants and participants were free to pass on as many phone numbers of eligible participants as they wanted. This method of identifying survey participants is also known as RDS – a chain-referral sampling method where survey participants share

information about other eligible participants – and was used to complement the RDD and IVR identification process. *Figure 2* illustrates the use of RDS with two waves. In this project, seeds were identified using RDD and IVR and each seed was asked to share phone numbers of other eligible survey participants etc.

Figure 2: Illustrative example of RDS



Each person interviewed provided on average 1.1 phone numbers of other post-2014 returnees and while most (72%) participants did not share any phone number, others shared more than 52 numbers of eligible participants. In addition, the enumerators got in contact with four community leaders who were each able to share between 99 and 237 phone numbers and ultimately, more than half (59%) of the total sample was identified through RDS.

RDS was an attractive way to increase the sample size in this project because returnees oftentimes group together in settlements and create networks. However, expanding one’s sample size by RDS risks biasing the sample if respondents provide phone numbers of people who resemble themselves in terms of other characteristics than simply being returnees. For example, respondents who work certain jobs might be more likely to know people in similar jobs and hence provide phone number of these people and RDS also risks oversampling individuals with large networks of 2014-returnees. The effect of RDS on the representativeness of the data was therefore tested throughout the data collection and the results of those tests will be presented in section 4.2.

3. Survey design and execution

The actual data on re-settlement of post-2014 returnees was collected through phone surveys. Typically, household survey data are collected in face-to-face interviews although the use of phone surveys has gained momentum in recent years (Dabalen et al, 2016). While phone surveys do not allow for lengthy interviews, they offer the opportunity to collect data at higher frequency and lower cost and perhaps most importantly, to collect data in remote and high-risk environments. In addition, the presence of violent anti-government groups along with low trust in official authorities can make survey participants reluctant to attract unwanted attention by participating in face-to-face interviews. Instead phone surveys offer more anonymity and, if executed well, can increase the participation rate as well as willingness to answer

sensitive questions. This section will present the design and execution of the questionnaire as well as how the collected data was cleaned.

1. Design of questionnaire

The questionnaire was intended to shed light on the re-settlement of post-2014 returnees and in particular their (i) movements and choice of location, (ii) labor market attachment and, (iii) access to services. The design of the questionnaire strived to balance the comparability to other available data sources (e.g. from the UNHCR and the IOM) and the essentials for development program design. To further ensure a broad usefulness of the survey, the questions were developed in consultation with the UNHCR in Kabul and various teams within the World Bank Group who works on project in Afghanistan.

The specific questions of interest were grouped under each of the three sections. Grouping questions according to topic makes it easier for the respondents to stay focused and tends to bring out more information for open questions. Also, the line of questions was ordered to gain the confidence of the respondents before moving on to more complicated and personal questions and after verifying that the respondents were willing to participate in the survey and had in fact returned to Afghanistan in 2014 or later, the interview started off by asking simple questions related to individual as well as household demographics. The questionnaire then continued to the sections on the physical re-settlement of returnees, their labor market attachment and finally their access to services including health related questions. The last part of the survey consisted of questions to further assess the household vulnerability, the size of the respondent's network of post-2014 returnees and whether the respondents were willing to share contact information of other eligible survey participants (for RDS).

The interview was carried out over the phone and allowed for only one household member to be interviewed. The person interviewed was typically the person who answered the phone⁷ and would act as a proxy respondent and report on behalf of other household members. The design of the questionnaire therefore had to balance the accuracy of the responses and usefulness of the collected data. The questionnaire generally addressed the household as a whole and asked questions such as "how many boys in the household attend school?" and "in the past week, did you or any household member skip a meal or eat less because there was not enough food?". Balancing was especially necessary in the labor market section of the survey where an accurate assessment of the financial situation of the household required information on the labor market attachment of the primary provider in the household. Because the primary provider of the household not necessarily was the person being interviewed, the questions in the labor market section primarily focused on the 'main bread winner' of the household.

Prior to folding out the survey, approximately 50 pilot interviews were conducted to test the questionnaire and review the answers. Additional questions were added to the questionnaire as a result of the pilot interviews including "Do all women and girls in the household have a Tazkera?" and "How did you find this job?" and the pilot interviews also showed that the process of identifying eligible participants through RDD-IVR was longer than originally anticipated. The target population was therefore extended from post-2015 to post-2014 returnees and it was agreed to supplement the identification process with RDS.

⁷ The only exceptions were in cases where the person answering the phone told the enumerator that it would be better to speak with someone else in the household

In total, eligible participants were asked between 44 and 77 questions (see questionnaire in appendix 1).

2. Execution of survey

Interviews were carried out over a period of three months from the beginning of 2018⁸. In total, 3,575 interviews were completed and the interviews took on average 26 minutes to complete. All interviews were conducted by trained enumerators and survey participants were happy to talk. Eligible survey participants identified through RDD/IVR were typically called the day after the screening interview and if an eligible participant did not answer the phone, they were called again the following day. As a rule, phone numbers were tried at least 4 times before being removed from the call lists.

All enumerators had prior experience with large-scale surveys involving returned refugees. In addition, enumerators received three days of training in using the software and understanding the questionnaire as well as two days of training on humanitarian principles from UNHCR. The enumerators were all fluent in Dari, Pashto, and Uzbek and initially addressed the respondent in the language that the respondent chose during the IVR survey (although they could change to a different language if necessary) and while male respondents were contacted by either male or female enumerators, all calls to female respondents (as identified through the IVR survey) were routed to female enumerators.

To ease the data collection, the survey software (Open Data Kit) included filters and automatically skipped sections. This ensured that rather than having to do subsequent interviews to clarify inconsistencies, the software carried out several embedded quality checks during the interview to detect inconsistent responses. For example, if a respondent said they lived in Herat province then the embedded filter ensured that the enumerator saw a list of districts in Herat only. Also, if a respondent indicated that no household member had visited a health facility in the last three months, then the questionnaire automatically continued to the next set of questions. Hence the software ensured that the data was cleaned automatically by the enumerators while collecting the data⁹.

Enumerators were instructed to ask open questions and code the responses based on the answer given. Asking open questions with multiple response options in a phone survey avoids confusing the respondents who otherwise would have to memorize numerous response options. Moreover, open questions allow us to learn things that were not expected prior to collecting the data. For example, “No difficulties” was added as a response category to the questions “What is the main difficulty you face in accessing health care?” and “No income/living from savings” was added as response category to the questions regarding income (“What is the main source of income for your household?” and “What was the main source of income for your household before you returned to Afghanistan?”). If the answers did not seem consistent or were too broad, the enumerators were trained to ask respondents to clarify their answers. For example, if a respondent indicated that a household member had visited a health facility in the last three months, then the enumerator asked which services they used. The response options usually did not match directly with what the respondents said – a respondent might say that a family member received treatment for diarrhea and in this case, the enumerator would follow up and ask whether the person receiving treatment was an adult or child, to determine the most appropriate response. To ensure that feasible categories were used

⁸ The pilot interviews were carried out from January 7th, 2018 to January 9th, 2018 and the actual survey was conducted from January 15th, 2018 to April 28th, 2018

⁹ Throughout the data collection the response option “Other” was reviewed and re-classified if possible.

to classify the responses, the responses were manually inspected over the course of the data collection and common patterns of responses were coded as standardized response options.

3. Response rates of phone interview

To achieve 3,575 completed interviews, enumerators made a total of 9,247 phone calls¹⁰. About 53% of these phone numbers were collected through IVR and 47% were collected through respondent-driven sampling (RDS) (see **Error! Reference source not found.**). Hence enumerators only had to make 2.6 calls in order to complete one interview with a 2014-returnee.

Overall, 93.5% of post-2014 returnees called by enumerators agreed to be surveyed (either at the moment of the initial call or when they were called back later). Looking more closely at the response rates, *Table 3* shows that of the 9,247 phone calls that were made, 5,624 (61%) of the calls were answered and the response rate did not differ between eligible participants identified through IVR or RDS. Of the calls that were answered, 75% were post-2014 who consented to be interviewed. The rate of consent was 20 percentage points higher among calls made using phone numbers identified through RDS than IVR. This difference is attributed to the fact that 25% of the respondents identified through IVR turned out not to be returnees while this was only the case for 5% of the respondents identified through RDS. The same pattern is found in terms of year of return to Afghanistan where 25% and 6% of returnees identified through IVR and RDS, respectively had returned prior to 2014. These rates support the hypothesis that returnees tend to establish networks upon return to Afghanistan, but the rates also highlight the importance of basic technological skills to operate mobile phones in the target population when relying on IVR technology for data collection.

Table 3: Response tree – phone interview

Response to interview call	IVR	RDS	Total
Phone call not answered	1,996 (40%)	1,627 (38%)	3,623 (39%)
<i>Phone call is answered</i>	<i>2,948 (60%)</i>	<i>2,676 (62%)</i>	<i>5,624 (61%)</i>
Not a returnee	740 (25%)	137 (5%)	877 (16%)
Declined to be interviewed	127 (4%)	68 (3%)	195 (3%)
Asked to be called back later	155 (5%)	193 (7%)	348 (6%)
<i>Consented to be interviewed</i>	<i>1,926 (65%)</i>	<i>2,278 (85%)</i>	<i>4,204 (75%)</i>
Pre-2014	472 (25%)	141 (6%)	613 (15%)
Post-2014			
(uncompleted)	6 (0%)	10 (0%)	16 (0%)
Post-2014			
<i>(completed)</i>	<i>1,448 (75%)</i>	<i>2,127 (93%)</i>	<i>3,575 (85%)</i>

Ultimately 3,575 phone interviews were completed and 41% of those were identified through RDD/IVR and 59% were identified through RDS.

¹⁰ Numbers were called multiple times if not responded to

4. Cost of the survey

One of the advantages of using mobile technology to identify and interview eligible survey respondents is the costs which are typically significantly lower than collecting data through face-to-face interviews. The total cost of this data collection project was \$100k of which \$60k was used to identify eligible survey participants using IVR/RDD and \$40k was used to conduct surveys. The average cost of identifying one post-2014 returnee who expressed interest in participating in the survey using RDD-IVR was therefore \$18¹¹ and the average cost of one completed interview was \$28¹².

4. Effects of applied sampling methods

RDD-IVR and RDS methods were used to identify eligible survey participants. This section breaks down the response rates of the identification of eligible participants and discuss the implications hereof. The section starts by comparing IVR respondents who were willing to participate in the survey with the ones who were not. It then analyzes how RDS affected the composition of the sample.

1. IVR selection

The IVR served as a screening tool to identify post-2014 returnees who were willing to participate in the survey. 14%¹³ of the eligible participants, who completed the IVR interview, were not interested in participating in the survey (see *Table 4*) and although the breakup of the response rates by demographic characteristics shows that the IVR respondents who were willing to participate in the survey generally resembled the ones who were not, a few differences stand out. The breakup on gender shows that twice as many women (24%) as men (12%) were unwilling to participate in the survey and while it is not possible to determine the reason why this is the case, it is likely due to cultural norms or, as previously discussed, because the IVR was carried out using a male voice. The breakup also shows that individuals whose highest completed education was religious or apprenticeship as well as individuals settled in the West-Central regions of Afghanistan tended to be less interested in participating in the survey. It is noted, however, that the latter groups make up small shares of the total survey.

¹¹ \$60k/3,254=\$18.4

¹² \$100/3,575=\$28.0

¹³ 531/(3,254+531)*100=14%

Table 4: Response to IVR by characteristics

Willing to participate in phone survey		Not willing	Willing	Total
	Total	14%	86%	(3,785) 100%
Gender	Males	12%	88%	(3,214) 100%
	Females	24%	76%	(571) 100%
Education*	No education	14%	86%	(1,066) 100%
	Primary	16%	84%	(569) 100%
	Secondary	13%	87%	(437) 100%
	High School	10%	90%	(731) 100%
	University	11%	89%	(620) 100%
	Religious	21%	79%	(99) 100%
	Apprenticeship	27%	73%	(75) 100%
Location prior to return	Pakistan	11%	89%	(1,902) 100%
	Iran	16%	84%	(1,192) 100%
	Other	17%	83%	(691) 100%
Current region	Central	14%	86%	(1,568) 100%
	East	14%	86%	(643) 100%
	North	13%	87%	(654) 100%
	Northeast	16%	84%	(129) 100%
	South	12%	88%	(217) 100%
	Southwest	13%	87%	(142) 100%
	West	17%	83%	(111) 100%
	West-central	4%	96%	(55) 100%
	Not sure	21%	79%	(266) 100%
Year of return	2014	14%	86%	(961) 100%
	2015	14%	86%	(702) 100%
	2016	13%	87%	(999) 100%
	2017	13%	87%	(737) 100%
	2018	18%	82%	(386) 100%

*188 individuals did not indicate their level of education. The ordering of the IVR questions was adjusted after the pilot

2. RDS selection

The identification of eligible survey participants was supplemented by RDS – a chain-referral sampling method where survey participants refer other eligible survey participants. RDS is especially useful to recruit populations that are difficult to reach. However, it is a non-random sampling method that risks biasing the sample if people tend to refer people who resemble themselves (also known as homophily) and groups with larger networks are likely to be oversampled. To ensure that the RDS did not bias the sample, RDS-estimators were surveilled throughout the data collection. RDS estimators adjust for network and

homophily bias by assigning smaller weights to individual who have large networks and possess similar characteristics as their recruiter.

RDS¹⁴ was the source of identification of 2,127 (59%) survey participants and an additional 1,325 eligible participants¹⁵ were nominated but never recruited for the survey. Survey participants were free to pass on as many phone numbers as they wanted and although some participants were referred multiple times, they were only interviewed one time¹⁶ (sampling was without replacement). The average network size of post-2014 returnees, among respondents with a network (40% of the sample), was 10.5¹⁷ and each survey respondent provided on average 1.1 phone numbers of other potential survey participants. While most respondents (95%) did not recruit new participants, others provided up to 52 phone numbers and recruited 31 respondents. In addition, enumerators reached four community elders who provided lists of 681 post-2014 returnees which ultimately lead to 408 interviews.¹⁸

Respondent Driven Analyses and homophily scores

The impact of RDS on the representativeness of the sample was assessed by comparing unweighted and RDS-weighted sample proportions and homophily scores were calculated (see *Table 5*). The weighted sample proportions were based on RDS-II¹⁹. Homophily scores measures the extent to which participants recruit individuals who resembles themselves rather than at random. A homophily score of one indicates that RDS participants are recruited at random while the homophily score for e.g. males of 1.3 indicates that 30% more men than women were sampled using RDS compared to the expected number if the sampling had been random.

The estimation results show no evidence of non-random recruitment by gender, age, education and job type and the unweighted sample proportions fall within the 95% confidence interval of the RDS-weighted sample proportions for all groups. These findings are reiterated by the low homophily scores indicating that participants did not systematically recruit other participants who resembled themselves. However, the results also show that RDS may have increased the share of returnees living in Nangarhar and returning to Afghanistan in 2016. This is echoed by the high homophily scores of living location. For example, respondents living in Nangarhar were 3.4 times more likely to recruit other returnees in Nangarhar than if the sampling had been random which suggests that RDS biased the sample towards certain geographical locations and in particular Nangarhar.

¹⁴ The RDS was initiated after approximately two weeks of interviews

¹⁵ In total, 4,003 phones numbers were shared by survey participants. 442 individuals were nominated between two and up to six times.

¹⁶ In the RDS analysis, survey respondents identified through RDS were linked to the individual first referring to them

¹⁷ 23% of the survey participants were not able to estimate their network and their network size was imputed using the median network size of all households corresponding to 1. The median network size of respondents with a network was 5.

¹⁸In addition, 103 phone numbers were shared and 65 recruitments were made by pre-2014 returnees who did not participate in the survey

¹⁹ RDS-II does not require an estimate of individual network size but assumes sampling without replacement as opposed to the successive sampling estimator RDS-SS. The RDS-SS estimator requires an estimate of a total population of post-2014 returnee households and an estimate of 200,000 was used.

Table 5: Population proportions, sample proportions, and RDS-II estimates with 95% confidence intervals (CIs)

		Full sample (unweighted)	RDS-II		Homophily scores
Variables			Estimate	95% CI	
Gender	Males	0.94	0.93	(90.92 :0.94)	1.30
	Females	0.06	0.07	(0.06:0.08)	-
Education	No education	0.48	0.46	(0.44:0.49)	1.19
	Religious	0.04	0.04	(0.03:0.05)	1.10
	Primary	0.25	0.25	(0.23:0.27)	1.09
	Secondary	0.13	0.14	(0.13:0.16)	1.26
	More than secondary	0.10	0.10	(0.09:0.12)	1.51
Job type	Daily wage non-agriculture	0.54	0.53	(0.50:0.55)	1.22
	Self-employed non-agriculture	0.20	0.21	(0.19:0.23)	1.12
	Salaried private sector	0.12	0.12	(0.10:0.13)	1.32
	Self-employed agriculture	0.07	0.08	(0.06:0.09)	1.44
	Salaried public sector	0.05	0.06	(0.05:0.07)	0.99
	Daily wage agriculture	<0.01	<0.01	-	0.54
	Employer	<0.01	<0.01	-	-
	Other	<0.01	<0.01	-	-
	Non-paid	<0.01	<0.01	-	-
	None	<0.01	<0.01	-	-
Current province	Nangarhar	0.47	0.43	(0.41:0.45)	3.42
	Kabul	0.17	0.18	(0.16:0.20)	2.96
	Other	0.37	0.39	(0.37:0.41)	3.43
Year of return	2014	0.06	0.07	(0.06:0.08)	1.34
	2015	0.10	0.11	(0.09:0.12)	1.44
	2016	0.50	0.46	(0.43:0.49)	1.30
	2017	0.30	0.32	(0.30:0.34)	1.24
	2018	0.04	0.05	(0.03:0.06)	1.20

*Highlighted in bold if unweighted sample share is significantly different from RDS-II weighted shares

In summary, RDS was able to successfully identify more than half of the survey participants and while RDS had otherwise limited effects on potential sample bias, it oversampled individuals from Nangarhar. If returnees living in Nangarhar are systematically different from returnees living in other provinces, this may bias the sample.

3. Discussion of sampling methods

The comparison of rates of interest in participating in the survey from the IVR, shows that respondents who expressed interest in participating in a survey generally resembled respondents who were not interested. The exception to the rule were women who tended to be less willing to participate in a future survey. This however is not surprising as women typically are hard to reach in Afghanistan, and mobile ownership tends

to be concentrated among men. Moreover, because the survey focuses on household level characteristics, women who share their household with one or more men will indirectly be represented in the survey.

RDS on other hand contributes to an oversampling of respondents from Nangarhar. The over-sampling of respondents from Nangarhar is a concern if returnees living in Nangarhar are systemically different than other returnees so to correct for the possible bias, survey weights are suggested (for more details see section **Error! Reference source not found.**). It is possible, that the over-sampling could have been avoided had there been an upper limit to how many phone numbers each respondent could pass along. For example, the four community elders who recruited 408 of the survey respondents all lived in Nangarhar and 88% of their recruitments also lived in Nangarhar. In comparison, only 20% of individuals identified through IVR lived in Nangarhar which resembles the number presented by the UNHCR²⁰. Restricting the number of recruitments to e.g. 4 or 5 could have allowed using RDS without biasing the sample towards – in this case – certain geographical locations.

5. Representativeness

Representativeness of the survey is important because it ensures that all types of post-2014 returnee households are included in the sample but relying on mobile technology to identify and interview post-2014 returnees implicitly excludes individuals without mobile phones and living in areas with poor mobile signal from participating in the survey. The survey sample is therefore ideally representative of returnees with mobile phones and The Asia Foundation (2018) find that mobile phones tend to be owned by the young, the educated, men, and urban residents.

This section examines the representativeness of survey sample. This section starts by discussing how to approximate the true population of Afghan returnee in lack of representative data sources. The representativeness of sample is then assessed by comparing key characteristics of returnees in the phone survey to key characteristics of the approximated true population of returnees^{21 22}. The comparisons are performed separately for documented and undocumented returnees and first, a preliminary assessment of systematic differences based on bivariate comparisons of the phone survey and the true population of returnees are provided. Bivariate tests do not account for intercorrelation of demographic characteristics and the analysis is therefore extended to multivariate comparisons of the two populations. Participation in the phone survey is the dependent variable in both the bivariate and multivariate analyses.

1. The true population of Afghan returnees

The representativeness of the sample is assessed by comparing characteristics of respondents of the survey to characteristics of full population of recent returnees in Afghanistan. The lack of representative information on returnees after they have returned to Afghanistan implies that there is no data available to characterize the true population of post-2014 returnees. Instead, the true population is approximated by

²⁰ Approximately 27% of Afghans returning to Afghanistan between Jan 1st, 2014 and April 30th, 2018 intended to return to Nangarhar according to data collected by the UNHCR upon return at the Encashment center (Afghan Voluntary Repatriation database)

²¹ Out of a total sample of 3,575 interviews, 1,798 of were undocumented returnees and 1,777 were documented returnees.

²² The phone survey in this analysis only includes respondents who were the main bread winner of the household and who settled in their province of origin

combining information on all Afghan refugees living in Pakistan in 2010 with UNHCR registries on returnees. This section describes the data sources used to approximate the true population of post-2014 returnees and how the individual-level data is processed to resemble the household structure of the phone survey data collected by the World Bank Group.

The target population of the survey is post-2014 returnees and the UNHCR and the IOM estimate that approximately 3.2 million documented and undocumented displaced Afghans returned to Afghanistan from January 2014 to April 2018. The majority of Afghans returning to Afghanistan came from either Iran or Pakistan and the core characteristics of the full population of Afghan returnees is approximated by the *Afghan Population Profiling, Verification and Response survey* (PPVR). The PPVR is a census-like survey conducted by the Government of Pakistan and the United Nations High Commissioner for Refugees (UNHCR) from November 2010 to 30th December 2011. It includes detailed information on individual and household characteristics of all Afghans who lived in Pakistan at the time of the survey. In total, 975,000 individuals were interviewed and 507,000 of those were documented refugees for whom the proof of registration (PoR) number was recorded. The subset of Afghans living in Pakistan in 2011 and returning to Afghanistan in 2015-2017 were identified in the *Voluntary Return Form survey* (VRF). The VRF is collected by the UNHCR and includes the PoR number of documented returnees who pass through one of their Voluntary Repatriation Centers.

To assess the representativeness of the phone survey data, core characteristics of the documented returnees were compared to the core characteristics of the subset of documented refugees living in Pakistan in 2011 (in the PPVR database) who returned to Afghanistan in 2015-2017. The return of undocumented refugees living in Pakistan in 2011 cannot be identified. Instead, all undocumented refugees living in Pakistan in 2011 were used as the comparison group to assess the representativeness of undocumented returnees in the phone survey. To ensure that the phone survey respondents and PPVR respondents are comparable, only phone survey respondents who returned in 2015-2017 were included in the comparison analysis. See Box 1.

Box 1: Comparison groups to assess the representativeness of the phone survey

Phone survey	PPVR and VRF
Documented Afghans who returned to Afghanistan in 2015-2017 Main bread winners only	Documented Afghans who lived in Pakistan in 2011 and returned to Afghanistan in 2015-2017
Undocumented Afghan who returned to Afghanistan in 2015-2017 Main bread	Undocumented Afghans who lived in Pakistan in 2011

The phone survey is a household-level dataset while the PPVR includes data on all household members and to make the two datasets comparable, only interviews with main bread winners in the phone survey were included in the analysis. In total, 2,643 interviews were conducted with respondents who were also the main bread winner of the household and 49% of those were documented returnees. The PPVR dataset was transformed into two household-level datasets; a *simple* transformation where the household representative typically is the head of household as the household representative and a more *advanced* transformation where the household member with the best job and highest education typically represents the household. *Table 7* and *Table 7* show the steps taken to transform the PPVR dataset into household-level datasets. The transformed PPVR data included 135,346 households of which 12% were documented refugees who returned to Afghanistan in 2015-17, 43% were documented refugees who did not return in 2015-2017 and 45% were undocumented refugees.

Table 6: Steps in turning PPVR into a household-level dataset – v.1 (Simple)

<i>Full sample:</i>	970,344 individuals 463,328 (48%) are undocumented 507,016 (52%) are documented
Step 1: Age* Keep individuals older than 13 years	- <u>208,424 individuals</u> 761,920 individuals
Step 2: Household head aged 18-50 Keep household head between 18 and 50 years	- <u>274,347 individuals</u> 487,573 individuals
Step 3: Employed aged 18-50 Keep (random) employed individuals between 18 and 50 years	<u>264,538 individuals</u> 223,035 individuals
Step 4: Gender Keep males	- <u>55,247 individuals</u> 167,788 individuals
Step 5: Age Keep individuals aged 18-50	- <u>19,409 individuals</u> 148,379 individuals
Step 6: Age - oldest Keep oldest if younger than 40 years and youngest if older than 40 years	- <u>12,975 individuals</u> 135,404 individuals
Step 7: Random Randomly dropping observations if more than one observation per household	- <u>58 individuals</u> 135,346 individuals
<i>Final sample:</i>	135,346 households 61,549 (45%) are undocumented & 73,797 (55%) are documented

* Age is right-side winsorized at the 99.95 percentile

Table 7: Steps in turning PPVR into a household-level dataset- v.2 (Advanced)

<i>Full sample:</i>	970,344 individuals 463,328 (48%) are undocumented 507,016 (52%) are documented
Step 1: Age* Keep individuals older than 13 years	- <u>208,424 individuals</u> 761,920 individuals
Step 2: Employment Keep employed individuals (if at least one household member is employed)	- <u>542,899 individuals</u> 252,631 individuals
Step 3: Job type Keep individuals with the best type of jobs	<u>46,787 individuals</u> 205,844 individuals
Step 4: Education Keep individuals with the highest level of attained education	- <u>31,194 individuals</u> 174,650 individuals
Step 5: Head of household Keep only head of household if there is one	- <u>28,380 individuals</u> 146,270 individuals
Step 6: Age Keep oldest if younger than 40 years and youngest if older than 40 years	- <u>10,808 individuals</u> 135,467 individuals
Step 7: Gender Keep if male	- <u>30 individuals</u> 135,437 individuals
Step 8: Random Randomly dropping observations if more than one observation per hh	- <u>91 individuals</u> 135,346 individuals
<i>Final sample:</i>	135,346 households 60,725 (45%) are undocumented & 74,621 (55%) are documented

* Age is right-side winsorized at the 99.95 percentile

Finally, two combined datasets of the phone survey and the PPVR dataset were constructed. The datasets

Variable	Categories	Mean			Test values	
		WBPS*	PPVR v.1	PPVR v.2	Chi2 v.1	Chi2 v.2
Age		32.40	36.14	42.12	***	***
Household size		9.23	7.00	7.00	***	***
Share of children (0-16y) in household		0.56	0.51	0.50	***	***
Gender	Male	99.2%	93.6%	92.8%	65.7***	77.2***
	Female	0.8%	6.4%	7.2%		
Language	Pashto	89.3%	88.2%	88.1%	1.3	1.4
	Other	10.7%	11.8%	11.9%		
Language	Pashto	89.3%	88.2%	88.1%	68.1***	92.9***
	Dari	10.7%	7.7%	7.7%		
	Other	0.0%	4.1%	4.1%		
Job type in Pakistan/prior to returning to Afghanistan	Unemployed/unpaid work	5.1%	19.1%	7.8%	229.9***	92.9***
	Day laborer	52.6%	45.6%	44.9%		
	Salaried	11.8%	16.0%	21.6%		
	Own account/employer	30.5%	19.4%	25.7%		
Province of origin*	Nangarhar	60.2%	23.9%	23.8%	594.6***	600.4***
	Kabul	5.3%	11.8%	11.8%		
	Other	34.6%	64.3%	64.4%		
Education	No schooling	48.5%	59.7%	58.2%	121.6***	95.1***
	Religious	4.0%	5.0%	5.9%		
	Primary	26.8%	18.5%	18.2%		
	Middle or secondary	11.6%	12.2%	11.8%		
	Tertiary, college or university	9.1%	4.6%	5.8%		

Note: Order of completed education is “No education”, “Religious”, “Primary”, “Middle/secondary” and “Technical/college/university”. “Don’t know/Other” education in PPVR is categorized as no education. Pearson’s chi-square test is conducted on categorical indicators and Mann-Whitney tests are conducted on ordered indicators. ***p<0.01, **p<0.05, *p<0.1. Only main bread winner in families returning to Afghanistan in 2015-2017 are included in the sample. Province of origin only exists for individuals living in their province of origin – in the model specifications above, only individuals living in their province of origin are therefore included.

included core characteristics such as gender, age, household size, educational background and province of origin as well as a variable indicating whether the individual was a participant in the phone survey or the PPVR.

2. Bivariate test of representativeness

The bivariate estimation results indicate that phone survey participants are not representative of the true population of post-2014 returnees (see

Table 8). Instead, the differences between the survey and the true population of returnees (approximated by two different sub-samples of the PPVR) generally align with typically observed characteristics of phone survey respondents

Households participating in the phone survey were significantly larger and the share of children significantly bigger than in the true population of Afghan returnee households. The larger household size may on the one hand reflect an over-sampling of large households for example because large households include more

people who may own a phone. On the other hand, it may also reflect that many returnees re-settle with family members and hence the average household size increases after returning to Afghanistan.

The estimation results also show that survey participants were significantly younger and more likely to be male than the household representatives in the true population of returnees. These findings reiterate the general idea that phone surveys tend to be biased towards young (and possibly more tech-savvy) males. Phone survey participants were also significantly more likely to have an education and origin from (and live in) Nangarhar province. Under the assumption that education is positively correlated with income, these findings are consistent with needing to own and be able to operate a mobile phone to participate in the phone survey. Moreover, although it is unexpected that the share of survey participants living in Nangarhar is 2½ times larger than the true share (which is influenced by the RDS method used), it is not surprising that individuals living in urban areas, which tend to have better mobile coverage, are over-represented in the phone survey.

The results furthermore show that phone survey participants were less likely to have salaried employment prior to returning to Afghanistan. This finding is surprising as salaried employees generally are expected to have better-paying and more stable jobs and hence be more likely to own mobile phones. The RDD was set-up to call every day between 7.30 AM and 9.00 PM and did not have the options for respondents to re-schedule the screening interview. If for example salaried employees had less opportunities to answer their phones during work hours compared to other types of employment, this may explain the relatively low share of salaried employees in the phone survey. Alternatively, survey respondents may have an incentive to misreport their income and employment status in the hope of attracting aid. Only the share of Pashto native speakers in the phone survey resembled the true share of Pashto speakers among returnees.

Table 8: Demographic characteristics and Bivariate Comparisons of 2015-2017 **documented returnees** – comparing the phone survey with two different sub-samples of the PPVR dataset.

Variable	Categories	Mean			Test values	
		WBPS*	PPVR v.1	PPVR v.2	Chi2 v.1	Chi2 v.2
Age		32.40	36.14	42.12	***	***
Household size		9.23	7.00	7.00	***	***
Share of children (0-16y) in household		0.56	0.51	0.50	***	***
Gender	Male	99.2%	93.6%	92.8%	65.7***	77.2***
	Female	0.8%	6.4%	7.2%		
Language	Pashto	89.3%	88.2%	88.1%	1.3	1.4
	Other	10.7%	11.8%	11.9%		
Language	Pashto	89.3%	88.2%	88.1%	68.1***	92.9***
	Dari	10.7%	7.7%	7.7%		
	Other	0.0%	4.1%	4.1%		
Job type in Pakistan/prior to returning to Afghanistan	Unemployed/unpaid work	5.1%	19.1%	7.8%	229.9***	92.9***
	Day laborer	52.6%	45.6%	44.9%		
	Salaried	11.8%	16.0%	21.6%		
	Own account/employer	30.5%	19.4%	25.7%		
Province of origin*	Nangarhar	60.2%	23.9%	23.8%	594.6***	600.4***
	Kabul	5.3%	11.8%	11.8%		
	Other	34.6%	64.3%	64.4%		
Education	No schooling	48.5%	59.7%	58.2%	121.6***	95.1***
	Religious	4.0%	5.0%	5.9%		
	Primary	26.8%	18.5%	18.2%		
	Middle or secondary	11.6%	12.2%	11.8%		
	Tertiary, college or university	9.1%	4.6%	5.8%		

Note: Order of completed education is “No education”, “Religious”, “Primary”, “Middle/secondary” and “Technical/college/university”. “Don’t know/Other” education in PPVR is categorized as no education. Pearson’s chi-square test is conducted on categorical indicators and Mann-Whitney tests are conducted on ordered indicators. ***p<0.01, **p<0.05, *p<0.1. Only main bread winner in families returning to Afghanistan in 2015-2017 are included in the sample. Province of origin only exists for individuals living in their province of origin – in the model specifications above, only individuals living in their province of origin are therefore included.

3. Multivariate test of representativeness

The analysis of representativeness was extended to multivariate comparisons to account for intercorrelation of demographic characteristics and various Probit model specifications were estimated using the two sub-samples of the PPVR and the phone survey (see Table 9). The results reiterate the findings from the bivariate analyses and suggest that the phone survey is not representative of the true population of returnees but generally differ in ways that are typical for phone surveys.

The multivariate estimation results, like the bivariate results, show that households participating the phone survey were significantly larger and the share of children bigger than in the true population of returnees. Phone survey participants were similarly more likely to be male and significantly younger than the household representatives of the true population of returnees. The multivariate results also echo that significantly more phone survey respondents were from Nangarhar and work as own account/employers or day laborers.

Overall, the estimation results also show that returnees with higher education were more likely to participate in the phone survey. However, the size and significance of the effects depend on which sub-sample of the PPVR the phone survey is compared to. Comparing the phone survey with the simple sub-sample of the PPVR generates larger and more significant coefficients than comparing the phone survey with the more advanced sub-sample of the PPVR. These differences reflect the selection of the household representative of the PPVR data when transforming it into a household level dataset and highlight the sensitivity of the comparisons to the selection of the true population of returnees. The multivariate analyses also show that the relative effects of employment on the probability of participating in the survey are notable larger in the simple PPVR sample selection than in the more advanced sample selection. The relativeness of the marginal effects suggests that undocumented returnees in the phone survey resemble the true population of returnees better than documented returnees in the phone survey.

Table 9: Demographic characteristics and Multivariate Comparisons, Marginal effects Probit models – Documented returnees

VARIABLES	(1) Documented Simple (1)	(2) Undocumented Simple (1)	(3) Documented Adv (1)	(4) Undocumented Adv (1)
Age of the respondent	-0.00171*** (0.000129)	-0.000816*** (3.87e-05)	-0.00217*** (0.000103)	-0.000704*** (3.38e-05)
Gender of the respondent (0:Female, 1:Male)	0.0266*** (0.00231)	0.00688*** (0.000483)	0.0221*** (0.00166)	0.00510*** (0.000364)
job_type==Day labourer	0.0704*** (0.00619)	0.0230*** (0.00189)	0.0176*** (0.00503)	0.00556*** (0.00127)
job_type==Salaried worker	0.0551*** (0.0110)	0.0177*** (0.00340)	-0.0102*** (0.00394)	-0.00348*** (0.000856)
job_type==Own account/Employer	0.127*** (0.0138)	0.0500*** (0.00596)	0.0212*** (0.00672)	0.00512*** (0.00174)
Share of children(0-16y) in hh	0.0373*** (0.00628)	0.0112*** (0.00144)	0.0322*** (0.00501)	0.00715*** (0.00106)
Household size	0.00275*** (0.000265)	0.000472*** (5.36e-05)	0.00284*** (0.000235)	0.000507*** (4.49e-05)
province==Kabul	-0.00711* (0.00379)	-0.00448*** (0.000706)	-0.00514* (0.00306)	-0.00316*** (0.000518)
province==Nangarhar	0.0643*** (0.00454)	0.00827*** (0.000866)	0.0529*** (0.00402)	0.00608*** (0.000685)
education_psm==No education	-0.0374*** (0.00599)	-0.00984*** (0.00162)	-0.0191*** (0.00434)	-0.00408*** (0.00101)
education_psm==Religious	-0.0203*** (0.00300)	-0.00492*** (0.000695)	-0.0154*** (0.00241)	-0.00383*** (0.000451)
education_psm==Primary	-0.0199*** (0.00327)	-0.00362*** (0.000817)	-0.0147*** (0.00260)	-0.00293*** (0.000565)
education_psm==Middle/secondary	-0.0213*** (0.00291)	-0.00477*** (0.000707)	-0.0175*** (0.00214)	-0.00371*** (0.000479)
Observations	16,800	62,267	17,133	61,559

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: Excluded binary variables are Technical/college/tertiary, Unpaid/family worker, Other province than Kabul and Nangarhar due to multicollinearity

Overall assessment of the representativeness of the phone survey

Collecting survey data using mobile technology restricts individuals who do not own a phone, live in an area with poor phone signal or who are not able to operate a mobile phone and the assessment of representativeness of the phone survey show, as expected, that individuals who tend to be over-represented in phone surveys are also over-represented in this survey.

6. Adjusting for survey bias: Survey weights

A survey is ideally a representative sub-set of the population it was collected from. However, for various reasons this may not always be the case and instead survey weights can be applied to account for sample biases. The WBPS is ideally representative of post-2014 returnees who own mobile phones but as shown in section 4.2, RDS caused an over-sampling of individuals from Nangarhar. This section discusses and compares two sets of suggested survey weights for the WBPS to account for regional imbalances.

Survey weights indicate how much each observation should count in statistical computations and are typically computed based on demographic and geographic characteristics from the full population or a representative sample of the target population. Computing survey weights for the WBPS was challenged by a lack of representative data of returnees in Afghanistan and two sets of survey weights were computed to account for the geographical imbalance of the WBPS (see **Box 2** for more information on the construction of the two sets of weights). The two sets of weights rely on different assumptions: the first set of weights assumes that the IVR sample represents the true regional distribution of returnees whereas the second set of weights assumes that documented returnees' intended destination upon return at the Encashment center represents the true regional distribution of returnees. Moreover, both sets of survey weights rely on the assumption that post-2014 returnees own mobile phones at the same rate as the average Afghan living in their province of settlement.

Box 2: Construction of survey weights

$$\text{Survey Weights 1} = \frac{\text{IVR shares}_i * \#\text{Returnees}_i * \text{Mobile owners}_i}{\#\text{Respondents in WBPS}_i}$$

$$\text{Survey Weights 2} = \frac{\text{UNHCR shares}_i * \#\text{Returnees}_i * \text{Mobile owners}_i}{\#\text{Respondents in WBPS}_i},$$

where i =province, IVR shares_i =share of returnees by province (based on the IVR sample), Mobile owners =share of Afghans owning a mobile phone (based on ALCS 2013/14), $\#\text{Returnees}$ =Number of post-2014 returnees (based on UNHCR and IOM), $\#\text{Respondents in WBPS}$ =Number of respondents in WBPS (based on WBPS) and UNHCR shares_i =share of documented returnees by province (based on UNHCR)

To illustrate the effects of applying the two sets of survey weights, the distribution of selected characteristics is compared in *Table 10*. The survey weights are constructed to account for regional biases and as expected, the comparisons in *Table 10* show significant differences between the unweighted share of post-2014 returnees living in Nangarhar (47%) and the weighted shares (15% and 20%, respectively). The results also show that applying the survey weights tend to influence the distribution of other characteristics of post-2014 returnees but that the relative distribution remains unchanged. Choosing the best set of weights depends on which assumptions are assessed to be most credible.

Table 10: Characteristics of WBPS responses – weighted and unweighted

	Unweighted	Survey weights 1	Survey weights 2
Province of settlement			
Nangarhar	47%	15%	20%
Kabul	17%	31%	28%
Other	36%	54%	53%
Is household living in province of origin?			
Yes	71%	64%	68%
Education of bread winner			
No education	51%	46%	46%
Primary school completed	25%	26%	25%
Middle or secondary school completed	12%	13%	13%
Technical/college/university completed	9%	10%	12%
Religious school	4%	4%	5%
Other/Don't know	0%	0%	0%
Is bread winner employed?			
Yes	89%	87%	88%
Job type of bread winner			
Daily wage labor, non-agriculture	54%	48%	48%
Self-employed, non-agriculture	20%	24%	23%
Salaried, private sector	12%	13%	13%
Self-employed, agriculture	7%	8%	8%
Salaried, public sector	5%	7%	7%
Other	1%	1%	1%
Income of bread winner			
Less than 1,500 Afs	1%	1%	1%
1,500 - 3,000 Afs	6%	5%	5%
3,001 - 5,000 Afs	28%	23%	24%
5,001 - 10,000 Afs	49%	51%	52%
More than 10,000 Afs	15%	19%	18%

Note: Staff calculations based on the WBPS. Weights are calculated based on the WBPS, UNHCR, IOM and ALCS 2013/14

7. Conclusion

Most household survey data are collected through face-to-face interviews which can be both costly and time-consuming and is oftentimes not feasible in high-risk environments such as Afghanistan. Moreover, a high-risk environment combined with limited official data further complicate the identification of eligible survey participants. This note suggests using RDD and IVR to identify eligible survey participants and conduct household surveys over the phone. While this method is likely to bias the sample towards respondents who are more educated, living in urban areas and better off financially, the analysis also shows that the participation of poor household is high. Collecting data using mobile devices rather than face-to-face interviews also allows collecting data quickly often necessary for programs responding to sudden changes e.g. natural disasters or deterioration of political stability.

Sources

ALCS 2013/14: *Afghanistan Living Condition Survey* (previously known as National Risk and Vulnerability Assessment) conducted by the Central Statistics Organization of Afghanistan

Dabalen, A., Etang, A., Hoogeveen, J., Mushi, E., Schipper, Y. and von Engelhardt, J, 2016: *Mobile Phone Panel Surveys in Developing Countries: A Practical Guide for Microdata Collection*, Directions in Development--Poverty. Washington, DC, World Bank.

IOM, International Organization for Migration, *Return of Undocumented Afghans*. Based on last weekly update (in December) for 2014, 2015, 2016 and 2017 and last weekly update for April 2018

L'Engle K., Sefa E., Adimazoya E. A., Yartey E., Lenzi R., Tarpo C., et al., 2018: *Survey research with a random digit dial national mobile phone sample in Ghana: Methods and sample quality*. PLoS ONE 13(1): e0190902.

NRVA 2007/08: *National Risk and Vulnerability Assessment* conducted by the Central Statistics Organization of Afghanistan

The Asia Foundation, 2018: *"A Survey of the Afghan People"*. Afghanistan in 2017.

UNHCR, United Nations High Commissioner for Refugees, Afghan Voluntary Repatriation database

World Bank Group Databank, Health Nutrition and Population Statistics, Total population, Afghanistan, 2016. URL: <http://databank.worldbank.org/data/source/health-nutrition-and-population-statistics>

Appendix 1

Questionnaire used for Phone survey interviews

Presentation
Do you agree to be interviewed?
What is your gender?
In what year did you return to Afghanistan?
What was the season when you returned?
How old are you?
How many people (including men and women, boys and girls, and yourself), live in your household?
How many baby BOYS age 4 years or younger are in your household?
How many BOYS age 5-16 are in your household?
How many of the BOYS in your household age 5-16 are attending school?
How many baby GIRLS age 4 years or younger are in your household?
How many GIRLS age 5-16 are in your household?
How many of the GIRLS in your household age 5-16 are attending school?
What is your education level?
Did you obtain a Voluntary Repatriation Form (VRF) from UNHCR when you returned?
Do you know the VRF number?
Please read out your VRF number.
Do you have a Tazkira?
How many members of your household have a Tazkira?
Do ALL women and girls in the household have a Tazkira?
In which province do you currently live?
Which district?
How long have you been living in this community?
Is this your province of origin?
Is this your district of origin?
What is the main reason why you decided to settle in your place of origin?
What is the main reason why you chose to settle here (and not your place of origin)?
What is the most important difficulty faced by your household in this area?
Are you planning to leave this place?
Where are you planning to go?
Which province?
Why are you planning to leave?
Upon returning to Afghanistan, is this the first place you came?
In which province did you first settle?
Which district?
Why did you go there?
How many months did you stay there?
Why did you leave?
How many places did you move after leaving this first place and before coming here?
What did you leave behind before returning to Afghanistan?
What is the main source of income for your household?
Was this your main source of household income before you returned to Afghanistan?
What was the main source of income for your household before you returned to Afghanistan?
Why is your main source of income different now?
Do you currently have a job?
How long have you had this job?
How did you find this job?
How long did it take you to find this job?
How many MEN (aged 14-64) in the household are currently working for pay?
How many MEN (aged 14-64) in the household are currently NOT working / staying at home?

Are you the main breadwinner for the household?
Now think about the main breadwinner for the family (the one who usually earns the most income)
Does the main breadwinner currently have a job?
How long have they NOT had a job?
What did you leave behind before returning to Afghanistan?
What is the main source of income for your household?
Was this your main source of household income before you returned to Afghanistan?
Why is your main source of income different now?
Do you currently have a job?
How long have you had this job?
How did you find this job?
How long did it take you to find this job?
How many MEN (aged 14-64) in the household are currently working for pay?
How many MEN (aged 14-64) in the household are currently NOT working / staying at home?
Are you the main breadwinner for the household?
Now think about the main breadwinner for the family (the one who usually earns the most income)
Does the main breadwinner currently have a job?
How long have they NOT had a job?
How many days did they work last week?
On average, how many hours a day do they work (on the days they go to work)?
How often do they get paid?
What type of job do they have?
Do they run their own business / home enterprise?
What is the main challenge they face in running a business?
What type of job did they have before returning to Afghanistan?
Did they want to start their own business when they came here?
Why could they not start their own business?
Are you in debt?
How much?
Are you able to pay your debt?
How much income does the bread winner make in a typical month?
Did you or any household members use medical / health facilities during the last 3 months?
Which services did they use?
What are the nearest / closest types of medical care you can access?
What is the main difficulty you face in accessing medical care?
Are there schools available within walking distance (1 hour)?
What type of school is available within walking distance (1 hour)
What is the main reason why some BOYS in the household are not attending school regularly?
What is the main reason why some GIRLS in the household are not attending school regularly?
What is the main challenge faced by children who are attending school regularly?
In times of need, have any children in the household below the age of 14 contributed to the household income?
In the past week, did you or any household member skip a meal or eat less because there was not enough food?
How many different households do you know that have returned to Afghanistan in 2015 or later?
Did you or any household members use medical / health facilities during the last 3 months?
Which services did they use?
We are trying to contact people who returned to Afghanistan in 2015, to better understand the situation. Can you provide us with the mobile phone numbers of any other people we could call?
Thank you for participating!