A hard-to-reach population can be defined - in the research context - as [1] any population who is difficult to establish a list from which a sample can be selected (e.g. list of undocumented workers) or [2] a population that it is difficult to get in touch with (e.g. don’t have a phone number).

Hard-to-reach populations described in the literature are typically marginalized by social-economic factors. For example:

- homeless and precariously housed (e.g. no fixed address, move a lot, no phone number)
- drug users, sex workers, undocumented immigrants (e.g. afraid to disclose)

Marginalized populations, however, are not the only hard-to-reach groups. Rich, powerful and famous people (e.g. high level politicians, famous artists, top executives) and people with busy schedules can also be hard-to-reach, since they tend to make themselves less accessible than the general population.

**Sampling hard-to-reach populations**

Sampling hard-to-reach populations requires a variety of strategies tailored to each specific population and research project. Most of these strategies will include an understanding of the populations’ social networks - how one member of the population relates to another as well as how they relate to different organizations and venues.

All strategies should be based on the development of a meaningful and respectful relationship with the population being studied.

Examples of sampling strategies that build upon the relationships between participants from the same population are snowball sampling and respondent-driven sampling. An example of a sampling strategy that builds upon the relationship of potential participants with organizations and/or particular venues is time-location sampling.

**Snowball sampling**, also called chain referral sampling, is a non-probabilistic sampling technique where existing research participants recruit and/or refer people they know to also participate in the research project. For example, with a goal of recruiting 100 participants, researchers identify 10 undocumented immigrants and ask them to participate in a research study and refer 10 other participants to the study that are also undocumented immigrants.

**Respondent-driven sampling** (RDS) combines snowball sampling with a mathematical model that weights the sample to compensate for the fact that the sample was collected in a non-random way. The first recruited participants are carefully selected to represent different segments of the population and act as seeds for the recruitment of other participants. Participants are given a fixed number of serialized referral coupons to recruit peers. More information can be found at http://respondentdrivensampling.org/.
Time-location sampling, also known as venue-based and facility-based sampling, uses venue-day-time (VDT) units to develop a sampling frame. For example, researchers select 4 public spaces in a particular neighborhood of interest and approach people in these places during selected timeframes. Identifiers can be collected to prevent duplicate participation by the same respondent at different VDTs.

An effective approach is to utilize a combination of facility-based sampling with cluster sampling (e.g. researchers select organizations that provide services to homeless people (cluster sampling) and approach participants within the organizations that utilize meal programs in these organizations during selected VDTs).

It is important to keep in mind that Snowball Sampling is a non-probabilistic technique and both, RDS and Time-Location Sampling although non probabilistic techniques are able to offer probability approximations based on mathematical models of participant relationships (RDS) and venue-day-time sampling frames (time-location sampling).

Recruiting and following-up with hard-to-reach populations

Reaching hard-to-reach populations requires a planned approach and innovative strategies. Some of the strategies that have been used successfully in the past are:

(1) The development of clear and adaptable recruitment and participant management protocols that can guide the process.

(2) The development of a project specific participant management database, which will help provide access to accurate and timely participant contact and project information.

(3) Having access to other databases, such as social services, drug and health, housing providers, etc. This usually requires a contract and will require prior participant consent.

(4) Having an established network of partners that can assist with obtaining contact information (e.g. service providers, community partners, etc.). This will require prior participant consent.

(5) Outreach recruitment strategies, including visiting service providers, public places and participants’ homes.

(6) Different modes of communication, including phone, mail, email, posters and social media (Facebook).

(7) Development of clear messages that inform participants of the proper incentives for participation (e.g. advantages to the community, honorarium, etc.).

(8) Establishment of frequent loops of communication with participants in-between follow-up appointments (e.g. calling in to update their contact information every month).

(9) Continuous development of a relationship with the study participants based on mutual respect and benefits (e.g. hiring community-based coordinators and interviewers who are known and trusted by the population of interest).