REACHING THE UNREACHABLE. INNOVATIVE STRATEGIES TO APPROACH HARD-TO-REACH POPULATIONS

Nadia Nur, Federico Di Leo

Abstract. This paper presents an overview of the methodological approaches used by Istat to study hard-to-reach populations, focusing on Roma and Sinti people living in settlements and the homeless in Rome. These populations, characterized by their marginal status and preference for anonymity, challenge traditional sampling methods, necessitating innovative strategies. Our objective is to outline and compare innovative methodological strategies that Istat is implementing to carry on: a) a survey on housing conditions and social inclusion of Roma and Sinti people living in settlements; b) a count/survey of homeless people. For the Roma and Sinti survey, we employ a random walk technique for household selection in settlements, relying on self-identification and specific guidelines for contacting households. In contrast, the homeless count utilizes a point-in-time approach to estimate demographic and social characteristics. This involves a "snapshot" count conducted in a high-density homeless neighborhood. By comparing these methodologies, we aim to contribute to the advancement of official statistics in the realm of investigating hard-to-reach communities.

1. Data collection on Roma households living in settlements

1.1. Background

The Roma community in Italy has long faced significant housing challenges, residing in settlements with substandard living conditions and lacking access to proper infrastructure and basic services. These issues are compounded by historical discrimination, poverty, and social exclusion. To tackle these pressing concerns, policies aimed at relocating Roma households to adequate housing and enhancing social inclusion have been implemented. In pursuit of the objectives identified by the National Strategy for Roma, Sinti, and Caminanti inclusion 2012-2020 (Unar, 2022), Istat built up a set of surveys, with the cooperation of the National Platform of Roma

associations and Unar[1]. The overall project carried out by Istat adopted a two-stage selection approach to assess the effectiveness of policies and for conducting a comprehensive survey on housing conditions and social inclusion of Roma households. The first stage involved a broad survey covering municipalities with over 15,000 residents. The primary objectives were to identify Roma households that had been successfully relocated to adequate housing from informal settlements, evaluate the efficacy of policies related to such relocations, and quantify the number of Roma people still residing in both formal and informal settlements (Nur, 2021). Through stage one, municipalities with significant Roma populations have been selected for further investigation in stage two. Starting from the information collected from the municipalities, we built a list of households that were living in adequate housing, and we located Roma settlements (information has been updated in 2021). The second stage of the research focuses on in-depth investigation and involves two target groups: Roma households currently residing in adequate housing after relocation and households living in formal settlements[2]. The overarching objectives of the second stage are twofold: a. To assess the living conditions and level of social inclusion of Roma households that had been successfully relocated to adequate housing. This involves evaluating the improvements in housing conditions, access to basic services, and overall well-being of the residents. b. To examine the housing conditions and social inclusion status of households living in formal settlements, with an emphasis on identifying the challenges they face and informing future policy initiatives. To ensure accurate and culturally sensitive data collection, the study employs Computer-Assisted Personal Interviews (CAPI) with a structured questionnaire. This section outlines the use of the random walk technique for data collection and its adaptation and the potential benefits and challenges of implementing this technique in Roma settlements, providing a comprehensive overview of the process and procedures followed by interviewers to ensure the survey's representation of the Roma population[2].

### 1.2. *Data collection techniques in Roma settlements*

Studying Roma households living in settlements is crucial for understanding their living conditions and social dynamics and identifying areas for improvement although they often present unique methodological challenges due to their marginalized status, social complexities, and lack of official recognition. Traditional

---

[1] ISTAT-UNAR collaboration agreement for the establishment of a statistical information framework on social disadvantage and housing conditions of Roma, Sinti and Caminanti (Rsc) people. The execution of the agreement is supported by PON Inclusion ESF 2014-2020 to Axis4 - Specific objective 11.1, Action 11.1.3.
[2] Data on Roma settlements have been collected from March to June 2023. Results of the survey will be published by mid 2024.

sampling methodologies often face significant challenges, making them less effective or not suitable for these populations, Typically, they rely on a well-defined and accessible sampling frame, such as a list of households or individuals. Hard-to-reach populations, like Roma camps dwellers or homeless often do not have fixed addresses or are not included in standard registries, making it difficult to create a comprehensive and accurate sampling frame. Moreover, these populations can be highly mobile and difficult to locate and follow up with. Thus, traditional data collection methods may yield limited results. To achieve accurate and comprehensive data collection, an unconventional and tailored methodology is necessary, as outlined by other European surveys (Fra, 2023 and Fra, 2022)). The random walk technique offers a promising alternative for collecting data in such contexts, potentially mitigating some of the difficulties encountered by researchers. Among the key factors contributing to the survey's success, the correct implementation of household selection procedures plays a crucial role. These selection procedures are built upon an innovative random walk approach, which incorporated a direction matrix to enhance accuracy. To ensure systematic coverage during the random walk, researchers must delineate the boundaries of the informal settlement. By defining the survey area, the possibility of missing households or biased data collection is reduced. The approach adopted by the Italian National Institute of Statistics (Istat) involves a random walk technique with a predefined selection interval from an assigned starting point within the "Roma camp", usually the main entrance if visible (figure 1). Interviewers proceed to move randomly within the settlement, following specific rules for each step. To maintain consistency, interviewers move in concentric circles, following a counterclockwise route. This systematic movement allows for comprehensive coverage of the settlement, reducing the likelihood of missing households. This method eliminates potential interviewer bias, allowing for an equitable representation of households.

Figure 1 – *Random route.*
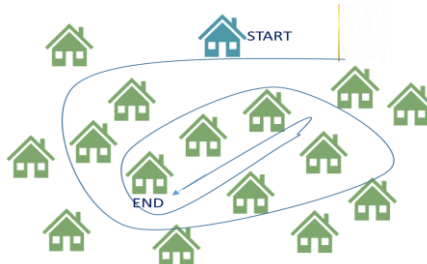
Interviewers apply a sampling step of five, meaning they make contact with every fifth dwelling unit in the camp. In the case of small camps, a sample interval of three is preferred to obtain a sufficient number of data points (Figure 2).

Figure 2 – *Selection interval.*

By employing a predefined selection interval and adhering to specific guidelines during data collection, the random walk technique proves effective in capturing the heterogeneity of the settlement while avoiding discretionary decisions and enhancing the screening process for accurate identification of the target population.

Owing to the sensitivity of asking if they have a Roma background, enumerating and screening for eligible households is usually a step conducted by the interviewers. A mandatory question to ascertain whether the respondent belongs to the Roma community is posed by the interviewers. This method helps avoid misclassification and enhances the accuracy of data collection, while ensuring that only willing participants are included in the survey. The interview can be done only if the head of the household, his/her spouse or partner, or an adult member of the family (+18) who lives in the selected house accepts to be interviewed. If no suitable respondent is found, the interviewer schedules an appointment. If nobody is found in the selected dwelling unit, the interviewer leaves his business card to schedule an appointment with the head of the family or another person suitable for the interview. By employing this approach, researchers can overcome some of the barriers associated with conventional methods and gather more comprehensive and representative data.

1.3. Challenges and potential benefits of random walk data collection

Data collection in Roma settlements is challenging due to various factors, including:
- mistrust and stigmatisation: the Roma community may be reluctant to participate in data collection activities due to historical discrimination and stigmatisation. Some may be reluctant to self-identify or be unwilling to participate in research for various reasons, including mistrust of researchers.
- language barriers that may hinder effective communication between researchers and residents, impacting the quality of data collected

- conducting field research in settlements can be logistically challenging since it requires gaining access to the camp and obtaining permission from its residents.

Engaging with the community and seeking their participation and input throughout the research process can help to mitigate some of these challenges and ethical concerns. Collaboration with associations (Roma and pro-roma) is essential to the success of this study. Local associations play a crucial role in facilitating access to Roma communities, building trust, and ensuring cooperation with the survey. The involvement of associations and Roma communities strengthens the survey's outreach and credibility. Moreover, this study includes Roma interviewers where possible and recognizes the significance of cultural mediation. Roma interviewers possess a unique understanding of the community's culture and norms, establishing a comfortable environment for respondents. Their presence fosters mutual trust, thus yielding more accurate and insightful data. Using random walk techniques in vulnerable communities like Roma settlements also raises ethical concerns. Interviewers must be mindful of the potential impact on the residents and ensure that their work respects the dignity and privacy of the individuals involved. Interviewers with Roma background and cultural mediators can assure more awareness of ethical aspects. To ensure data quality, proper training is provided to the interviewers regarding the methodology and approach. Additionally, regular supervision and monitoring of the data collection process are conducted to address any issues promptly. Although the random walk technique can potentially mitigate some of the difficulties that can be encountered by researchers, still we can outline certain challenges and limitations such as:

- irregular layout of the settlement that might not fit the assumptions of traditional random walk models, which are often designed for more regular environments. This irregularity can introduce challenges in accurately modeling the movements and interactions of individuals within the camp.
- the duration of data collection for random walk experiments can significantly impact the results. A longer observation period may provide a more comprehensive understanding of movement patterns, but it can also be resource-intensive and impractical in certain situations.
- there may be concerns about the representativeness of the sample, as the randomness of the approach might miss certain segments of the population.
- the technique can introduce biases, as certain areas or groups may be more likely to be included in the sample.

By addressing some of the challenges associated with traditional methods, the random walk technique offers a means to gather more comprehensive and representative data, leading to better insights into the realities of the Roma communities residing in settlements. However, its successful implementation

requires careful planning, ethical considerations, and adaptability to the unique circumstances of each Roma camp.

2. Methodological strategies to investigate Homelessness: the Street Count in Rome

2.1. Introduction and background

The aim of this section is to provide an overview of methodological approaches adopted by Istat and by other research teams to study a specific hard-to-reach population, homeless people (Marpsat, M., & Razafindratsima, N. 2010).

The city of Rome is used as case study, as the total homeless population has been repeatedly estimated over time using different techniques[3]:

- PIT - point-in-time sampling (in 2014);
- TLS - time-location sampling (in 2011, 2014 and 2021);
- AD - administrative data (Population Census 2011 and 2021).

This population constitutes a very small proportion of the general population but represents a very sensitive target for national and local authorities to fine-tune contrast policies due to their marginalised status. Homeless is a hard-to-reach population, for whom the traditional sampling methods are rarely feasible. The goals of the research are basically two: homeless count and the study of the individual condition and story. Census data are focused only on homeless count while the other approaches provide information on both aspects of the statistical investigation. In principle PIT is more directed to make a "snapshot" of people experiencing homelessness, on a single night. TLS is more suitable for and investigation of the condition of unsheltered people and the causes which led to a situation of extreme poverty. The Population Census, however, provided in 2011 [AD_PC 2011] and 2021 [AD_PC 2021] its own estimate of homeless people including an estimate of the presence of this population in Rome. In 2014 Istat considered also the possibility of a direct count of the people spending the night on the street, in public spaces, or in other places that are not considered appropriate shelter. The street count considered only the city of Turin because of the possibility to rely on a good network.

---

[3] A partial study has been carried out by Di Censi and others with the capture-recapture method, using a database (2022-2011) provided by the Municipality of Rome. The results are not used for this document. See Di Censi, L. (2013). Metodologie applicate per la misurazione della povertà urbana. FrancoAngeli.

In 2011 the ISTAT with the support of partners[4] realised the first National Research [TLS_NR 2011] on the condition of people living in extreme poverty. Three years later, in 2014, Istat and its partners carried out second research. [TLS_NR 2014]. "Fondazione Rodolfo Debenedetti" with the support of the Municipality (Roma Capitale) carried out in March 2014, with the Project "RacContami" a first attempt of direct count of homeless [PIT_RR 2014]. There are no further statistics on homeless people based on convenience sampling or "street count" at a national level or focusing on the situation in Rome after 2014. In 2021 a researcher with the support of the European University Institute[5] implemented a survey based on Istat approach, for the Municipality of Rome only [TLS_RM 2021].

2.2. *Homeless count: a first comparison of the results*

Homeless count in Rome according to various sources is summarised in Table 1. National Research (2011) estimated that the homeless population in Rome consisted of 7.827 while according to 2011 Population Census was 21.118 while ten years later, in 2021, was 22 182. These latter figures are mutually consistent, but they are different from the estimates based on the two National Research carried out by Istat (7.709 individuals in 2014). The "street count" conducted in 2014 with RacContami Project produced a result far different but consistent with the so-called "Rome Research". Before doing precise comparisons, it is necessary to understand why the homeless count according to the Population Census is so different from other estimates.

National Census defines homelessness according to official demographic records (*anagrafe*). The "anagrafe" registers considers people with two main characteristics:
- people who registered in conventional addresses and/or addresses of assistance bodies/institutions in which the homeless are present;
- people who are regularly present in Italy and therefore have the possibility to have an official residence address.

---

[4] The research was carried together with the Ministry of Welfare, the Italian Federation of Organizations for Homeless People (fio.PSD) and Caritas. See Istat (2014). L'indagine sulle persone senza dimora in Italia. Roma.

[5] See Gatta, A. (2023). A target on the head? Welfare targeting, conditionality and the persistence of disadvantage. PhD Thesis in Political and Social Sciences, European University Institute, Florence.

Table 1 - *Homeless count in Rome – Various sources and methodological approaches (2011-2021).*

Source	Point in Time	Time Location Sample	Administrative Data
National Research 2011		7.827	
Population Census 2011			21.118
RacContami 2014	3.276		
National Research 2014		7.709	
Population Census 2021			22.182

Istat (2011, 2014 e 2022), Braga (2014, Gatta (2022)

According to these two features many people who are registered in conventional addresses are not necessarily sleeping in open spaces or "unconventional shelters" but are persons who are eligible for an unconventional registration. In Rome, the Social Services Secretariat decides if there are conditions and there is not any official regulation to include people in this specific register. Therefore, data include many people who are in need of a special address but are not homeless like persons in collectives, civilian residents temporarily absent from the country, persons living in areas to which access is difficult, etc.

However, the estimates based on TLS and PIT approach are also partly inconsistent because of several reasons which is possible to list:
- difference in definitions;
- sampling vs snapshot approach;
- period of the year of the research (winter, spring, etc.);
- possible double counting;
- extension of period of the research (single night vs a period of research);
- special situations like the Covid-19 pandemic;
- non-sampling errors and quality of data collection.

This list can be extended, and it is possible to find solutions to arrive at a (partial) reconciliation between the different methodologies. However, I want to stress the fact that each approach has its own validity and naturally PIT is much more expensive, complex and resource demanding than TLS. There are no advantages and/or limitations adopting the three different techniques examined: point-in-time sampling (PIT); time-location sampling (TLS); administrative data (AD) estimates are not suitable for direct comparisons. However, a few further clarifications are necessary. Estimates based on time-location sampling provide information on homelessness over a period that varies according to the approach taken; in the case of the Istat surveys the time window adopted was one month wide. Therefore, the

PIT estimate and the TLS estimate only coincide under the assumption that there were no "entries" and/or "exits" from homelessness condition during the period considered. As regards to the administrative data the estimate is affected by the source characteristics: part of the homeless are not counted, i.e. those who are not registered in the "registry office-anagrafe" (or those who are not identified as homeless because they are still registered at their previous address of usual residence) while part of the "registry" as homeless may not be homeless and/or may not be in severe housing deprivation (addresses dedicated to homeless are also used by municipalities for categories that do not belong to the target groups such as unaccompanied minors in care, women victims of violence, caravanners, people who for personal/fiscal reasons find it convenient to be homeless, etc.).

2.3. *Street Count work in Paris and the Pilot Survey in Esquilino District.*

Both TLS and PIT surveys face a set of problems concerning, from the respondent side, willingness or ability to answer the questions, language barriers, fear etc. Therefore, an option for avoiding bias that can lead to underestimating the target population might be to collaborate with community organisations with access to these groups. The Pilot Survey in Esquilino District was implemented to verify the possibility to make a new PIT estimation of homelessness in the Municipality of Rome according to the approach used in Paris and other major cities. Aware of possible biases of previous research, the first goal was to arrive at a new homeless count with few demographic features (age, sex, continent of origin). Following the methodology used in Paris, starting around 9pm, small teams of volunteers walked every street in one of 24 sectors of Esquilino District (figure 3).

Figure 3 – *Map of Esquilino district.*
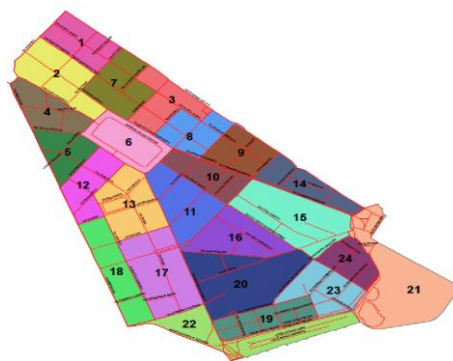
Four more teams were dedicated to the area of the main railway station (Stazione Termini) where it is possible to meet a large number of homeless. Each team had a

leader, but all the members were responsible for the various aspects of the survey: filling out the forms properly, detecting the people to be counted, and identifying the path to be followed. For every individual, the team completed a questionnaire with basic observable information. For those willing to be interviewed, the volunteers could ask more targeted questions – with responses to remain strictly anonymous. Once the team covered every street in the assigned sector, they returned completed questionnaires to the neighbourhood headquarters for an initial quality control.

2.4. *Challenges of PIT technique*

Practical implementation challenges associated with street counts can affect the accuracy and richness of the data. Given the transitory nature of rough sleepers during the day, the counts generally take place at night: but if you start too early, people may not have settled in for the night and you'll miss them; if you start too late, many people are sleeping, and data collection will be limited. Street counts are also resource-intensive, relying on many enumerators (around 2000 for the Paris count alone), who require training and supplies. Street counts also depend on relatively favourable external conditions: extreme weather or a global pandemic, for instance, complicates data collection efforts and can limit comparability across time.

The Pilot Survey results as well as the specific characteristics of Rome require specific adjustments for the final research, scheduled for the end of 2023.

Among the main changes to be introduced the most relevant are connected with:
- the questionnaire;
- the introduction of a CAPI approach;
- the correct identification of the Municipality to be covered (Rome is much bigger than Paris in term of square meters but not in terms of population);
- the number of teams to be selected for the total snapshot;
- the support of people belonging to charities vs people selected from universities (enumerators);
- the possibility to have a partner responsible for the organisation.

3. Conclusions

Homeless and Roma communities living in settlements are considered hard-to-reach populations when it comes to data collection. However, obtaining accurate and comprehensive data on these vulnerable groups is crucial for designing inclusive policies that address their specific needs and challenges effectively. To achieve this, survey design must consider several key considerations. Both surveys pose methodological challenges, such as the lack of a reliable sampling frame and the high mobility of the population, necessitating the use of innovative and flexible approaches. These approaches should strive to balance the need for accurate,

representative data with the practicalities of reaching and engaging these groups effectively. Moreover, both surveys face a set of problems concerning from the respondent side, willingness, or ability to answer the questions, language barriers, fear etc. Therefore, an option for avoiding bias that can lead to underestimating the target population might be to collaborate with community organisations with access to these groups. These organizations have established connections and trust within the target populations, making them valuable partners in the data collection process. Their involvement can facilitate access to the communities and help overcome barriers that might hinder conventional data gathering efforts. Another approach that can encourage participation and ensure that respondents feel comfortable sharing their experiences and perspectives is to involve interviewers or enumerators with specialised skills in cultural mediation or deep knowledge of the target population to create contact with the target population and gain their trust. In addition, it is essential to provide proper training to interviewers on the nuances of handling 'official' data collection. Finally, official statistics agencies should embrace innovative and adaptable data collection methodologies and techniques to obtain high-quality data from these hard-to-reach populations.

References

APUR. 2021. Les personnes en situation de rue à Paris la nuit du 30-31 janvier 2020. Available at
https://www.apur.org/sites/default/files/nuit_solidarite_2020.pdf?token=5LtIDAe
7.

BRAGA M. 2014. *Raccontami 2014 - prima indagine point in time dei senza fissa dimora a Roma*. Roma: Fondazione Rodolfo De Benedetti.

EMPLOYMENT AND SOCIAL DEVELOPMENT CANADA ESDC. 2021. Everyone Counts – A Guide to Point-in-Time Counts in Canada. Available at https://homelessnesslearninghub.ca/wp-content/uploads/2021/06/2829-Pit-Count-Report-edition-3.1-EN54.pdf.

FRA - FUNDAMENTAL RIGHTS AGENCY. 2023. Roma Survey 2021 Technical report. Available at http://fra.europa.eu/sites/default/files/fra_uploads/fra-2023-roma-survey-2021-technical-report_en.pdf.

FRA - FUNDAMENTAL RIGHTS AGENCY. 2022. Report Roma in 10 European countries – Main results: Roma Survey 2021. Available at https://fra.europa.eu/en/publication/2022/roma-survey-findings.

GATTA A. 2022. Non avere dimora a Roma. Accesso alla residenza, alle misure di sostegno al reddito e alle cure mediche per le persone senza dimora della Capitale.

Osservatorio Fio.PSD. Available at https://www.fiopsd.org/report-progetto-instrada/.

INGLESE F. AND MASI A. 2019. Una metodologia per la stima della homelessness attraverso le Unità di strada: il caso della città di Torino, *Istat Working Papers No.5*

Insee. 2021. Nuit nationale de la solidarité - Compter pour agir. Guide méthodologique, *Insee Méthodes,* No.139.

ISTAT. 2022, Popolazione residente e dinamica demografica. Available at https://www.istat.it/it/files/2022/12/CENSIMENTO-E-DINAMICA-DEMOGRAFICA-2021.pdf .

ISTAT. 2012. Le persone senza dimora. Available at https://www.istat.it/it/archivio/72163.

ISTAT. 2015. Le persone senza dimora. Available at https://www.istat.it/it/archivio/175984.

MARPSAT, M. RAZAFINDRATSIMA, N. 2010. Survey methods for hard-to-reach populations: introduction to the special issue. *Methodological Innovations Online*, Vol.5, No.2, pp.3-16.

NUR N. 2021. Abitare in transizione. Indagine sui progetti di transizione abitativa rivolti alle popolazioni Rom, Sinte e Caminanti. Available at https://www.istat.it/it/archivio/254436.

SCHNEIDER M., BRISSON, D. & BURNES, D. 2016. Do We Really Know how Many Are Homeless? An Analysis of the Point-In-Time Homelessness Count, *Families in Society*, Vol.974, pp. 321–329.

UNAR. 2022. Strategia nazionale di uguaglianza, inclusione e partecipazione di rom e sinti 2021-2030. Available at www.unar.it/portale/web/guest/strategia-rsc-2020-2030.

UNAR. 2012. Strategia nazionale di uguaglianza, inclusione e partecipazione di rom e sinti. 2012-2020. Available at https://www.minori.gov.it/it/minori/strategia-nazionale-2012-2020-dinclusione-dei-rom-dei-sinti-e-dei-caminanti.

_____

Nadia NUR, Istat, nur@istat.it
Federico DI LEO, Istat, dileo@istat.it