

## **THE FRENCH SURVEY OF HOMELESS PEOPLE USING SHELTERS AND SOUP KITCHENS**

Cécile Brousse, Bernadette Guiot de la Rochère, Emmanuel Massé<sup>1</sup>

### **ABSTRACT**

The French survey of homeless people using support services is unique because of its scope and the conditions under which it was conducted. About 4,000 users of shelters and soup kitchens were surveyed in January and February 2001. Because some users move from one service point to another, it is necessary to collect precise data on the number of times each respondent used such services (meals and person-nights) during the week preceding the survey. Data quality is extremely important since it has a major impact on the sampling weight assigned to each individual.

KEY WORDS: sample frame; exclusion; weight share; poverty; data quality; non-response; homeless.

### **1. BACKGROUND AND OBJECTIVES**

Though a major social phenomenon, homelessness is, in France as elsewhere, a form of exclusion whose exact magnitude is unknown. Few surveys have ever been conducted in France, and they covered very limited areas, such as Paris or Lyon.

In view of this gap in the public statistical system, the Conseil National de l'Information Statistique (CNIS) has been advocating a national survey since 1996 to determine the number and situation of people who do not have a permanent personal residence and to describe the processes leading to homelessness. The survey data should also provide a basis for comparing the situation of the homeless with that of people living in regular housing. The initial results are expected by early 2002.

In addition to socio-demographic characteristics, the questionnaire gathered information about the labour market status, income, indebtedness, use of social services and support, health, access to health care, and family history of respondents. The individual's residential history and difficulties in getting access to housing were of particular interest. The questionnaire also provided precise data on respondents' living arrangements: whether they were living in a group setting, a dwelling or a makeshift shelter. People contacted in soup kitchens who were not homeless, though strictly speaking out of scope, were asked to provide some information.

### **2. METHODOLOGY**

The methodology used was based on work by the Bureau of the Census (1996) and France's Institut National des Études Démographiques (1995 and 1998). It involved sampling support services intended primarily for the homeless.

---

<sup>1</sup> INSEE, 18 Boulevard Adolphe Pinard, 75675 Paris, France.

The authors would like to thank Maryse Marpsat (INED), Pascal Arduin (INED) and Pascal Ardilly for their advice on methodology, and Xavier d'Haultfœuille and Elise Montier for proofreading the article. Martine Quaglia and Nancy Viard assisted in organizing the survey and collecting the data.

## 2.1 Scope of the survey

The theoretical target population consists of homeless people who sleep in the street, in makeshift shelters or in free (or nearly free) accommodations provided by organizations. These accommodations may be in group facilities, hotel rooms or ordinary apartments. They may be offered for periods ranging from a few days, in emergency situations, to several months. There is no comprehensive, reliable sample frame for the homeless. Consequently, INSEE decided to use services (person-nights and meals provided by shelters and soup kitchens) as the survey unit.

This operational definition differs from the definition of the homeless in two respects:

1. A particular homeless person may never use the above-mentioned services (at least during the survey's reference period).
2. Conversely, a person who has a home may eat at a soup kitchen.

Other types of services could have been used, such as day shelters, clothing handouts, or medical or legal assistance. A specific "qualitative" survey of day shelters supports the hypothesis that a very small percentage of the homeless who use such shelters do not use any other service in the frame. This suggests that the undercoverage due to the omission of day shelters is minor.

## 2.2 Services sample frame

Using administrative files (files from organizations that assist people in difficulty, from the ministry of employment and solidarity, and from INSEE's inventory of medical and social institutions), INSEE's regional branches constructed sample frames in 80 agglomerations with population over 20,000. To that end, sources were compared region by region, and many contacts were made both locally with social organizations (associations, communities and governments) and nationally through a self-administered survey of 1,600 places of worship. In March 2000, a telephone survey of 2,800 organizations confirmed that the inventory was useful and collected additional information about the characteristics of the services they provide (types of users, days and hours of operation, average number of users during hours of operation). Nearly half of the organizations were found to be out of scope. To improve the quality of the frame, the heads of the organizations included in the inventory were sent a letter in July 2000 asking them to update the list of agencies in their communes. As a result of this mailing, 120 new organizations were added and eight were determined to be out of scope. The final frame consisted of 2,398 services in 1,463 organizations. These services are of the following types: group accommodation (emergency and non-emergency), individual accommodation (dwellings and hotel rooms), static meal service (lunch and dinner) and mobile meal service (lunch and dinner).

## 2.3 Sampling of persons

The services were selected by three-stage sampling.

In the first stage, agglomerations with population over 20,000 were selected by probability-proportional-to-size sampling. Geographically, the scope was restricted to agglomerations with population over 20,000, since smaller agglomerations have very few shelters, as shown by a supplementary study of agglomerations with population between 5,000 and 20,000.

In the second stage, sampled units belonged to the set formed by the Cartesian product of services in the frame and days in the survey. In other words, the sample consisted of ordered pairs (service, day). Before the sample was selected, the population was stratified by type of service and, for shelter services, by type of user (men only / women only / mixed). Selection was proportional to the average size reported by the organizations in the telephone survey, deflated by the agglomeration's probability of selection. The goal was to have the same probability of selection for each service. The result was a sample of 1,225 pairs

(service, day). For additional information, see the paper in the *Actes des journées de méthodologie statistiques* (Cécile Brousse and Emmanuel Massé, December 2000).

In the third stage, the services were sampled. For a given service and day, a fixed number of services were selected (two or four depending on the type of service). An individual was associated with each service. Two different sampling methods were used. For services included in a list, the selection was made by systematic sampling from the list. For services for which there was no list (primarily meal services), a location passed by all users (entrance, exit, food distribution table) was identified, and users were selected by order of arrival. Because non-response was so common, the interviewers had to select successive replacements for refusals to ensure that the sample did not become too small. Limited to four, these replacements were all recorded, together with the reasons for refusal, in a document referred to as a contact sheet.

## **2.4 Calculation of weights**

As we have seen, the statistical unit of the survey was the use of a service rather than the user. Under the sample design, the weights of the services can be determined, but not the weights of the users. The weight share method was used to go from the former to the latter (Ardilly, Leblanc, 2000; Lavallée, 1995). To compute the weights of individuals by that method, one must know how frequently they use the services.

Frequency of use was measured over a week. In the “diary” section of the questionnaire, which took about 10 minutes to complete, respondents were asked to describe the places where they ate and slept during the week. It would have been preferable to cover a longer period, but memory effects would have degraded data quality.

Ultimately, then, the weights of individuals depended essentially on two factors:

- the service selection probabilities, which can only be computed if the actual frequency of use of the services on survey day is known;
- the number of service contacts with respondents (services used by respondents) during the reference period.

Since frequency of use is measured on a weekly basis, the method leads us to an operational definition of the homeless “according to the survey”: a homeless person is a person who had no personal or family residence for at least one night during the reference period and therefore had to sleep in the street, in a place not designed for habitation, or in a place provided by a shelter service.

## **3. INITIAL INDICATIONS OF COLLECTION QUALITY**

Collection took place in January and February 2001. In all, 4,195 questionnaires were completed (the goal was to have at least 4,000 respondents), but about 50 of them will not be used because the respondents were selected by mistake (minors, organization staff, volunteers). The interviews averaged 56 minutes, but the length varied substantially with the survey location; 3% of the interviews took over 90 minutes.

Collection quality was assessed by means of three sources of complementary information. The first source was the collection documents themselves, where refusals and enumeration results were recorded. Second, in the last part of the questionnaire the interviewers were asked for their opinions on the quality of the responses: whether the respondent understood the questions; whether he/she was suspicious before and/or after the interview; whether he/she was interested in the interview; whether he/she had memory problems in completing the diary or distinguishing one night or meal from another. Third, two weeks after the end of collection, the interviewers completed a 15-minute evaluation questionnaire covering three of the visits in which they participated (the first, the second-last, and the last). The questionnaire has provided some valuable information about their perceptions concerning sampling, interviewing and training.

### 3.1 The collection process

A unique team-based structure was introduced following six tests with 300 questionnaires in seven of INSEE's regional branches. Each team assigned to survey a particular service consisted of one "sampler", who enumerated the population present and selected the subjects to be interviewed, and one or two interviewers, responsible for completing two questionnaires each. In services where there was a list of users, the samplers first selected respondents at random and then conducted two interviews. In other services, their task was simply to select the samples.

#### - preparation and organization of collection

Special emphasis was placed on training: 3.5 days for the 135 samplers, and 2.5 days for the 190 interviewers. In each training session, a "dry run" was conducted in a day shelter. This training under actual survey conditions was considered useful or very useful by 95% of the interviewers.

The sampler tasks were strictly differentiated from the interviewer tasks. In practice, however, there was a substantial amount of cooperation: in 25% of the cases, the interviewers helped the samplers make contact with the selected subjects, and in 10% of the cases, they also assisted in the enumeration. Conversely, when time constraints were tight (in mobile soup kitchens where service was very rapid), the samplers completed questionnaires to help out the interviewers. In addition, in places where collection was apt to be complicated, the survey teams were accompanied by regional survey managers.

#### - sampling

It is difficult to objectively assess the quality of the enumeration and sampling carried out in the field. Qualitatively, the main risk factors were (1) the unusual conditions involved in surveying without a list (19% of visits), which might make identifying individuals and tracking their movements (critical situations in mobile soup kitchens, for example) very difficult, and (2) the temptation that might affect a sampler, faced with a difficult population, to introduce an element of subjectivity into the selection of the people to be interviewed.

Nevertheless, a few objective data are available at this time, such as the perception of sampling training, which 90% of the interviewers judged satisfactory, and compliance with the selection and replacement instructions, which enabled the teams to enumerate the population correctly in 88% of the services visited and to use the selection table correctly 87% of the time. In 84% of the visits, all refusals and cases of inability to respond were recorded. There were significant differences depending on the survey location. Selection problems were more common in soup kitchens, while replacement difficulties were more frequent in dispersed shelter services (hotel rooms or dwellings).

**Factors that forced samplers to adjust their instructions, by service visited  
(more than one response possible)**

	Dispersed shelter	Non-emergency group shelter	Emergency group shelter	Meal service
The organization did not allow you to follow them	9.6%	16.7%	0%	0%
There were too few of you	0%	2.8%	12.5%	25%
The pre-visit information was incomplete	11.3%	16.7%	25%	13.8%
The service was delivered too quickly	1.9%	5.6%	0%	58.6%
There were too many refusals and cases of inability to answer	13.2%	8.3%	12.5%	34.5%
There were too few people	20.8%	13.9%	25.0%	13.78%
The people were difficult to contact	46.3%	36.1%	0%	6.9%
The numbers in the selection table were too close together	11.5%	2.8%	0%	16.7%
The instructions did not cover this case	17.0%	8.3%	14.3%	10.3%
Other reasons	26.4%	29.7%	37.5%	19.2%

Source: Survey of samplers (February-March 2001)

Another indication of the quality of sampling was compliance with the visit schedule: 95% of the interviews were conducted on the date scheduled; most of the survey teams conducted their interviews within three days following the planned date.

- consent to the survey

There was concern about the possibility of refusals at two levels: the support services and the actual respondents. In fact, very few of the services refused to cooperate; only 1.7% of them denied the survey teams access. This low refusal rate was due to the backing of the major associations, which were consulted throughout the process of preparing the survey, and to the hard work of the survey teams.

As for the interview subjects, initial processing of the contact sheets puts the non-response rate at about 35%. The conception and enumeration of “final” non-response were considerably different from what one might find in ordinary household surveys. The interview teams were allowed to select no more than four replacements for the selected person. After the fifth failed contact, the questionnaire was set aside, and the sample size decreased by one. Replacements were permitted so that the sample would not become too small, but obviously that in no way solved the problem of bias in the associated estimator. In all, 6,281 contacts were needed to obtain the 4,195 questionnaires.

The failed contacts can be broken down as follows: 13.5% refusals, 14% unable to answer (12.5% unable to speak French) and 6% impossible to contact. The refusal and “unable” non-response rate was higher for men than for women (33% compared with 21%), and it increased with the age of the person contacted (21% for people under 30, 30% for people 30 and over). Soup kitchens (static or mobile) had both the highest refusal rate and the highest “unable” rate: the consent rate ranged from 74% in non-emergency group shelters to 32% in mobile dinner services.

In about 60% of the cases, the survey teams were able to collect a “stated” reason for refusal. The reasons were worded in many different ways, of course, but qualitatively, the most common were lack of time and lack of interest.

- interviewing

Overall, 96.8% of the interviews went well. However, a few difficulties were noted in mobile soup kitchens, as 10% of the interviews went poorly or, in a few cases, were interrupted. The most common reasons cited by the interviewers were drinking alcohol, taking drugs and taking medication. An atmosphere of trust prevailed during the interviews. Prior to the interview, 16% of respondents were slightly suspicious (compared with 6.8% after the interview), and 3% were very suspicious (1.2% afterwards). Three quarters of the respondents understood the questions well or very well; only 4% misunderstood the questions. The figures for respondents’ ability to express themselves were similar.

### **3.2 Sample frame**

The results described below are preliminary. They were obtained by processing 3,000 questionnaires from 11 of the 17 regions covered by the survey. Processing will be completed in December 2001. If we consider which regions have already been processed, the final results should not change the analyses presented here.

The quality of the sample frame can be assessed ex post with the information collected by the survey teams.

- quality of the frame used for collection (January-February 2001 frame)

The frame did not deteriorate much between March 2000 and January 2001, as only 2.7% of the sampled services disappeared during that period. Pre-visits determined that 1.7% of the services sampled were out of scope.

The information collected in the diaries, especially the addresses of the services used, provided a measure of the frame’s completeness. Detailed analysis of that information revealed that, of the 673 services mentioned in the 2,998 diaries, 77 were not listed in the frame. Of those 77 services, 21 were inadvertent omissions from the frame. Most of them were soup kitchens. About 30 of the 77 required further investigation. This was the case for parishes and work camps, for example, which were in scope only if they provided a regular meal service. For about 25 of the services, the descriptions were too imprecise to permit

identification. It was impossible to determine whether they were already listed (for instance, “restos du coeur”, “numéro 115”, with no further information) or even whether they were in scope (group home, prefab, train, bus, etc.).

A few missing organizations may not have been picked up by the diaries because they provide both meal and shelter services. Their residents would not have been sampled in other services and thus could not have mentioned the organizations in the diaries.

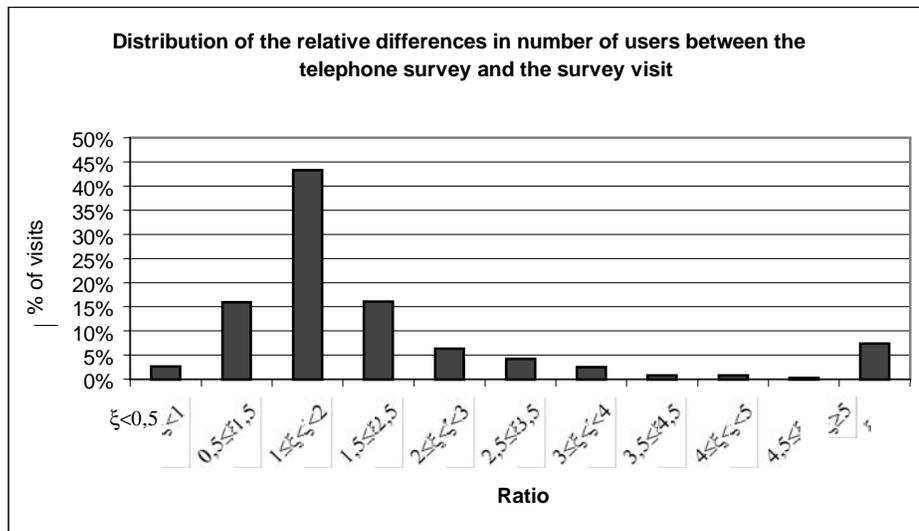
- difference in number of users between January 2000 and January 2001

There was a significant difference between the number of users recorded in the January 2000 telephone survey and the number counted during the survey in January-February 2001. On average, the telephone survey figures were 37% higher than the counts made by the survey teams when they enumerated users for the purpose of selecting interview subjects (Chart 1).

The number of users was measured on three occasions with three different collection methods. First, organization managers were interviewed by telephone in March 2000. Second, during a pre-visit a few weeks before the survey, the survey teams asked organization managers for an estimate of the number of users expected. Third, the sampler counted the users on the day of the survey. For services for which there was no list, managers were asked to provide estimates of the number of users.

The pre-visits gave the samplers an opportunity to identify differences between the descriptions provided by organization managers and the information obtained in the telephone survey. One quarter of the pre-visits uncovered discrepancies either in the nature of the services offered by the organizations or in the number of users. The samplers attributed these discrepancies to various factors: 17% of the differences were due to changes in the number of users between the two survey dates, 31% to the fact that the telephone survey questions were misunderstood, 13% to the fact that two different people responded to the two surveys, and 10% to differences in interpreting the categories suggested to the managers.

**Chart 1**



How to read the chart: The second bar from the left indicates that, for 16% of the visits, the ratio of the number of users reported in the telephone survey to the number counted on survey day was between 0.5 and 1.

Source: Survey of persons using shelters and soup kitchens (January-February 2001).

The number of users measured on survey day was 6% below the pre-visit estimate. For mobile soup kitchens in particular, the number of users is very difficult to forecast from day to day, and the managers' pre-visit estimates were 18% higher than the actual counts by the samplers on survey day.

For services for which there was no list, it was possible to compare the organization's counts with the survey teams' counts. The differences were small: the number of users recorded by the organizations on a survey day were, on average, 4.3% higher than the samplers' estimates.

### 3.3 Memory effects and diary quality

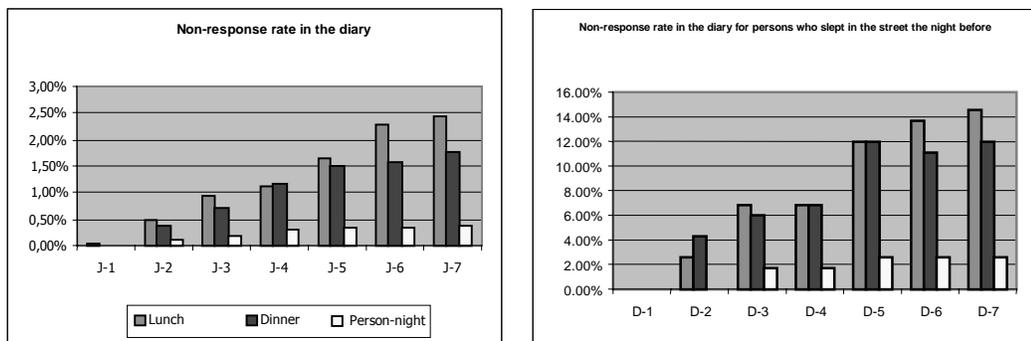
We have two sources of information we can use to measure the impact of memory effects on the quality of the diary.

According to the data provided by the interviewers in the final section of the questionnaire, only 4% of the respondents had difficulty remembering where they slept in the seven days preceding survey day. The rate was higher for meals, as 11% of the respondents had trouble remembering. Non-response on the diary section was very low. Again, it was higher for meals than for sleeping locations. In addition, the farther away the event was in time, the higher the non-response rate became. The rate was nearly zero for the day before the survey, compared with about 2% for meals seven days before.

Memory problems varied substantially with the type of service where respondents were contacted. According to the interviewers, 6% of people contacted in dispersed housing or non-emergency shelters had difficulty answering questions about the meals they had the week before, compared with 32% of those interviewed in mobile soup kitchens serving dinner. These differences are related to the stability of the lifestyle: 37% of people who slept in the street the night before had trouble answering questions about the meals they had during the week (20% for sleeping location).

The same patterns are evident in the diary non-response rates. Over 10% of respondents who slept in the street the previous night were unable to say where they ate more than four days before the survey. The non-response rates for sleeping locations were under 3%.

Chart 2



Source: Survey of persons using shelters and soup kitchens (January-February 2001)

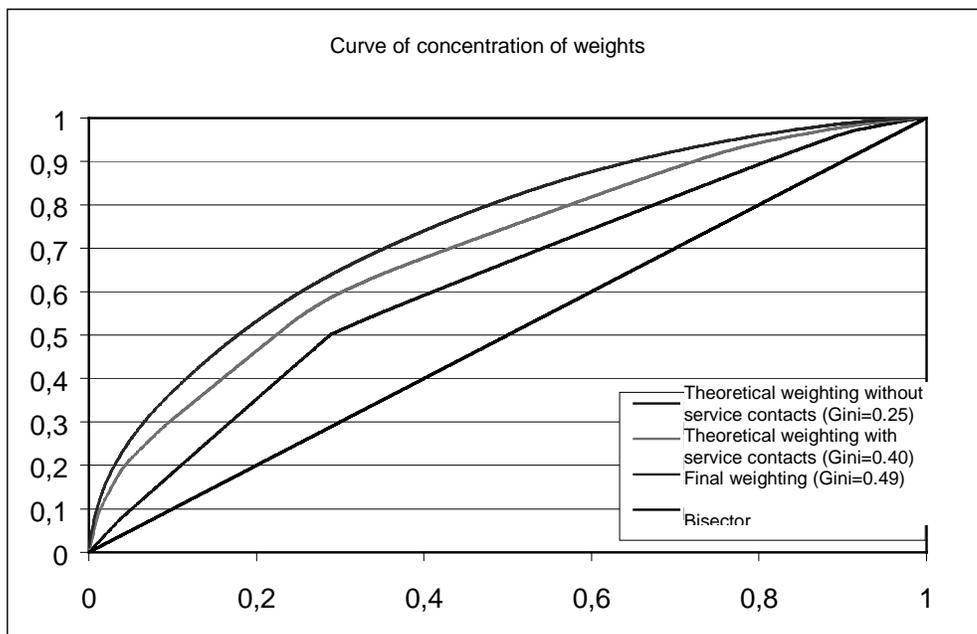
### 3.4 Weighting

While the method used was designed to produce equal or nearly equal weights for all respondents, significant variation in weights was noted. That variation was due to a number of factors:

- the initial theoretical sampling weights of the services (because of the collection workload, the Île de France region, which includes Paris and contains about a third of all the services, was undersampled);
- the dispersion of service contacts with individuals who were interviewed (under the weight share method, the weight is inversely proportional to the number of service contacts);
- discrepancies between the sample frame and measurements made in the field (in particular the number of users reported in the telephone survey and the number counted during the survey).

The successive effects of these factors on the curve of concentration of the weights are shown in Chart 3 below. Some individuals affected by more than one factor (sampled in Paris, did not use any of the services in the week preceding survey day, contacted in a service whose number of users increased between frames) had comparatively high weights.

**Chart 3**



Source: Survey of persons using shelters and soup kitchens (January-February 2001)

Since the special sampling method used in this survey resulted in a high dispersion of the weights (50% of the questionnaires account for 81% of the weight), particular attention must be paid to the quality of the high-weight questionnaires.

## REFERENCES

Ardilly, P. and Le blanc, D. (2001), "Sampling and Weighting a Survey of Homeless Persons: a French Example", *Survey Methodology*, 27, pp 109-118.

Arduin, P. and Viard, N. « Méthode de collecte de l'enquête auprès des utilisateurs de service d'hébergement et de restauration gratuite. Acte du colloque francophone sur les sondages » June 2000.

Brousse C., Massé E. (2000) : L'enquête auprès des personnes fréquentant les lieux d'hébergement et de distribution de repas, *Actes des journées de méthodologie Statistiques* (à paraître).

Firdion, J. M., Marpsat M. (2000) : La rue et le foyer, *Travaux et Documents INED* n°144

Lavallée, P. (1995) : Cross-sectional Weighting of Longitudinal Surveys of Individuals and Households using the Weight Share Method, *Survey Methodology* vol. 21, p.25-32.