



Survey of Homeless in Canada
Street component
Feasibility Study

François Laflamme
Statistical Consultation Group
Statistics Canada
April 11, 2001



Statistics
Canada

Statistique
Canada

Background

Statistics Canada (STC) has been contracted by Human Resources Development Canada (HRDC) to look into the feasibility of conducting a survey of homelessness in several Canadian cities. The primary objective of a survey of the homeless would be to obtain absolute counts of the number of homeless individuals. A secondary objective would be to collect socio-economic information to better understand the issue of homelessness.

In the first phase of this feasibility study, Statistics Canada conducted a literature review of the methodology used in previous Homeless studies in Canada and the USA¹. The purpose of this review was to focus on the lessons learned from these studies in order to present some preliminary options to be investigated for the HRDC/STC study. In addition to this literature review, Mantel and Yung produced a preliminary report on *Sample Design Options for a Survey of Homeless in Canada*². In it, they recommended an investigation of a multi-frame approach based on the observation that on any given night the homeless population can usefully be divided into two components – the shelter population and the street population.

The shelter component is thought to be the much larger contributor to total homeless counts. It is anticipated that it will be the easier of the two to obtain accurate counts of homeless individuals due to the fact that government almost universally supports these facilities with grant money and their universal registration. There are currently two activities underway which could facilitate a survey of the shelter component-HIFIS and the 2001 Census. Both are administrative in nature and are attempts to list all facilities providing shelter facilities to homeless individuals. In the case of HIFIS it will also be possible to obtain some information on the clients using these facilities during a particular time period.

The primary objective of this, the second phase of this project is to investigate the feasibility of conducting a pilot survey in selected Canadian urban areas in order to obtain accurate counts for the street component of the homeless population. This component is thought to be the much smaller contributor to total homeless counts and thought not to be a significant contributor in rural and even some urban areas. On the other hand, it will be the component for which it will be the most difficult to obtain accurate counts. Before going further in the development of proposed methodologies for this study, additional discussion must take place in order to clarify basic concepts and to have a common and uniform understanding of various issues.

Objectives

One objective of this document is to review, from a methodological perspective, the first steps of the survey to address and raise specific issues that must be discussed prior to any pilot survey-taking place.

The first part of this document entails a general description of all components of the homeless population. It then describes potential survey frames to survey this 'difficult to reach' population. Some concepts and definitions are presented and discussed throughout. A discussion on potential reference periods for the survey is finally presented.

¹ Mantel, H. and Yung, W., *First Steps towards a Survey of Homelessness in Canada – Lessons from Previous Studies*, April 2000

² Mantel, H. and Yung, W., *Sample Design Options for a Survey of Homeless in Canada*, August 2000

The second part of this document focuses on the street component of a Survey of the Homeless. The coverage of this sub-population constitutes Statistic Canada's mandate for the second phase of this feasibility study. Some definitions are first presented and sampling strategies are proposed in splitting the street population into two more refined sub-populations. A discussion on data collection issues is provided and some other considerations for the field test are described.

Section 1: Overall Homeless Study

Objectives

The general objective of the HRDC\STC Homeless Study is to get a point-in-time measure of the homeless population by capturing information about individuals who either have no housing at all or are staying in temporary forms of shelter.

1. Target Population and Frames

The target population denotes the population from which inference is required. In most survey contexts it is generally not possible to reach or cover the entire target population. We may, for example, have to drop some segments of it, for example, not covering specific areas of a city for security reasons when trying to enumerate the homeless population. Once the target population is defined then it is necessary to determine a means of reaching this population. This is what is referred to as a sampling frame (or the frame). This might be a list of persons in the target population, a list of shelters and/or service providers through which homeless people might be reached. Given that a multi-frame approach is more likely to be adopted for the pilot survey it is obvious that the potential sampling frames do not meet all these criteria.

1.1 Target Population

The population of interest for this survey is individuals who are absolutely homeless according to the 1987 United Nations definition – i.e., individuals or members of a family who either have no housing at all or are staying in temporary forms of shelter – within certain specified Canadian cities. The homeless population staying temporarily in private residences of friends or acquaintances will not be covered by this survey. On any particular night³ the population of interest can usefully be divided into two distinct components – the shelter component and the street component. The shelter component means homeless individuals who are staying in some form of temporary shelter such as shelters for the homeless. The street component refers to people staying overnight literally on the street or in other locations not intended for human habitation. The street component can be further subdivided into those who have some contact with services such as soup kitchens and drop-in centers, and those who do not have any such contact. Thus in summary two components and three sub-populations are being discussed – the shelter and street components and the shelter, services and street sub-populations.

More specifically, the target population includes:

- Individuals in shelters (establishments for persons lacking a fixed address)
- Individuals in motels rented by a municipality on a temporary basis
- Individuals in shelters for youth/youth in crisis
- Individuals in shelters for women in crisis and their children
- Individuals in shelters for partners who have been abused such as transition houses
- Individuals in shelters for ex-inmates

³ A one-day reference period is assumed. Further discussion on the reference period is presented in section 3 of this report.

Individuals lacking a fixed address that used soup kitchens and drop-in centers during the reference period

Individuals lacking a fixed address that live on the street, or in abandoned buildings, alleys, under bridges, in parked cars, underground parking garages, parks, bus shelters, subway stations or other locations not intended for human habitation

The target population excludes:

Individuals staying temporarily in the private homes of a friend/relative

Individuals in a general hospital/emergency room of hospital

Individuals in other types of hospital/medical institutions/nursing homes

Individuals in establishments for children/minors (court, assistance services)

Individuals in establishments for delinquents and young offenders

Individuals in penal/correction institutions/jail

As some of these inclusions and exclusions are by definition and are quite arbitrary there must be discussion and agreement amongst all interested study parties.

1.2 Frames

Given that the homeless population can be divided into three sub-populations (shelter, services-street, and pure-street population), Mantel and Yung recommended investigating the adoption of a multi-frame approach. With this approach, we can construct sample frames of the homeless population and develop sample designs for each of these sub-populations. Since each sub-population will have its own sampling frame⁴, the sample design adopted for each sub-population can be different depending on the particular characteristics of that specific homeless group. Unfortunately, this multi-frame approach does create some problems in terms of overlapping units. It is possible, and probably likely, that some individuals will be covered by more than one sampling frame. For example, an individual spending the night in a shelter may have visited a soup kitchen during the day and thus would also be covered by the shelter and the services sampling frames. In order to avoid double counting this individual, some method of unduplication needs to be developed. This method could be as simple as asking the respondent if he/she had used a shelter/soup kitchen during the day or as complicated as deriving a unique identifier for each respondent and using this identifier to unduplicate the collected data. Reliably identifying this overlap is likely to be one of the most difficult challenges facing the survey, especially if more than two frames are used.

1.3 Relationship between Target Population and Frame

In the multi-frame approach, the target population might be broken down into 6 distinct (non-overlapping) sub-populations as described and shown in figure 1 below:

Sub-population A: Refers to the “shelter” homeless population staying in some form of temporary shelter who have no contact with services such as soup kitchens\drop-in centers during the reference period. – In-scope population

Sub-population B: Refers to the “shelter” homeless population staying in some form of temporary shelter who have some contact with services such as soup kitchens\drop-in centers during the reference period. – In-scope population

⁴ It is possible to think about potential survey design options where two or more sub-populations of the homeless population are sharing the same sampling frame. For example, an area-site frame could be used to survey both the services-street and the pure-street sub-populations. This will be discussed further in section 2.

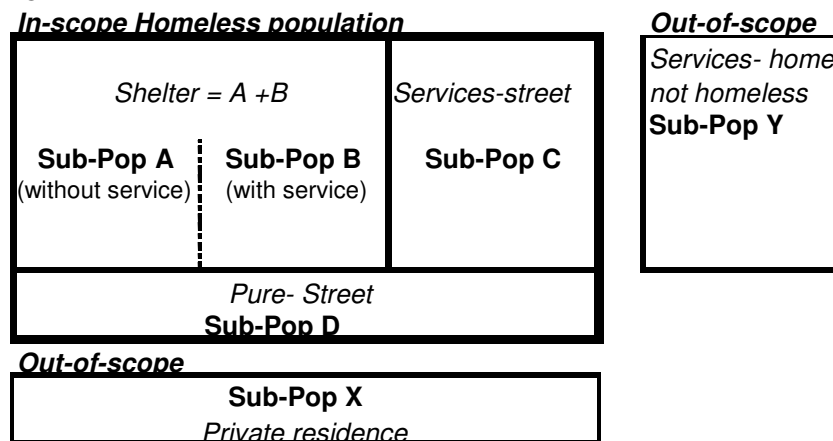
Sub-population C: Refers to the “Services-street” homeless population that do not stay overnight in shelters but who have some contact with specific services such as soup kitchens\drop-in centers during the reference period. – In-scope population

Sub-population D: Refers to the “Pure-street” homeless population that do not stay overnight in shelters and have no contact with services such as soup kitchens\drop-in centers during the reference period. – In-scope population

Sub-population X: Refers to the “Out-of-scope” homeless population staying temporarily in private residences of friends or acquaintances during the reference period. – Out-of-scope population

Sub-population Y: Refers to the “Services-home” population that do not stay overnight in shelters but who have some contact with specific services such as soup kitchens\drop-in centers during the reference period. – Out-of-scope population

Figure 1



In summary,

Shelter population: Equal to sub-populations A and B. If an individual uses a shelter in a given day, this individual belongs to the shelter population (even if services are also used as with sub-population B).

Services-street population: Equals to sub-population C

Pure-street population: Equals to sub-population D

Note: Sub-populations X & Y are out-of-scope

Here is a list of potential sampling frames to reach these sub-populations:

Frame 1: “Shelter” frame: list of shelters for a particular area (e.g., a particular urban area). This frame could be used to estimate sub-populations A and B.

Frame 2: “Services” frame: list of service providers (e.g., soup kitchens, drop-in centers, etc.) for a particular area. This frame could be used to estimate sub-populations B, C, Y.

Frame 3: “Area/site” frame: list of areas or sites for a particular urban area. This frame could be used to estimate sub-populations A, B, C, D, X and Y.

The correspondence between the target population and the sampling frame is summarized in Figure 2 below.

Figure 2

	Frames		
	Shelter	Service	Area
Population Covered	Sub-Pop A	Sub-Pop B	Sub-Pop A
	Sub-Pop B	Sub-Pop C	Sub-Pop B
		Sub-Pop Y	Sub-Pop C
			Sub-Pop D
			Sub-Pop X
			Sub-Pop Y

2. Reference period

What is the best choice for the reference period for this type of survey? Is it a single night? A week? A month? The choice of the reference period for the survey will impact the composition of the population in at least two ways:

- (I) It will affect the overall population of homeless individuals, since it will include any person who is homeless at any time during the reference period and, similarly,
- (II) The relative contribution of each sub-population (shelter, services-street, pure-street) to the overall population of homeless will also be affected. As an example, let’s assume on day 1 that shelter, services-street and pure-street populations represent respectively 50%, 35% and 15% of the overall population of homeless in a given area as shown in Table 1. All homeless individuals that belong to the services-street and pure-street sub-populations on day 1 that use shelter on day 2, move into the shelter sub-population for the two-day period. Similarly, individuals in the pure-street sub-population on day 1 that used some kind of services on day two would move to the service-street population for the two day period. In other words, to be considered part of the pure-street sub-population an individual could not use shelter and have no contact with services for the whole reference period. In summary, the relative contribution of the shelter sub-population to the overall population in a particular reference period will always be smaller to its relative contribution in the same reference period plus one day. However, it is very difficult to predict at this point this marginal daily rate as well as the rate and/or the sign of the change for either the services-street or the pure-street sub-populations.

Table 1

Population	Length of the reference period (days)						
	1	2	3	4	5	6	7
Shelter	50.0%	52.5%	54.2%	55.4%	56.4%	57.3%	58.0%
Service	35.0%	33.8%	32.9%	32.3%	31.8%	31.4%	31.0%
Street	15.0%	13.8%	12.9%	12.3%	11.8%	11.4%	11.0%

Moreover, the use of an extensive reference period will increase the probability of overlap between frames and will require the development of more complex unduplication methods. Finally, the use of longer reference periods could introduce a potential memory problem for many respondents. This last assumption however may have to be tested in the field. For all these listed reasons, it is proposed to use a single night reference period for the HRDC/STC study.

The current section attempted to clarify the objectives of the study, the target population and frames, as well as the reference period that could be used. As the mandate of STC in the second phase of this project is to investigate the feasibility of conducting a pilot survey for the street component of the homeless population the next section will focus on this component and its two sub-populations, the services-street and pure-street sub-populations.

Section 2: Street Component of the Homeless Study

Objectives

The general objective of the ‘street component’ study is to get a point-in-time measure of those individuals who have no housing at all. However, analogously to the overall Homeless Study, it is necessary to clearly define the target population, the potential frames as well as their relationship in order to facilitate further discussion about survey design options. Discussion about sampling strategy options for the street component, data collection and field operations and other considerations (including data analysis concerns) follows.

2.1 Street component

As mentioned previously, the street population refers to people staying overnight literally on the street or in other locations not intended for human habitation. On any particular night, the street population can be further subdivided into two sub-populations as shown in Figure 1 of Section 1.3, namely:

- Individuals lacking a fixed address that used soup kitchens and drop-in centers but do not use any shelter accommodation during the reference period (i.e. “Services-street” –Sub-population C)
- Individuals lacking a fixed address that live on the street, or in abandoned buildings, alleys, bridge, parked cars, parks, bus shelters, subway or other locations not intended for human habitation during the reference period (i.e. “Pure-street” –Sub-population D)

In order to facilitate discussion, the “pure-street” population will be further subdivided into those who had some contact *from time to time* with services (except on the reference period) over a one month period⁵, and those who do not have any such contact as shown in Figure 3. The first group will be called the “*Temporary Pure-Street*” population while the second group will be labelled the “*Permanent Pure-Street*” population. More specifically, the definitions of each sub-population of the street components of the homeless population (that do not stay overnight in a shelter) are:

Sub-population C: Refers to the “Services-street” homeless population who have some contact with specific services such as soup kitchens and drop-in centers during the reference period. – In-scope population

Sub-population D₁: Refers to the “*Temporary Pure-Street*” homeless population who have no contact with services such as soup kitchens and drop-in centers during the reference period but had some contact over the last month. – In-scope population

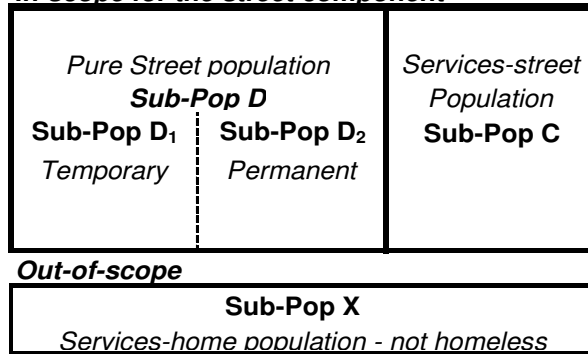
Sub-Population D₂: Refers to the “*Permanent Pure-Street*” homeless population who have no contact with services such as soup kitchens and drop-in centers over the last month including during the reference period. – In-scope population

Note: Sub-population D defined in section 1 is equal to Sub-population D₁ + Sub-population D₂

⁵ This period is arbitrarily fixed here at one month. An optimum period will need to be established.

Sub-population Y: Refers to the services “home” population who have some contact with specific services such as soup kitchens and drop-in center during the reference period.
 -> Out-of-scope population

Figure 3
In-scope for the street component



2.2 Survey Design Strategy for the Street Component

Mantel and Yung in their paper, *Sample Design Options for a Survey of Homeless in Canada* recommended the investigation of various multi-frame approaches in order to survey the individuals that belong to the street component of the homeless population.

Option 1 proposes to use a two-phase strategy to enumerate the street component of the homeless population, to collect more information about homeless individuals and to assess the quality of the first-phase census coverage. In the first phase, an area/site frame will be used in order to enumerate all homeless individuals that do stay overnight on the street on a given night while a service frame is intended to be used for the second study phase on the following day.

Other multi-frame strategies could be used to survey separately the “Services-street” and the “Pure-street” sub-populations. For the “Services-street” component, only one survey design approach is retained while three different strategies are presented for the “Pure-street” sub-population. In other words, each of the following three options combines the single “Services-street” strategy with each of the three approaches proposed to survey the “Pure-street” sub-population.

The four proposed options to survey the individuals that belong to the street component of the homeless population are summarized in Table 2 and are further discussed in section 2.2.1 that follows.

Table 2. Street Component Sample Design Options

OPTION #	Services-Street	Pure-Street
Option 1	Phase 1: Area/site frame – Enumeration Phase 2: Service Frame – Characteristics & coverage	
Option 2	Service Frame	Area/Site Frame
Option 3	Service Frame	Service Frame/external information
Option 4	Service Frame	Model/External information

2.2.1 Option 1: Area Frame for both “Services-street” and “Pure-street” Populations

The first option proposes a two-phase strategy. The first phase of the survey would enumerate both sub-populations of the street component i.e., the “services-street” and the “pure-street”. The second survey phase proposes to collect detailed information about homeless individuals and at the same time, to provide a qualitative measure of the first phase census coverage.

A) First phase

For the first phase, an area and/or site frame could be used to survey the street component of the homeless population (i.e. both “Services-street” and the “Pure-street” sub-populations) on the given night. More specifically, a census of homeless individuals would be done in the area/sites on a given night to determine the size of the homeless population in each area/site. Only a minimum number of questions would be asked to individuals to establish their “homeless” status and to collect some basic demographic information such as age and sex. In particular, questions about where they are going to sleep during the night and how many consecutive nights they have spent on the street (split between “services-street” and “pure-street”) are probably the only two important questions at this stage. This first phase is, in many ways very similar to the approach adopted by the city of Calgary⁶ in their last studies.

The determination of which areas to survey as well as the creation of the site list, if required, would need to be established through focus group sessions with various “knowledgeable individuals”, defined as social workers, shelter and service providers, policemen, etc. in each surveyed city. The focus groups would also be used to obtain preliminary estimates of the magnitude of the street population within each city to assess the need of conducting the street component survey within that city. Pre-determined criteria would be needed to guide these decisions. Enumeration area⁷ census maps could be used to completely divide the urban area into mutually exclusive an exhaustive areas (including the identified potential sites) and to ensure that the workload is most efficiently divided amongst the interviewers.

Area versus Site Frame for Phase 1

As part of the multiple frame approach, either area and/or site frames could be used in the first phase survey of Option 1. A site frame is a list of potential places where homeless individuals could spend the night, e.g., in abandoned buildings, under bridges, in parked cars, in bus shelters etc. This list would have to be built before the survey takes place using advice from “knowledgeable persons”. Moreover, this frame has to be maintained and updated for each subsequent phase or replication of the survey since these sites could vary over time, especially for large urban areas. Finally, additional information about each site is very likely to be required in order to improve the survey design. On the other hand, an area frame that consists of enumeration areas could be used without the need of secondary information. Table 3 shows some advantages and disadvantages of site and area frames.

Table 3. Some Advantages and Disadvantages of Site and Area Frames

⁶ John de Linde, The City of Calgary: Homelessness -Year 2000 Count.

⁷ The enumeration area is the smallest geographical area used by census. Census Enumeration Area (EA) maps are available for the 57 cities identified by HRDC.

	Site Frame	Area Frame
Advantages	More suitable for smaller urban areas	More suitable for larger urban areas
	More efficient survey design	Easy to develop and maintain
	Better use of resources	Complete coverage (no missing sites)
	Quicker data collection	Less costly to develop
	Less costly to survey	Minimum knowledge necessary
		Stable over time
Disadvantages	More costly to develop	Less efficient survey design
	Maintenance/updating required	More costly to survey
	Possibility of coverage errors (missing sites)	Extensive use of resources
	Additional information required	
	Dependent on advice of 'knowledgeable' individuals	
	Changes over time	

In light of the previous discussions, it is proposed to use an area frame for high-density street populations where multiple potential homeless sites could be found in a well-defined geographical area, such as downtown sections of large cities. In addition, a site frame could be used (if necessary) for other less densely populated areas within the same city where a limited number of other potential sites were identified by “knowledgeable” individuals.

B) Second phase

This sample would be drawn according to a two-stage probability sample design. At the first stage, a sample of service locations will be selected with probability proportional to their size, i.e., the total number of clients served, while at the second stage, a sample of individuals will be sampled according to pre-defined selection procedures. In this second phase, in-depth interviews would be conducted with selected homeless individuals to collect more individual/socio-economic information and to get a qualitative measure of the coverage of the first enumeration phase. As the “Pure-street” population is, by definition not reachable through the service frame, no in-depth information will be available unless it is decided to collect more detailed information in Phase I.

C) Precision and Sample Size for Phases I & II

Given that a census is used to enumerate the “street” component, the total homeless count will have no sampling error i.e., no error due to the fact that only a sample is taken from the population. However, some non-sampling error would occur in the first phase, e.g., some individuals could be missed during the census or some sites could be missed in the site list. The second phase of the survey could provide some quality indicators of the first Phase census coverage.

As Phase II involves sampling homeless individuals from the service frame, this phase will have sampling error since only some units of the target population are observed. The margin of error associated with a given estimate will depend on the variability of the characteristic in the population, the size of the population, the non-response as well as the survey design. Table 4 provides an overview of the margin of error for estimated characteristics associated with various population and sample sizes. For example, in the case of Calgary where the overall street population was evaluated to total approximately 200 individuals in May 2000, a sample of 50

would provide a margin of error of about 12% compared to 6.9% for a sample of 100 units. If we have a preliminary estimate that the number of homeless street population in large cities such as Toronto, Vancouver or Montreal is somewhere in the 2,000 and 5,000 range, we would expect a 5% sampling error attached to estimates derived from a sample of 300 units.

Table 4. Expected Margin of Error for Various Sizes of Street Population and Sample Sizes (Design effect=1.5).

<i>Example</i>	<i>Street Population (census)</i>	<i>Sample Size</i>	<i>Expected Margin of error (second phase)</i>
Calgary	200	50	15%
	200	75	11%
	200	100	8%
Montreal / Toronto / Vancouver	2000	100	12%
	2000	300	6%
	2000	500	5%
Montreal / Toronto / Vancouver	5000	100	12%
	5000	300	7%
	5000	500	5%

D) Advantages and Disadvantages of Option 1

A two-phase design has multiple advantages. First, Calgary has used the first phase methodology with success for many cycles of their “Homeless” surveys. This strategy is simple, straightforward, transparent, and easy to explain and defend. It represents an enormous advantage over the three other strategies proposed. It is also replicable across geographical units and time in order to allow comparison between regions as well as to monitor change over time. The first phase is relatively unobtrusive in the sense that the amount of information collected represents a minimum data set requirement. The first phase of the survey should also provide good precision of the homeless street-population count given that a census is taken. Only non-sampling error could impact on the overall precision of the total count. Finally, this strategy will have no overlapping problem since only one frame is used in the first phase⁸.

The second phase of the survey will provide additional information on the homeless population. More extensive and better quality information can be collected with face to face interviews in the safety and comfort of a service area compared to a street location as is proposed in Phase I. Moreover, the second phase could take place during the day, considerably reducing the security and safety issues generally associated with surveying at night. Additionally, this second phase could also be used to monitor the quality of the first-phase census coverage. The ability to collect in depth information on the homeless will be seen as a major advantage over other proposed methodologies by many data users.

On the other hand, this option has the following disadvantages:

- ◆ This strategy is labor intensive and could be very costly especially if professional interviewers are intensively involved.

⁸ The only overlap problem that could occur is between the shelter frame and the street component frame during the complete homeless study. However, this overlapping could be monitored and controlled.

- ◆ A limited amount of information can be obtained from interviews in Phase I. This approach will most likely be used solely to count people though even a minimal amount of information is necessary to determine if a given individual is part of the population of interest.
- ◆ The control of the field operations and the security of the personnel involved in these field activities are an important factor to be considered.
- ◆ Extensive cooperation of shelters, service providers, municipal governments and local police is essential for this survey option (in particular) to be successful.
- ◆ Inhospitable weather on the survey collection could impact the street component counts
- ◆ Any social intervention , e.g., special programs to relocate homeless individuals that occurs prior to the survey

2.2.2 Other Potential Options for the Street Component

Other multi-frame strategies could be used to survey separately the “Services-street” and the “Pure-street” sub-populations. For the “Services-street” component only one survey design approach is retained while three different strategies are presented for the “Pure-street” component. First, the “Services-street” approach will be described followed by the three “Pure-street” component potential strategies.

A) Services-street Strategy

Individuals from the “Services-street” sub-population can be contacted through the services that they use. The sampling unit for a survey of the ‘services’ population is the recipient of the service itself, for example, a person served a meal in a soup kitchen, or a person found in a drop-in center at a particular time. A sampling frame could consist of a complete list of service providers. It may, in fact, be better to consider this as two separate frames – one of soup kitchens and the other containing all other service providers. Note that shelters that also serve meals would need to be included if those meals were served to people who are not staying in the shelter. Additional information, such as estimates of the numbers of clients served, or list of service periods, could be used to improve the efficiency of sampling. The “Services-street” sample could be drawn according to a two-stage probability sample design. At the first stage, a sample of service locations will be selected with probability proportional to their size while at the second stage, a sample of individuals using those services will be sampled according to pre-defined selection procedures.

Unfortunately, this multi-frame approach creates many problems in terms of overlapping units between frames. It is possible, and probably likely, that some individuals will be covered by more than one sampling frame, for example by shelter and service frames. In addition, duplication within the service frame is another issue to be considered. For example, it is also possible that the same individual will use multiple services during the reference period – they could have lunch and dinner at the same (or at a different) soup kitchens. Special constraints need be put on soup kitchens where two or more meals are served on a given day. Finally, as was shown in Figure 1, it is possible to reach non-homeless individuals with the service frame. How likely is it to happen? We will be able to get an estimate of this hit rate (i.e. probability to reach a homeless individual in a service center) with a field test. The estimated hit rate could be used later at the survey design stage of the study in order to achieve the targeted sample size for the service component. The effective hit rate resulting from the study itself will be used at the estimation phase to get an estimate of the number of out-of-scope individuals that used the services.

Furthermore, depending on the strategy retained to survey the “Pure-street” sub-population, other overlapping possibilities are likely to occur. For example, overlapping between the service and area/site frame is likely to occur with Option 2. On the other hand, duplication between successive samples selected from the service frame must be monitored and considered if Option 3 is retained. This will be discussed in more detail later in this section.

In order to avoid double counting individuals appearing in the multiple frames, some method of unduplication needs to be developed. This method could be as simple as asking the respondent if he/she has used a shelter/soup kitchen/drop-in center during the day (and previous day for Option 2) or as complicated as deriving a unique identifier for each respondent and using this identifier to unduplicate the collected data. Reliably identifying these overlaps/duplications is likely to be one of the most difficult challenges facing the survey when using this methodology.

B) Three Strategies for the Pure-street Sub-population

By definition, the “Pure-street” population cannot be contacted through any shelter or service center on the reference period. A majority of surveys that have been done in the past have used an area/site frame to get a census of the ‘street’ population (i.e. both “Services-street” and “Pure-street” sub-populations) for practical reasons. While this strategy is probably the most advantageous option for this study, other multi-frame approaches were investigated in order to survey individuals that belong to the street component of the homeless population. Following, three different survey design options are presented for the “Pure-street” population with discussion on the advantages and disadvantages of each approach. The first option is similar to Option 1 in using an area or site frame to count and collect information about the “Pure-street” population. The second approach proposes to use successive samples selected from a service frame in order to estimate the *temporary pure-street* population in conjunction with external information that provides counts for the *permanent pure-street* population. Finally, the third option uses only the *permanent pure-street* population count provided by external information sources. This last option would essentially be used only in small/medium size urban areas where resources were limited.

For the three approaches, a preliminary list of the *permanent pure-street* population should be created in collaboration with ‘knowledgeable’ individuals from the surveyed area. This could be also done through focus groups. Depending on the selected survey design, this external source of information could be used as an “estimate” of the *permanent pure-street* population for the survey or as an independent source to be used to compare survey estimates. Since this specific component of the street population could be very difficult to reach and costly to survey, this external source of information could represent a valid compromise for that specific segment of the population. In particular, if that specific population represents a very small proportion of the overall homeless population⁹, it would be much more cost-efficient to use this external source of information.

B1) Option 2: Area Frame and/or Site Frame

Similar to Option 1, an area/site frame population could be used in order to enumerate homeless individuals that belong to the “Pure-street” sub-population. In order to get the count and the characteristics of the “Pure-street” population, a census of homeless individuals would be done in the areas/sites on a given night to determine the magnitude of the homeless population in each

⁹ This assumption should be verified in some way by the field test or through consultation with sponsors and knowledgeable individuals.

area/site visited. After or during the enumeration (depending on the specific situation), interviews could take place with randomly selected individuals in those sub-areas/sites:

- ◆ to determine the proportion of the homeless population that falls into each of these categories: “Services-street” population, “Pure-street” population and the out-of-scope population;
- ◆ to collect additional information on the characteristics of the “Pure-street” population component (only a sample of the pure street population will be interviewed “in-depth”).

This second option differs from the first one by the basic objective of the area/site survey. The first option proposes to count all street individuals while the second option is restricted to the “pure-street” sub-population. The surveyed areas or the site list could be different according to the option selected. In addition, the quality of the coverage of the area/site survey could be monitored with this option. However as with the first option, Option 2 represents in some ways a classical methodological approach. However, though this strategy has numerous advantages such as simplicity, reliability and reproducibility, it will remain very labor-intensive for many municipal jurisdictions. In addition, the choice of this option for enumerating only the “Pure-street” population could be difficult to defend given the additional marginal cost of surveying the “Services-street” population at the same time as with Option 1.

B2) Option 3: Successive Samples Selected from a Service Frame

By definition, the “Pure-street” population cannot be contacted through any shelters or services on the reference period. However, by asking questions about the use of services for the last few days using different samples over two or more consecutive days, it may be possible to find and get an estimate of those who use services occasionally but didn’t on the reference period. The idea here is that in the long run it will be possible to reach a large proportion of the “Temporary Pure-Street” population through the “Services-street¹⁰” sampling frame that represents a complete list of service providers.

Let’s take an example to illustrate the strategy. Suppose a first sample (sample 1) of individuals from the soup/drop-in centers will be asked about the use of services for the current day (reference period), previous day and the day before. Let’s apply the same survey design for the three following days (sample 2, 3, & 4). Table 5 shows the output of this design. In using this strategy, it should be possible to estimate the number of ‘Temporary Pure-Street’ homeless individuals for a given reference period. For example, on Reference period 3, $x_{-1}^4 = A+B$ where A equals the homeless “Temporary pure-street” and $(x_{-1}^4 - A)$ equals “not homeless pure-street (ex. services, shelters, out-of scope)” on reference period 3 using the sample of the following day. In addition, $x_{-2}^5 = C+D$ where C equals the homeless “Temporary pure-street” and $(x_{-2}^5 - C)$ equals “not homeless pure-street” on Reference period 3 using the sample of the second following day.

Table 5

	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6
Sample 1	x_{-2}^3	x_{-1}^3	x_0^3			
Sample 2		x_{-2}^4	x_{-1}^4	x_0^4		

¹⁰The shelter frame could also be used in conjunction with the service frame if required.

Sample 3			X_{-2}^5	X_{-1}^5	X_0^5	
Sample 4				X_{-2}^6	X_{-1}^6	X_0^6

According to this survey design, it could be possible to improve the services-street survey estimates and to estimate the number of “*Temporary Pure-Street*” homeless individuals using the relationship (links) between samples¹¹. However, additional investigation will need to be made in order to evaluate and determine what is the best method of estimation to use with this survey design in light of other potential problems, such as, respondent recall problems or change in the population from day to day. Notwithstanding the fact that the estimation methodology could be complex, this proposed approach has many advantages over other designs, as it would be possible to avoid directly surveying the street component using an area/site frame. Furthermore, more extensive and better quality information could be collected with face to face interviews in a service area as compared to a street location as is done in the first phase of Option 1. What’s more, all data collection could take place during the day, considerably reducing the security and safety issues generally associated with the ‘classical’ late night approach. Implicitly, this approach also means that it will be necessary to rely on an “estimate” of the “*permanent pure-street*” population created in collaboration with “knowledgeable” individuals. This adjustment should have a minimal impact on the final estimate. Option 3 is more likely to be used in medium to large urban areas where the absolute number (and the proportion) of the “*pure-street*” homeless population probably will be relatively important compared to smaller urban areas where this proportion will likely be very small in many instances.

B3) Option 4: No Survey

This last option also proposes to use the “*permanent pure-street*” population count provided by external information sources as was done in Option 3. However, two different approaches could be used to estimate the “*temporary pure-street*” population. The first proposes to adopt a model-based approach in using the data collected by Option 3. Modelisation of “*temporary pure-street*” population estimates could be performed taking into account the area/urban size and structure. The second approach proposes to use only the “*permanent pure-street*” population count provided by external information sources (no estimation of the “*temporary pure-street*” will be attempted). The first approach could be used in medium size urban areas while the second strategy could be applied in small urban areas where the absolute number of pure-street individuals is very small. In summary, Option 4 could be used only in small to medium sized urban areas where resources could be limited.

2.2.3 Advantages and Disadvantages of the Four Proposed Options

Table 6 shows some of the advantages and disadvantages for each of the four proposed options for surveying the “*pure-street*” population.

Table 6

¹¹ In theory, the estimate of *Temporary Pure-Street* for the reference period 5 is A+C- (A1C) but additional investigation needs take place in order to account for the survey design.

	Option 1 Area/site Only Phases 1 & 2	Option 2 Services Frame & Area/site	Option 3 Services Frame & Successive Service samples	Option 4 Services Frame & External information & modelization
Plus	Classical approach	Classical approach	Less costly to survey	Easy to explain (except for the model component)
	Easy to explain/defend	Easy to explain/ defend	Less extensive use of human resources	Minimum cost
	Complete coverage (no missing sites)	Very large coverage	Needs less survey control	Minimum knowledge
	Not costly to develop	Not costly to develop	Security acceptable (interviewing in daytime at service locations)	Negligible impact on survey estimates at national/provincial level
	Minimum knowledge One frame and 1 phase	Minimum knowledge 2 frames & one phase	More data collection possible	
	Very reliable estimates	Reliable estimates (service sample)	Better quality control/ Better response rates	
	Amount of data collected	Amount of data collected		
	Easy to replicate	Easy to replicate		
	Easy to compare estimates by region	Easy to compare estimates by region		
	Partially tested by Calgary			
	Coverage check			
	No overlapping frames with one phase			
Minus	Less efficient survey design	Estimation of the service-street sub-pop. requires additional info.	New approach	Reliable estimates?
		More costly to survey- Data collection	More difficult to explain. Estimation could be very complex	No variability /measure of quality
	More costly to survey- Data collection	Extensive use of human resources	Coverage issues	More difficult to defend
	Extensive use of human resources	Needs considerable control	More costly to develop	
	Need considerable control	Security issues	Needs prior knowledge	
	Security issues exist for first phase	Response rates/ data quality- first phase	Use of external information	
	Response rates/ data quality–first phase	Collaboration from street worker essential	Need additional information for estimation purposes	
	Collaboration from street workers essential	Overlapping frames		
		Similar to option 1 with less advantages for about the same cost		

2.2.4 Recommendations

Based on a review of the options discussed some recommendations will be made for the pilot survey for the street component as well as a general proposal for an overall national homeless survey.

A) Recommendation for the Pilot Survey

The objective of any pilot survey is to test on a subset of the sample units the instruments and procedures developed with the ultimate goal of standardizing these and replicating them across all units for the complete study. To this end the units selected for the pilot survey should be representative of all units that would be found in the complete study. If this is not the case then unplanned situations will arise at the time of the complete study and, if possible, will have to be dealt with at that time. This may introduce procedures, which are not standard across all the target sample units and may lead ultimately to bias in the estimates.

HRDC has preliminarily identified 57 urban areas across Canada that were thought should be part of a comprehensive Survey of the Homeless in Canada. Each province and territory has at least one urban area identified. Upon reviewing these 57 urban areas the question which comes to mind is, do all these areas have a (significant) street component? Thus it is believed that the first task will be to group the 57 cities into two categories, one where a street component should be conducted and the other where it is not necessary and where no significant impact on the total homeless count is expected. This grouping can be achieved through consultation with “knowledgeable” individuals in a focus group setting. What remains to be determined is how many focus groups would be necessary to obtain an acceptably accurate division of these urban areas. Perhaps one focus group is all that is necessary. Perhaps expertise is confined to a province or in the worst case scenario to each urban area. To be able to confidently advance to a national Survey of the Homeless in Canada, the pilot survey must include all types of units that will be covered in the national survey. To this end, at least one urban area from each of the two categories should be selected for testing in the pilot survey.

Aside from an overall count of the homeless street component, one other objective of the pilot survey will be to determine the effectiveness of using a service frame for estimating the street component as is being done in Phase 2 of the proposed option. Acceptance of the use of a service frame for enumerating the “pure street” component would be a major breakthrough in attempting these counts. The benefits are many fold: quality control and more extensive data; dramatic cost savings; less intrusive for homeless; and increased safety for interviewers.

At this juncture it is uncertain whether one pilot test will be sufficient. More than one may be necessary, with each building on the last. This question will have to await the results of the first pilot test.

B) Recommendation for the Survey of Homeless in Canada: An Integrated Approach

For the national Homeless survey, an “integrated design” is proposed. It adopts the most cost-effective survey design specific to each enumeration or urban area targeted, taking into account the potential homeless population size, the structure of various components of the homeless population and the survey frequency. The adoption of specific survey design methods has as its goal the maximization of precision while minimizing costs of data collection. To ensure data compatibility, standard and consistent survey concepts, procedures, and content will need to be enforced across all surveyed areas and components. In addition, standard survey processing and estimation methods will also be the prominent integrating features underpinning this proposed integrated design for enumerating the homeless population.

This report will assume that a frame for the shelter component will exist at the time a national survey is undertaken. Different sample design strategies are proposed for the two categories of urban areas, those where a street component is thought necessary and those where it is not

thought necessary. Table 7 summarizes the approach recommended by category of urban area. Furthermore, other Canadian cities should also be included in the survey to get (complete) national coverage of the Homeless population in Canada. Information on homeless counts for these additional urban areas and for Category 2 urban areas is recommended to be collected through the use of administrative records (e.g., HIFIS administrative records for shelters).

Table 7. Category of Urban Area by recommended frame/option

Component	Category 1	Category 2
Street	Option retained	None
Shelter	Census/HIFIS listings	Census/HIFIS listings

Given that a census will be used for the shelter as well as for the street component in the pre-selected cities, the total count of homeless individuals will have no sampling error while some sampling error will be attached to the more detailed characteristics collected in the second survey phase. However, non-sampling error is likely to occur given that some homeless individuals will most likely be missed by the census approach. As the shelter component count is expected to represent a very large proportion of the total homeless population in Canada, it is reasonable to believe that the potential non-sampling error attached to the street component count would have negligible impact on a comprehensive national Homeless count.

2.3 Data Collection

Assuming Option 1 is retained a one night reference period is recommended for Phase I. For the second phase a one night reference period is not as crucial. The data collection duration could be based on the number of interviews required as well as other operational considerations. If it is determined necessary to collect data beyond a single day one possibility is to divide the sample of service locations into a number of individually representative samples that could be covered on different days. In such a scenario, as it would be possible for the same person to be in the sample on two different days it would be necessary to decide how to treat these cases and to identify them in the field. In addition, it would also be possible that the composition of the overall homeless population had changed over the reference period. For example, the relative proportion of the services component could increase at the end of the month or the relative proportion of the street component could decrease on cold or rainy days. The relative movement of the homeless population between its different components should be addressed and discussed prior to the field test. Moreover, some “births” and “deaths” could arise from day to day in the homeless population. Notwithstanding, what is the magnitude of this potential change? Is it more important at the end and at the beginning of the month? Again, this is something to ask “knowledgeable” persons or something to assess in the pilot test.

2.3.1 Questionnaire Content

Unlike some surveys where very detailed information is asked to individuals, it is proposed to ask only what can be tolerated and to tailor data collection to the conditions under which it is collected. For example, Phase I will be conducted at unorthodox times and places, therefore only very basic information can be observed and recorded by the interviewer. On the other hand, data collected in the convenience of a ‘service provider’ area in phase 2 will favour the collection of more elaborate information. It is for this reason that it is proposed to restrict the question content for the first phase to two questions (in addition to age and sex) while the second phase would expect to gather more detailed information about individuals and family. Such a strategy will improve response rates, minimize respondent burden, and most likely yield better quality data. Based on Statistics Canada policy on informing survey respondents, respondents would need to

be informed about the purpose of the survey, the authority under which the survey is taken, the mandatory or voluntary nature of the survey, confidentiality protection and any data sharing agreements in place. Following these introductions, a hierarchy of information collection is proposed for phase 2 that could include:

- Interview Conditions (geographical code of enumeration area or site, total count of homeless individuals found at that location, date and time of day, weather)
- Gender, age, family status (and if applicable, number, sex and age of dependants)
- Homeless status on the reference period and any other screening questions to determine eligibility for homeless enumeration
- Ethnic background (visible minority status/aboriginal status/citizenship status, etc.)
- Flows into and out of homelessness (duration of current episode, number of episodes, reason for current episode, mobility during current episode (inter/intra urban, provincial), conditions prior to current episode)
- Uses of shelter/services (during the reference period)
- Other relevant characteristics such as health, education, labor force participation, income, etc.
- History of medical treatments received
- Psychological Status
- Perceived bridges to escape homelessness
- Dependencies
- Legal Issues

The questionnaire should use simple concepts and wording. To that end, focus group testing should be carried out in order to clarify and test the questionnaire prior to a pilot-test. In addition, confidentiality issues will also have to be looked at closely.

2.3.2 Field Operations and Issues

There are several issues surrounding data collection. Some are unique to Statistics Canada - its regional infrastructure organization, unionized interviewer workforce and its near term projected workload. The principal issues are listed and discussed below:

- (1) Logistics of data collection: if a 'classical' approach is adopted then there is the real likelihood of data collection at non-traditional times and in non-traditional locations. Precedence for surveys of the street component of the homeless has favoured data collection when activity is at its low point for a day – between the hours of 2:00 A.M. and 5:00 A.M. and at the locations where the “pure street” homeless population are found – in parks, under bridges, in abandoned buildings, etc. This presents a number of problems for a traditional organization like Statistics Canada. STC interviewers tend to be predominantly female and older. Such an undertaking would tend to favor a younger interviewing staff, those better able to stand the rigors of getting to and around these unorthodox sites. Because of the nature of the interviewing, interviewers would have to be given the option of ‘volunteering’ to participate or not, and for those that do, it would be necessary to provide adequate security, perhaps an escort provided by an off-duty police officer or perhaps a doubling-up of interviewers. Costs would be dramatically higher than for a regular survey because they probably would be at over-time pay compensation levels. There is also the issue of whether the interviewer unions would ever endorse such an undertaking even with the increased security and compensation.

- (2) There is also the question of whether sufficient expertise could be established and built-up consistently in the regional offices concerning the myriad of connections that would be necessary in each urban area for such an undertaking. In a 'community undertaking' of a study of this nature there generally are at least three benefits of 'community involvement' that would not be achieved with only STC organization – (1) by involving the community more likely to obtain a valid reliable count (2) they serve as a communications tool, informing the community of the study objectives, it's schedule and generally serve to make the community aware of the magnitude of the problem and (3) they unite the community in coming to a solution for the problem.
- (3) Given that Statistics Canada doesn't do the data collection itself and contracts it out – to whom? Is there a private firm with a national infrastructure in place sufficient to handle the collection? If volunteers are used as was done in the Calgary study how could they be organized consistently to ensure a 'standard' data collection protocol? One thing seems clear – it probably would be impossible to use unpaid 'volunteers' as was done in Calgary. There are great advantages as noted above, in using community professionals who work in the social arena but to ensure a consistent application of the methodology in conducting a national count it is recommended that their respective organizations be reimbursed.

2.4 Concluding Remarks

Homelessness is a complex issue that cannot be discussed or attacked with a single approach. The present proposal presents various options for looking at the UN definition of absolute homelessness and its two dimensions of shelter and street. A point-in-time measure of the homeless situation is proposed. Previous studies have used a measure of prevalence that attempts to count the number of unique homeless individuals over a certain period of time. Such an approach is generally more appropriately used for only the shelter component where it is potentially easier to keep track of homeless individuals that used the shelter system. In any event, either of these two measures is sufficient by its own to fully address the homeless issue in Canada.

In the end, it has to be remembered that the survey results are only one of the tools that should be used to analyze and understand the Canadian homeless situation. For example, the overall shelter capacity overtime, the occupancy rate and the number of homeless individuals turned away¹² for a given urban area are other dimensions to be considered in a complete analysis of the homeless issue. In addition, this survey does not include individuals who are at risk of becoming homeless, i.e., the focus is the absolute homeless and has excluded the relative homeless. For example, individuals that pay too much of their income for rent and/or who are living in unsafe, inadequate or insecure housing as well as individuals staying temporarily in the private home of a friend/relative because they have no permanent home of their own. These facts have also to be considered when analyzing and interpreting the survey results. For example, an increase in the overall shelter capacity that is currently saturated in a given urban area could lead individuals currently living in inadequate housing or those living temporarily with friends and relatives to move into shelters as soon as they become available.

¹² People not admitted to shelter for various reasons.