Session 1A:

Critical Issues in Counting Homeless Persons

What’s Behind the Numbers? Definitional Issues in Counting the Homeless
David S. Cordray and Georgine M. Pion

Federal Homeless Information Needs and Local Practice
Christopher Walker
What’s Behind the Numbers? Definitional Issues in Counting the Homeless

David S. Cordray
Vanderbilt University

Georgine M. Pion
Vanderbilt University

Abstract

Any estimate of the number of homeless persons involves several definitional issues, including the underlying conceptual definition of “homelessness,” the intended use of and rationale for the count, how the conceptual definition is translated into operational procedures, and methodological choices. These issues are at least partially responsible for the variation in existing estimates of homeless individuals. To best serve multiple constituencies and demands for information, the adoption of a broad definition is recommended. Furthermore, it is necessary to develop and report estimates for specific policy-relevant subgroups of homeless persons, as well as to develop a family of research studies that can yield an understanding of the conditions facing the homeless and identify “markers” for those who are at risk of becoming homeless.

Introduction

Definitional issues pertinent to counting homeless individuals, including who has been counted and who could be counted, play an integral and controversial role in shaping estimates of the homeless population. At the same time, an exclusive focus on these issues can prove unproductive, often serving as just one more interesting intellectual exercise. Such a concern with definitions rapidly becomes devoid of meaning unless it is linked explicitly to practices of counting homeless persons. Thus, in delineating issues embodied in the task of defining homelessness (e.g., “Who could or should be counted?”), a broader approach that examines the interplay among definitions, methods, and reasons for counting the homeless is essential. As with any attempt to count or estimate the magnitude of a phenomenon, an understanding of estimates of the size of the homeless population requires information about how the estimates were produced, who is and is not represented, and why the estimates were undertaken in the first place. By using this broader lens (one that focuses on what’s behind the numbers), it is possible to better understand the points of convergence and divergence in estimates produced by different investigators, and to identify
methodological implications for improving our knowledge about homelessness, along with the quality of information that is generated.

What’s behind the numbers?

As evidence in both scholarly journals and the popular press proves, considerable variation exists among national estimates of homeless individuals in the United States.¹ A cursory examination of the bases for these varying estimates highlights the fact that definitional issues are partly responsible for their diversity. However, definitional issues are only one element; therefore, it is useful to briefly review the set of factors that may contribute to the differing estimates.

Some examples

Nearly ten years ago, Hombs and Snyder, members of the Community for Creative Non-Violence, estimated that there were 2.2 to 3 million homeless persons in the United States.² What’s behind these numbers? According to Rossi, they essentially represent “extrapolations from local estimates by knowledgeable persons.”³ It appears that the definition of homelessness was left to the discretion of the 100 key informants residing in the 25 cities that CCNV contacted.⁴ Rossi further concludes that “not enough information is given about the Hombs and Snyder estimates to let us evaluate their worth.”⁵ Kondratas levels a much stronger judgment, calling the numbers a “clear leap of fantasy.”⁶ She also points to the “staying power” of these numbers and their political significance in setting the stage for further debate about the size of the homeless population.

Using similar methods (key informant plus other “count-like” strategies), the U.S. Department of Housing and Urban Development (HUD) placed the number of homeless during any given night during December 1983 and January 1984 at between 250,000 and 350,000.⁷ Although this study was conducted shortly after Hombs and Snyder’s effort, HUD’s estimate was about one-tenth that of Hombs and Snyder. What factors account for this difference? Although identifying all the disparities is arduous, HUD researchers, unlike Hombs and Snyder, did provide experts with a definition of what they meant by homelessness (i.e., individuals in shelters or public/private places not designed for human habitation), a specific time frame (one average night), and a specific period for which the basic estimates were derived.
The data collected by HUD serve as the basis for three additional estimates. Tucker, extrapolating from data gathered in the HUD study, estimated the 1987 homeless population at 700,000. What’s behind this number? Tucker’s procedure essentially involved using the estimates in “the most reliable range” obtained by HUD from service organizations for 35 individual cities, and duplicating HUD’s procedures to develop estimates for 15 additional cities; in contrast to HUD, however, he based his projected “homeless rate per thousand” for each of the 50 cities on the population of the city itself rather than the population for the Standard Metropolitan Statistical Area (SMSA). A different estimate was put forth by Freeman and Hall, who developed their own street-to-shelter ratios, applied these estimates to the HUD estimates of the shelter population, and then adjusted this total based on an assumed growth rate. Using these procedures, Freeman and Hall estimated 343,000 to 363,000 homeless persons in 1985.

Drawing on the 1984 HUD estimates, the National Alliance to End Homelessness estimated that as many as 736,000 persons were homeless on any particular night, with between 1.3 million and 2 million different individuals experiencing homelessness at some point during 1988. The Alliance’s nightly prevalence estimate was based on a presumed 20 percent annual increase in the number of homeless, stemming from the reported annual growth in the demand for shelters. The latter, in turn, was based on an estimate taken from a report by the U.S. Conference of Mayors. Thus, these three numbers are essentially estimates based on estimates.

In a separate data collection effort, the Urban Institute recently estimated that 567,000 to 600,000 individuals in the United States were homeless on any given night in 1987. The figure is derived from a probability-based national estimate of the number of homeless persons using shelters or soup kitchens in cities with populations over 100,000. Rather than assuming that all users of soup kitchens were homeless, the Urban Institute researchers screened respondents using these guidelines:

- Not having a home or a permanent place to live;
- Residing in a shelter or hotel/motel paid for by a voucher or other instrument;
- Staying in an indoor or outdoor space that was not intended for habitation; and
• Staying with a relative or friend “with whom they did not have a regular arrangement to stay for five or more days a week.”

Based on these definitions, the study found 229,000 persons homeless over a seven-day period. To develop the national estimate of 567,000 to 600,000, projections were made, using findings from within the target cities and shelter/soup kitchen locations that were sampled. Based on new information and assumptions about the distribution of homeless in other places and the nonutilization of services by homeless individuals, additional analyses have resulted in a new estimate of about 350,000 to 500,000 homeless in the United States.

In his recent book *Down and Out in America: The Origins of Homelessness*, Peter Rossi concluded that the best estimate of the literally homeless (i.e., those in shelters and on the streets) was 250,000 to 350,000, and that another four to seven million individuals were extremely poor and thus at risk of becoming homeless. This estimate of the literally homeless is based on what Rossi regards as the most reliable evidence collected. Because the majority of previously reported estimates are not actual counts, he is not sanguine about the analytic basis for his choice of the best estimate of the literally homeless. On the other hand, his appraisal of the size of the group that others have found difficult to estimate (i.e., those at risk of homelessness) is grounded in analyses of rather firm national data bases.

*Some definitional issues accounting for variation in estimates*

The examples cited above suggest several definitional issues that have been previously confronted by researchers. With regard to the estimate offered by Hombs and Snyder, we can see where definitional issues appear to have been skirted altogether. In other instances, different definitions have been adopted by different investigators. The numbers reported by the Urban Institute reflect a broader view of who is homeless (i.e., they include those who are doubled up) than those produced by either HUD or Rossi. We will refer to this as an issue associated with the conceptual definition of homelessness. Even when the definitions are the same, however, researchers possess considerable latitude in how they operationalize their conceptual definitions—a point we will revisit later. For example, selecting a specific time of year to conduct the count is an operational detail that can influence the resulting numbers. Would HUD’s figures have been smaller if the time frame chosen for
sampling had been spring or summer? This issue was explicitly addressed by Rossi and his colleagues, who incorporated into their study of the Chicago homeless provisions for estimating the use of shelters and streets in winter and fall.18

Looking across the examples, several other definitional issues emerge. The time span associated with the estimate is important. For the most part, previous estimates have referred to the prevalence of homelessness at any given point in time (i.e., on an average night), although the National Alliance to End Homelessness did attempt to project the number of individuals who would experience homelessness during a given year. Given the frequent movement of this population into and out of homelessness,19 it is not surprising that vast differences exist between the prevalence estimates and this annual incidence figure. Upon which criteria should definitions be based? Is one set of definitions preferable over another? Do some definitions bias results in a predictable fashion? In general, the answer to these questions is “It depends.”

Other reasons for variation

As previously noted, multiple factors have contributed to varying estimates. In the case of technical culprits, some definitions impose biases that tend to understate or overstate the size of the population. Yet with regard to other contributing factors, bias is not the correct attribution. Rather, the estimates are based on different conceptions of the problem or estimation task. So as not to leave the impression that estimates depend only on definitional issues, a brief review of other contributors to varying estimates appears in order.

Since estimates address different time periods, actual growth in the size of the population could certainly account for the disparities between estimates derived in the early 1980s compared to those produced for the latter part of the decade. The methodological difficulties confronting efforts to identify and count this transient and ever-changing population, along with the technical transgressions actually committed, have been admirably catalogued in the literature.20 In addition, it is clear that some methods produce more bias than others.21 Also well-articulated have been the obvious and more subtle roles played by such factors as prevailing views about the causes of homelessness.22 Time and resource constraints are culpable in terms of limiting the methodological choices that are available.23
Scratching even slightly below the surface to see what lies behind a given estimate of the number of homeless persons reveals a rather complex interplay among definitions, methods, and the reasons for counting or estimating in the first place (see figure 1).

In a nutshell, our scheme asserts that conceptual definitions of homelessness establish the foundations upon which subsequent assessments are built. Important research issues, that is, the specification of operational definitions and methods of enumeration, flow more or less directly from this initial conceptual definition. Simply stated, it is impossible to make meaningful decisions about whom to count as homeless and how to derive that estimate without a firm
grasp of the concept that one intends to measure. Of course, this assumes that an agreed-upon definition of homelessness is possible—a dubious assumption in many minds.

The authors adopt a different perspective by arguing that waiting for consensus to develop around a *singular* conceptual definition is not productive. There are simply too many political pushes and pulls associated with desires to expand or contract any definition. What is more important for researchers is that counts be reported in a fashion that allows users to understand who, on the continuum of homelessness, has and has not been included. The authors refer to this method of reporting counts as *intelligent segmentation*. That is, total estimates, regardless of the conceptual definition embraced, would be reported for meaningful subgroups (e.g., shelter residents, persons on the streets, and residents of institutions). Thus, the conceptual definition of homelessness is comprised of an ordered series of subconstructs, each with its own conceptual meaning. Throughout the articles in this journal, homelessness seems to be presented as the anchor point of a continuum that is probably best labeled as a standard of living. Although the line between those who have adequate living arrangements and those who do not is often fuzzy, the real difference among researchers appears to be the extent to which they are willing to count as homeless the persons who are further up this continuum.

There are several reasons for abandoning the idea of an agreed-upon definition for homelessness. Hopper notes that conceptions of homelessness have changed dramatically over time, and the conditions that serve as “markers” are culturally determined. This does not mean, however, that we should forsake our concerns about definitions. Rather, what is needed is a perspective that will permit us to avoid “chasing our tails” over such definitional issues as which definition is correct. As shown in figure 1, the authors depart from the customary idea that operationalizations and methods of enumeration flow directly from a singular conceptual definition (as typically embodied in classical measurement notions) by recognizing that these relationships are moderated by the notion of the intended use or the rationale for the estimate.

The reasoning behind this position is simple. A primary obstacle to defining homelessness by consensus is that multiple purposes exist for embarking on an estimation task, along with many constituencies clamoring for different kinds of information. For example, Congress may be interested in obtaining a count of all homeless individuals, and this broad ambition dictates the use of a broad conceptual definition. On the other hand, some requestors may have
more modest demands involving a specific subgroup of the homeless population. Their purposes would be adequately served by a more circumscribed definition (e.g., the literally homeless). Thus, the intended use of and impetus for asking the question “How many?” plays a role in framing the parameters for a particular study.

Similarly, the anticipated uses of a count enlarge or shrink the conceptual definition that is entertained, which in turn influences the development of operational procedures. As an illustration, if the main impetus for counting homeless persons is to discern current and future service needs, a broad conceptual definition is necessary—one that includes multiple subgroups, such as those who are literally homeless and those who are precariously housed. For planning purposes, the logic of adopting a broad definition is that it supplies the means for assessing the needs of both today’s homeless and those who are at risk (tomorrow’s potential homeless population). Long-term resource planning would be of little value if those who were vulnerable now and likely to need supportive assistance in the future were disregarded. If, on the other hand, the requestor’s needs are more discrete (e.g., knowing how many shelter beds are needed or how many meals must be prepared), only the relevant portion of the homeless continuum would be considered. The operational definition would be highly filtered in this circumstance.

Of course, it is not only the operational definition of homelessness that is altered by the breadth of the conceptual definition adopted and the uses to which the results are to be applied; the methodologies for counting are also affected (see figure 1). A broad definition in all likelihood will require the development of multiple sampling frames (e.g., for shelters, for nonconventional dwelling spaces, and for soup lines). Thus, the methodological complexity expands dramatically when national estimates, based on a broad definition, are stipulated. Time and resources often are insufficient for carrying out such ambitious enumeration efforts, so methodological trade-offs must then be made, involving the use of projection procedures and other less precise or rigorous methods. Therefore, it should be clear that methodological decisions are not independent of the impetus for the study or its scope, but rather are conditional upon these previous decisions.

**How might conceptual definitions be generated?**

The generally accepted practice in social science research is first to articulate what is intended to be measured, that is, to state the conceptual definition of the phenomenon to be investigated. For
example, studies aimed at examining the causes and consequences of chronic mental illness must clearly state what “chronic mental illness” means; similarly, efforts to investigate the relationship between stress and diathesis for the onset of depression must delineate what the term “stress” encompasses. For knowledge to most effectively be accumulated and advanced through the work of multiple and independent researchers, a consensus regarding the proper definition is ideal. At the very least, however, it is imperative that researchers be conscientious and forthright about specifying the definition(s) that guide their individual studies.

The conceptual definitions underlying previous efforts to characterize the size and composition of the homeless population often have been ambiguously or incompletely articulated. In some cases, no such definition is provided, and thus the definition may or may not be understood by the bulk of researchers and/or users of the results. In other cases, studies suffer from “semantic imprecision,” allowing interpretational degrees of freedom on the part of investigators. Thus, it comes as no surprise that no generally agreed-upon definition of homelessness exists. In fact, some would suggest that there is little merit in arriving at a definition of homelessness. For the purposes of the decennial census, the Census Bureau studiously avoided providing any definition of homelessness. As Taeuber and Siegel asserted, “the 1990 Census did not impose one”; rather, it skirted the issue by simply designating locations where and time when persons would be counted (e.g., at shelters in the evening and in the street and open public locations in the early hours of the morning). The establishment of a conceptual definition was bypassed, and the Census Bureau proceeded directly to developing an operational definition. Although it is difficult to envisage that the Census Bureau’s selection of where to look was devoid of any conceptual underpinning, however tacit, such definitional “leapfrogging” is in all likelihood common in homelessness research.

Common-sense approaches

One way to define homelessness is to rely on common sense or natural language. The mere mention of the term homeless generates a variety of images. The most obvious one is of a person who has no home. In fact, the Oxford English Dictionary’s definition of homeless is a common-sense one: “having no home or permanent shelter.”

Other sources also are instructive for obtaining insight into the meaning and circumstances customarily associated with homelessness. As shown by a casual perusal of the subject index for
newspaper articles on homelessness, journalists have contributed to the process of defining homelessness. Some headlines underscore the more standard view of the literally homeless: “Alone and Homeless, ‘Shutouts’ of Society Sleep in Doorways,” “New York City Resists State on Shelters for Homeless,” “Police Close Down Makeshift Home for Skid Row Residents.”31 Other circumstances associated with homelessness also are revealed by newspaper accounts (“A devastating typhoon . . . left 180 people dead, 370,000 homeless”).32 More recently, those who are temporarily displaced and have doubled up with family or friends have sometimes been considered homeless by the press. Other conditions highlighted by the press include families living in motels on an apparently permanent basis due to their inability to find affordable housing or to accumulate the cash deposits required for rentals, and those in which there is a significant gap between income and rents.33 While journalistic accounts can be helpful, each represents a limited segment of the homeless population. It remains unclear how these segments could be stitched together to provide a broader and more workable conceptual definition. Another strategy for exploring how homelessness has been defined, and thus who should or could be counted, is to consult definitions that have been developed previously.

Formal definitions

The *Oxford English Dictionary* is often viewed as the ultimate source for definitions, but its definition of *homeless* (“having no home or permanent shelter”) is in fact somewhat blurry. In trying to apply this definition to enumeration efforts, we would immediately have questions concerning what exactly is meant by the terms *home* and *permanent.*34

Defining *home* could be relatively straightforward, if one wished to rely on the Internal Revenue Service (IRS). In its 1990 income tax filing instructions, *home* is defined as “a house, condominium, cooperative, mobile home, boat, or similar property. It must provide basic living accommodations, including sleeping space and toilet and cooking facilities.”35

Obviously, any dwelling that is missing one or more of these features provides a definitional rationale for classifying a resident of that dwelling as homeless. This source is less instructive in terms of elucidating what is meant by permanence, but operational parameters could be identified. And, while the spirit of the IRS definition is consistent with common-sense notions of what constitutes a home,
one could easily imagine a shelter for the homeless, which may contain shared sleeping, toilet, and cooking facilities, as qualifying as a home within the letter of the definition. Clearly, reliance on this definition is inadequate for the purposes of providing a conceptual framework for enumerating the homeless.

One researcher who has articulated a conceptual definition of homelessness is Rossi: “Homelessness, at its core, means not having customary and regular access to a conventional dwelling; it mainly applies to those who do not rent or own a residence.”

This definition focuses on the “literally homeless” population. Rossi recognizes that other conditions put individuals at risk of homelessness (e.g., precarious housing situations such as being doubled up or being weakly attached to the housing market), but prefers to label these persons as “the extremely poor.” The literally homeless are considered to occupy the lowest level of this standard-of-living continuum.

The definition incorporated by the McKinney Act (P.L. 100-77, see 103(2)(1), 101 stat. 485 (1987)) adopts a somewhat broader view of homelessness:

The term “homeless” or “homeless individual” includes an individual who (1) lacks a fixed, regular, and adequate nighttime residence and (2) has a primary nighttime residence that is (a) a supervised, publicly or privately operated shelter designed to provide temporary living accommodations (including welfare hotels, congregate shelters, and transitional housing for the mentally ill), (b) an institution that provides a temporary residence for individuals intended to be institutionalized, or (c) a public or private place not designed for, or ordinarily used as, a regular sleeping accommodation for human beings.

In addition to the literally homeless, this definition also allows inclusion of individuals who are in such institutions as psychiatric hospitals (prisons are explicitly excluded in the act’s definition), if they have no specific dwelling to return to upon their release; this expanded coverage is similar to that incorporated by other federal statutes. The National Institute of Alcohol Abuse and Alcoholism (NIAAA), in its recent Cooperative Research Demonstration Project for Homeless Individuals with Alcohol and Other Drug Problems, further broadens the McKinney definition by stating that those “doubled up” should be eligible for services funded by its project. Although it is not clear whether Burt and Cohen or Sosin, Colson,
and Grossman specified a conceptual definition prior to conducting their research, their operational practices are consistent with this broadest conceptual framework.39

Operational definitions

What should be clear from the preceding section is that conceptual definitions can guide decisions to include or exclude an individual by limiting or expanding the target population to be counted.40 In making the translation between explicit or implicit conceptual definitions and the specific operational procedures to be employed, several additional definitional issues must be addressed. These include the relevant individual characteristics (e.g., age) and the settings in which appropriate individuals are likely to be located (e.g., in shelters, in institutions, or on the streets). Persons in those settings often may not be homeless; for example, Burt and Cohen and Sosin, Colson, and Grossman found that not all clients of soup kitchens and meal programs are homeless.41 Such findings support the need to implement specific rules for identifying who should be counted, including the development and use of screening questions that may ask whether the individual is homeless and the duration of homelessness. Counts also can vary depending upon the time frame covered by the enumeration process (incidence versus prevalence). Specifically, estimates can be derived for a particular point in time (e.g., a given night) or over an extended period of time (e.g., in a given year); for the latter, the unduplicated count provides an estimate of the total number of individuals experiencing transitory or extended periods of homelessness.

Individual characteristics. Given that the homeless population is diverse and becoming more so, the central question, beyond the total number of homeless individuals, addresses the composition of the population. For planning services, such concerns are important because not all populations are likely to require the same set of services. Individual characterization of the population can assume an endless number of configurations; for example, key subgroups include homeless substance abusers, homeless chronically mentally ill, homeless families, and unaccompanied youths. In its most recent report, the Interagency Council on the Homeless catalogues the important subgroups and summarizes the federal initiatives that have been undertaken.42

For the purposes of this paper, what is important is the recognition that segmenting the total homeless population into policy-relevant components is central and has methodological implications. In particular, an interest in describing the population requires
individual-level data; therefore, a reliance on settings as a means of
determining who is homeless is insufficient. Nevertheless, settings
provide a rough (i.e., probabilistic), \textit{a priori} categorization scheme
for framing the enumeration task.

\textit{Settings where homeless persons are likely to be found.} In an effort
to summarize prior studies aimed at providing estimates of the
homeless population, Burt and Taeuber surveyed the authors of
seven studies.\textsuperscript{43} To identify who had been counted by the
researchers, Burt and Taeuber compiled an extensive listing of pos-
sible settings and conditions where homeless individuals might be
found. These were classified into seven generic categories with
numerous specific subcategories:

1. \textit{Shelter-type institutions.} These include homeless shelters,
domestic violence shelters, subsidized temporary
hotels/motels/apartments, and runaway and homeless youth
centers.

2. \textit{Non-shelter institutions.} In this category are jails, mental health
facilities, detoxification centers, and quarter-way, half-way, and
three-quarters-way houses.

3. \textit{“Gray area” institutions.} These include single-room-occupancy
dwellings, hotels that are paid with the individual’s own
resources (e.g., YMCAs and residential hotels with long-term
occupants), and transitional/permanent housing projects (e.g.,
group homes and arrangements targeted at the “once
homeless”).

4. \textit{Long-standing institutions.} This group includes hospitals
housing boarder babies, facilities designed to shelter victims of
natural disasters, and housing facilities for other emergency sit-
uations (e.g., facilities for abused and neglected children).

5. \textit{Nonresidential institutions.} This category is different from the
previous settings in that it does not include sleeping arrange-
ments, but rather service settings specifically targeted at home-
less individuals: soup kitchens, mobile food vans, drop-in
centers, and health clinics.

6. \textit{Noninstitutional locations.} These are nonconventional dwellings,
including streets, parks, transportation depots, abandoned
buildings, parked cars, sections of highways (e.g., underpasses),
public transportation (e.g., subways), parking garages, and rail-
road boxcars.
7. **Conventional dwelling units.** The focus here is on identifying those individuals who are “doubled” up or otherwise precariously housed.

While the list is daunting, it is instructive in terms of the potential scope of the search process. It should be noted that Burt and Taeuber did not identify the conceptual definitions used in these seven studies or in developing this list; rather, these settings appear to be based on logical and plausible locations.

The number of sites that served as the basis for enumeration varied dramatically across the seven studies reviewed, and not surprisingly, no study included all of the settings listed earlier. The choice of settings that were investigated appears to have been influenced by different conceptual definitions (either tacit or explicit). For example, Rossi, Fisher, and Willis focused on homeless shelters (both short-term and transitional) and nonconventional dwelling locations (e.g., streets, parks, and abandoned buildings).\(^4^4\) Rossi has articulated the clearest linkage between the conceptual and operational definitions.\(^4^5\) Stemming from his notion of literal homelessness, “conventional dwellings” include homes, apartments, mobile homes, rented rooms in hotels, rooming houses, or private homes, but exclude such public places as subway stations, scrap-material shacks, abandoned buildings, and dormitory arrangements (as in shelters). In addition, the identification of appropriate settings is also defined by his notion of “customary and regular access”; this access is granted to those who rent or own a home or who are members of a household that occupies a dwelling (e.g., children). Individuals who fall outside these boundaries are considered literally homeless. However, Rossi’s definition excludes those residing in institutions (e.g., hospitals and jails), and those doubled up or precariously housed fall into the extremely poor, rather than literally homeless, category.

Adhering to the McKinney definition provides additional latitude, given that such terms as “fixed, regular, and adequate nighttime residence” are subject to interpretation in terms of who should be counted. In fact, this definition explicitly states that those in institutions can be classified as homeless. Our reading of the definition suggests that the McKinney Act incorporates a compound definition. That is, not only must an individual lack a fixed, regular, and adequate nighttime residence, but also the individual’s nighttime residence must fall into one of the three categories identified in the act. Therefore, the institutionalized should be counted only if they lack a fixed, regular, and adequate residence upon discharge.
Some efforts have understood the McKinney definition to encompass persons who double up. For example, NIAAA explicitly views those doubled up as at risk for homelessness and eligible for services provided by NIAAA research demonstration projects. Similarly, space in which persons are doubled up may be viewed as not “fixed, regular, and adequate” because of the potential of extremely overcrowded and unsafe conditions. In contrast, other efforts to enumerate homeless persons (e.g., those of some state departments of education) have more rigidly construed the definition to exclude those who are doubled up from their reporting to the U.S. Department of Education on homeless children and youths.46

**Screening questions.** The list of settings is introduced by a key phrase—“likely to be found.” This phrase implies that not all individuals observed in those settings are in fact homeless. Therefore, simply counting the number of occupants in a particular setting at a given time may overestimate the size of the homeless population. For example, some residents of prisons are unlikely ever to be released, and thus their housing status is immaterial. Others may have families and spouses (and homes) to return to, and others may have sufficient resources to procure housing upon their release. The methodological implications are straightforward: it is necessary to interview residents of certain locations. For other settings (e.g., homeless shelters), the use of screeners is less important; in fact, most researchers have merely assumed that an individual’s presence in a shelter is sufficient to classify him/her as homeless.

**Duration and prior history of homelessness.** It is well recognized that homelessness is dynamic, with some individuals cycling in and out of various settings that represent more or less extreme forms of deprivation. For example, of those using meal programs or in treatment programs, nearly two-thirds of Sosin, Colson, and Grossman’s Chicago sample were not currently homeless, although a large majority of these had previously been homeless.47 Santiago et al. published similar findings, and their expanded definition (from “currently homeless” to “homeless within the last three months”) increased the number of individuals classified as homeless by 50 percent.48 The determination of duration and prior history of homelessness hinges on the investigator’s initial definition of homelessness.

**Incidence versus prevalence.** As noted earlier, counts of homeless individuals vary depending upon whether they are counts for a particular night or point in time (point prevalence), counts of the number becoming homeless within a given year (annual incidence), or counts of the cumulative total (annual prevalence). In deriving
these estimates, it is necessary to take into account the duration and prior history of homelessness, which are generally self-reported and subject to all the limitations of this approach.

The intended uses of the estimates and their methodological implications

The complexities associated with defining, locating, and counting the number of homeless individuals are unlikely to be the result of serendipitous or whimsical decisions, regardless of whether the estimate is for local or national use. Furthermore, for each estimate, there are one or more specific requestor(s) of the information who have a plan for how this information will be used (e.g., Congress might want to determine financial allocations required for services to the target population). For example, Section 722 of the McKinney Act requires each state to gather data on the number and location of homeless children and youths for reporting to Congress via the Secretary of Education. This information is then to serve as a basis for ensuring that this group has access to free and appropriate education.

If the main reason for attempting to understand the size and composition of the population is to improve service delivery, the definitions (who gets counted and where) and methods (how the estimate is derived) are likely to differ from those used if the goal is simply to derive a figure for ascertaining the magnitude of the problem. If the primary reason for conducting a count of the homeless is to develop a short-term plan for delivering services (e.g., to ensure adequate numbers of beds in local shelters), the focus will be on counting those individuals who are potential shelter occupants. Thus, the effort is directed toward counting the literally homeless, rather than those marginally housed or on the streets and not likely to use shelters. In contrast, the design of a long-term delivery plan requires that the definition of “homeless” include not only shelter users but also those likely to become homeless in the near future (those at imminent risk). To accommodate these various interests and uses, it is desirable to have a conceptual definition that is inclusive.

But as the definition of homelessness is expanded, methods for gathering data become more complex. Rossi makes this point quite clearly and argues that a definition must “cover the essence of that term and [be] also practical to use in actual research” (italics added). There are good technical reasons for constraining the definition in such a way that it can be studied, but political arguments exist for not restricting definitions in accordance with methodological convenience.
One important practical consideration is the cost of obtaining statistically sound estimates. Although it is possible to construct an elegant study for each component embodied in the broadest definition imaginable, the breadth of settings outlined above would make such an undertaking formidable and tremendously expensive. Even with adequate time and money, the rapid transition in the composition of this population would almost certainly call precise estimates into question. Although longitudinal studies (panel and cross-sectional) could be undertaken, political changes in the acceptability of definitions suggest additional sources of slippage. So, although money and time are important, and good studies take both, they are not sufficient to solve all problems.

If we lower our sights somewhat, we can reduce the time and resources needed to conduct a study in a variety of ways. For example, a precise national estimate could be obtained if we imposed a limited definition of homelessness (e.g., only those individuals located in shelters or on the streets). As the Census Bureau knows so well, concentrating on specific segments of the population invites accusations of bias.

To counter these claims requires a broadening of the definition of homelessness. However, this does not resolve the problem. Embracing a broad definition of homelessness—one similar to that specified in the McKinney Act or NIAAA’s expansion—is desirable, but not practical for all purposes. It appears unlikely that every estimate would or could include members of all possible groups that meet a broad definition’s criteria for inclusion. Therefore, unless research practices are altered, estimates will continue to be based on selected subgroups of the homeless population. Confusion will prevail.

Intelligent segmentation

The definition of homelessness—regardless of its breadth—is not really as intractable a problem as it may appear. The real problem is in how the resulting numbers are reported. Two issues are important—the completeness of reporting and the level at which components or segments of the homeless population are reported. First, “counters” can be clearer about who is being counted and who is not. It is often quite difficult to ascertain with confidence exactly who was included in the enumeration (or estimation) process. In some cases, it is impossible to tell. Second, even when operational definitions (e.g., settings and conditions) are described, the numbers found in each segment are not disclosed separately. The authors use
the term *intelligent segmentation* to denote how reporting of estimates could be carried out. In particular, a total estimate of the homeless would be broken into individual estimates for each of the subgroups examined. Rossi provides a nearly perfect example of this strategy in reporting his results on the shelter and street components of his Chicago study. In a later section of this paper, an illustration incorporates a broader definition than Rossi’s literal homelessness concept.

**The need for a family of studies**

Counting the number of homeless in the various settings encompassed by a broad definition will require the use of multiple methodological tactics. We refer to this as a *family of studies*, with each member targeted at a specific component of the conceptual definition. Because of the uncertainty associated with each method and the difficulty of implementing these strategies with transient populations, the necessary family of studies must also include “side studies” that check on the quality of each “member” study. This tactic is readily seen in the work of Taeuber and Siegel, Dennis, and James, to name a few. The constellation of member studies will, in all likelihood, pattern the complexity of the definition that is used. The specific locations where the homeless are likely to be found entail sampling designs that are tailored to each segment. In other words, multiple studies are needed. Unlike conventional household surveys, probability-based estimates of individuals in settings involving nonconventional dwellings requires an alteration in the traditional way we think about sampling frames. Dennis provides an accounting of these modifications. For example, blocks need to be sampled. Because not all blocks in a given geographical area will have the same probability of containing homeless individuals, it is necessary, for the purposes of sampling efficiency, to rely on informants to classify blocks; in this manner, differential probabilities of finding homeless individuals can be obtained so that stratifications can be built into the sampling design. The sampling frame developed for shelters is likely to be more similar, in practice, to conventional survey tactics. At the same time, constructing the list is likely to depend on multiple sources (e.g., key informants, listing of community service agencies, and telephone contacts).

Establishing the technical credibility of each member study is the primary intent of the second type of study within our so-called family. This focus is not novel, of course, inasmuch as almost every competent survey researcher is aware that transgressions can occur in the execution of the various study phases. Sampling frames are
not always complete, interviewers do not always follow instructions, respondents may not cooperate, and some respondents are incapable of providing accurate information on their prior sleeping arrangements or shelter status. The list of sampling and nonsampling errors and biases that might be encountered is substantial. Estimating the magnitude of these concerns is the primary objective underlying this type of methodological study.

Intelligent segmentation and multiple methods: an illustration

To make the idea behind intelligent segmentation clear and to show how the family of studies might be configured, the next section provides an illustration. This two-pronged approach was used in a recently completed study issued by the U.S. General Accounting Office (GAO). Although some of the issues were idiosyncratic to the particular question of interest, others were generic and have immediate application to other issues in the field (e.g., estimating prevalence of subgroups). Similarly, the analytic tactics (e.g., sensitivity analysis) are sufficiently general to be useful in other areas. It should be noted that we do not believe that this illustration necessarily represents the best way to estimate the number of homeless children and youths. However, we are convinced that it provides a useful illustration of how future results should be reported.

Origins of the GAO study

In reauthorizing the McKinney Act (P.L. 100-77), Congress requested that GAO provide an estimate of the number of homeless children and youths in all 50 states. That same law also required the secretary of education to compile and submit to Congress, through state education agencies, data on the number and location of homeless children and youth. The Department of Education issued its report on February 15, 1989, and in essence, the mandated count assigned to GAO was intended to function as a check on the accuracy of the department’s efforts. On June 15, 1989, GAO issued its report. It should be noted that the congressional language basically requested complete, national enumeration within 12 months. Through a series of negotiations, an estimate based on a representative sample was deemed sufficient to meet congressional needs. Other trade-offs will be made clearer as the example unfolds.

In a nutshell, different methods were used to derive an overall estimate. Because not all methods were equally trustworthy (actual
enumerations, expert opinion, and population extrapolations contain different strengths and weaknesses), the accuracy of the results produced by each method needed to be documented and disclosed. This differential trustworthiness resulting from statistical and nonstatistical sources of error and bias can be incorporated into confidence intervals or ranges based on sensitivity analyses.

The bottom line

Using a variety of methods, GAO estimated that on a given night (October 24, 1988), about 68,000 children and youths aged 16 years or younger were members of families that are literally homeless (see table 1). As argued earlier, a broad interpretation of the McKinney Act would suggest that the estimated 68,000 homeless children and youths represent only part of the potential homeless population, namely, that part that is literally homeless. Many advocates and stakeholders view this as a fairly narrow conceptualization of the problem. As a means of broadening the definition, GAO attempted to estimate those believed to be precariously housed (e.g., those doubled up with relatives or friends). This group is quite large, representing an additional 186,000 children and youths who could be considered homeless on any given day. GAO stated that these estimates did not include homeless runaway children and youth, nor did they account for those families who may be on the brink of homelessness by virtue of their economic situation. Despite these omissions, adding the number of precariously housed to the number of individuals who are literally homeless reveals that, on any given night, more than 250,000 children and youth might be considered homeless.

An illustration of intelligent segmentation

Simply reporting an aggregate figure (say 250,000) would have made it impossible to know how much each component of the definition contributed to the overall total. Segmentation of this estimate into its component parts makes it clear that the numbers differ across settings. As might be expected, homeless children and youths were not evenly distributed across the different locations where homeless families are thought to congregate or “reside.” Nationwide, urban shelters and hotels housed families with roughly 25,500 children and youths; about 21,800 were likely to be in suburban and rural areas; churches accounted for about 4,000; abandoned buildings, cars, or public places were likely to be called home by
Table 1. Estimated Number of U.S. Homeless Children and Youths at Any Given Time

<table>
<thead>
<tr>
<th>Category</th>
<th>Best Estimate</th>
<th>Rangea</th>
<th>Sourceb</th>
<th>Confidencec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Literally homeless</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shelters, hotels</td>
<td>25,522</td>
<td>18,265</td>
<td>32,779</td>
<td>Surveys</td>
</tr>
<tr>
<td>Churches</td>
<td>4,094</td>
<td>2,340</td>
<td>6,570</td>
<td>Opinion</td>
</tr>
<tr>
<td>Public places</td>
<td>9,016</td>
<td>4,512</td>
<td>24,072</td>
<td>Opinion</td>
</tr>
<tr>
<td>Other</td>
<td>7,651</td>
<td>5,168</td>
<td>10,446</td>
<td>Opinion</td>
</tr>
<tr>
<td>Suburban</td>
<td>14,427</td>
<td>7,213</td>
<td>21,641</td>
<td>Population Rates</td>
</tr>
<tr>
<td>Rural</td>
<td>7,357</td>
<td>3,678</td>
<td>11,035</td>
<td>Population Rates</td>
</tr>
<tr>
<td>Total</td>
<td>68,067</td>
<td>41,176</td>
<td>106,543</td>
<td>All</td>
</tr>
<tr>
<td>Precariously housed</td>
<td>185,512</td>
<td>39,362</td>
<td>296,452</td>
<td>Opinion</td>
</tr>
</tbody>
</table>


a The low and high estimates represent a plausible range of values based on various assumptions.

b The three primary sources of information upon which estimates were based were surveys of shelter providers and agencies providing vouchers (conducted on October 24, 1988), the application of homeless rates to a population base, and expert opinion.

c The confidence rating reflects an assessment of the level of certainty that can be expected of the estimates, given the reliability of the data sources and the range of estimates provided.

About 9,000; and about 7,000 may have been in various other settings (e.g., institutions). Not only is the utility of survey results greatly enhanced if data are presented setting by setting, as shown in table 1, but such practices sidestep the issue of which definition is correct. Furthermore, such reporting implies (and indeed, encourages) that the broadest possible definition (within resource constraints) should be used in future surveys. When idiosyncratic
definitions and operationalizations of homelessness are used across different sites, presenting separable estimates (by subgroup) is essential if the results of such estimation exercises are to be used intelligently.

An illustration of the family of studies

Opting for a broad definition of homelessness and one that fairly represented the settings where the literally homeless were likely to be found had substantial consequences for the estimation procedures that could be employed. The purist in all of us would probably like to use a common set of procedures (across settings) that could be defended on statistical grounds. In attempting to meet the request of Congress, GAO had to rely on a unique mixture of methodological strategies in piecing together its overall estimates. Basically, the strategy involved three steps and multiple methods (a count, expert opinions, and population-based extrapolations).

The first phase was very traditional. It entailed the use of survey methods whereby an unduplicated estimate of the number of children and youths in shelters and hotels (or motels) in 40 large urban areas (populations in excess of 250,000) was obtained, representing 27 cities. A multistage probability sample was drawn to select shelters; a telephone survey was used to obtain counts of the number of children housed in shelters on a particular night (October 24, 1988); and the number of vouchers for hotel or motel accommodations issued by the county were obtained for that same night. This method was intended to provide a nationally representative estimate of the number of children and youths in shelters and hotels or motels in urban counties. It also served as the foundation for the other estimate procedures.

To use a sample-based methodology for obtaining estimates in other settings (e.g., streets) would have been prohibitively expensive and time-consuming. GAO opted for an approximation. In particular, the second phase of the study involved developing separate estimates of the number of children and youths in some of the other settings used to define literal homelessness. These were derived by using the survey-based county estimates in conjunction with expert opinion. The experts’ ratings reflect their estimates of the proportions of homeless children and youths in each of these other settings. The same procedures were used to obtain estimates of the number of children and youths who might be precariously housed (e.g., doubled up).
Using sampling methods to simply identify children and youths in urban and suburban areas would have been too time-consuming and expensive. Again, GAO used an approximation based on a different methodology. For estimates of children and youths in rural and suburban areas, actual survey-based counts were used in conjunction with other estimates (derived from the empirical literature) and population counts in nonurban areas. Here, estimates based on the median homelessness rate across the 27 cities were used to project the homelessness rate relative to the population base in rural and suburban areas. Rather than relying on expert opinion, these estimates used extant data on population size in nonurban areas and estimates from the study’s first phase. The third phase assessed the accuracy of the estimates produced by these procedures by comparing the results with those reported in other national studies.

**Side studies on methodological adequacy**

Although estimates were produced for each setting within the definition of homelessness, they were obviously not of equal integrity. To account for the precision of each estimate, the GAO study also calculated ranges similar to confidence intervals (see table 1). With the exception of the survey results, these were not ordinary confidence intervals in the sense of classical statistics. Rather, sensitivity analyses were used as a basis for determining the robustness of each estimate. This produced an upper and lower boundary for each setting × method combination. Whereas an ordinary confidence interval uses the error of estimate as the basis for specifying the interval within which the true population value will be found, the interval estimation procedures altered key parameters or assumptions underlying each calculation. In this way, GAO was able to provide the user with a sense of the stability and sensitivity (to alternative assumptions) of each value. GAO also supplied a verbal description of the overall confidence placed in each estimate; ratings ranged from “low” to “high.”

Little is known about the prevalence of homelessness in suburban and rural areas. Because the initial sampling frame was restricted to large urban areas, it would have underestimated the number of homeless children and youths by omitting suburban and rural areas. Prior studies of the homeless population in suburban areas assumed that prevalence is about one-third the rate of that for central cities. In creating the lower boundary for the confidence interval or range, one-third of the median rate of homelessness found in the sample of cities was used. In constructing the upper
boundary, the assumption was made that the median rate was appropriate. Here, the best-guess point estimate ended up being the average of the high and low boundary values. Although these values are derived from extrapolations of estimates (GAO’s own estimates from the survey results) and compounded estimates (an estimate from another study was applied to its survey-based estimate), the confidence that might be placed in these values was judged as moderate. That is, it was judged to be below the confidence level placed in the survey results and above the confidence level ascribed to the opinion-based estimates for the number of individuals housed in churches, in public places, in other settings (such as institutions), and those in doubled-up situations (see table 1).

Specifically, the opinion-based estimates were derived from interviews with shelter providers, advocates, and knowledgeable government officials in the sample of 40 counties (covering 27 cities). More than 300 individuals provided their countywide estimates of the relative number of homeless families residing in settings other than public or privately sponsored shelters and hotels. These responses were converted to ratios that, when applied to the estimated number of families in shelters, provided estimates of the number of families in other settings. The median ratios for each county were computed along with lower and upper bound values (first and third quartile, respectively). The results showed that opinion-based estimates were very sensitive to the choice of values (median, first, or third quartile), depicting a substantial range in values. As seen in table 1, using the median ratios derived within counties produced an estimate of approximately 186,000 doubled-up children and youths nationwide. The range, however, suggests that there may be from 39,000 to 296,000, depending upon how the data are analyzed at the individual county level. The ranges for other estimates using this method are also quite broad. This is especially true for expert-based assessments of the number of children and youths that are likely to be in public places. Because of this hypersensitivity, the GAO report judged the confidence that should be placed in opinion-based estimates as low.

Sensitivity analyses are helpful to a certain extent, but they cannot establish with certainty the overall sensibleness of a set of calculations. The decisions that were made on high versus low estimates—although based on logic and, where possible, prior data—are but a subset of all the possible values that could have been chosen. And, although it is better to provide the client or user with a sense of the confidence that should be placed in the numbers that have been produced, bracketing in this way does have inherent limitations.
As a means of judging the adequacy of these methods, a second form of multimethod research can be used—namely, comparing competing estimates from parallel studies. Two recent reports served as the basis for other national estimates of the number of homeless children and youths. The Institute of Medicine (IOM), using data from the National Alliance to End Homelessness, estimated that 100,000 children and youth were literally homeless, in contrast to the GAO estimate of roughly 68,000. A careful review of the IOM methodology suggests several noncomparabilities across procedures. Adjusting the IOM estimate to reflect new information on service utilization—a key assumption in the IOM estimate—reveals an adjusted estimate of 87,000; although this figure is closer to GAO’s best-guess value of 68,000, it is about 23 percent higher than GAO's. However, it is well within the confidence range established for the estimate of the number of literally homeless.

In another report, the Urban Institute, under contract to the U.S. Department of Agriculture, conducted a study of homeless in shelters and soup kitchens in 20 cities with populations greater than 100,000, and extrapolated its findings to the nation. The Urban Institute reported approximately 61,500 homeless children in cities and suburban areas. GAO’s comparable estimate was 60,710, excluding rural settings. Although none of these three studies is, by itself, able to stand up to close technical scrutiny, the fact that they differ in approach and converge within a reasonably close confidence range suggests that GAO probably derived a sensible understanding of the magnitude of the problem.

Conclusions

In this paper, the authors have shown that several definitional issues must be addressed to achieve an understanding of what’s behind the numbers. The thorniest of these issues concerns conceptual definitions—that is, what is meant by “homeless.” Some definitions are quite literal, whereas others are more expansive, including individuals who technically have shelter but are tenuously housed. Which definition is correct is not an easy question to answer. Should we worry about such definitions? Of course. Are definitional issues worth agonizing over? Probably not. Rather, we see several sensible rationales for incorporating a broad definition at the onset; there are a variety of legitimate reasons for being concerned about the size of various subgroups that compose the segment of the population labeled the “extremely poor.” In particular, policy formulation and services planning require a broad look at who is homeless and who may be homeless in the future.
However, broad conceptual definitions expand the number and types of locations within which homeless persons are likely to be found, and this diversity constitutes a sizable methodological challenge. This situation should not restrict inquiry as to the size and composition of the population. We would prefer to see more concerted efforts to clearly delineate which segment(s) of the population compose any aggregate numbers. The concept of intelligent segmentation as a vehicle for reporting estimates of the components of a broad definition is a means of dealing with the obstacles inherent in developing a consensus as to who should be counted in the aggregate. Having a broad definition that can be decomposed into its separate segments allows for a better understanding of who has and who has not been included in any particular study. This facilitates the understanding of differences across studies and the use of the numbers by other interested policy makers, researchers, and the public.

We also suggest that the complexity inherent in broad definitions will require the development of a family of studies to capture the full range of conditions facing the homeless and those who are at risk. The methods underlying each study within the family are likely to be uncertain due to a host of sampling and nonsampling biases that we may know little about for particular subpopulations. The values that are reported should reflect these conceptual and methodological uncertainties. In this way, consumers will be better informed about the likely “true” range of homeless individuals in the United States.

Authors

David S. Cordray is professor of public policy and psychology and chair of the Department of Human Resources at Vanderbilt University. He also is director of the Center for the Study of At-Risk Populations and Public Assistance Policy at the Vanderbilt Institute for Public Policy. Georgine M. Pion is research associate professor of psychology and human development at Vanderbilt University and fellow in the Center for the Study of At-Risk Populations and Public Assistance Policy. Both are involved in a national, NIAAA-funded evaluation of 14 research demonstration projects to examine the effectiveness of interventions for homeless individuals with alcohol and other drug abuse problems.

The authors would like to thank Peter Rossi, Howard Sandler, and Mark Appelbaum for their helpful comments on this paper and Kelly Johnson, Ann Dismukes, and Nell Thelin for their assistance in its preparation and production. The U.S. General Accounting Office study discussed here as an illustration of the use of intelligent segmentation and multiple methods was conducted while David Cordray served as assistant director within the Program Evaluation and
Methodology Division of the U.S. General Accounting Office. The work summarized here was a collaborative effort involving Lois-Ellin Datta, James Onken, and Peggy Murray, among others.

**Endnotes**


5. Rossi, *Down and Out in America*, 54.


15. Ibid.

16. Rossi, Down and Out in America, 81. Although no specific dates were provided by Rossi, we assume that he refers to the year just prior to the publication of his book.

17. Hombs and Snyder, Homelessness in America.


24. Kim Hopper, “Homelessness Old and New: The Matter of Definitions,” this journal. Hopper provides a nice account of the historical changes in definitions of homelessness. He takes a social constructionist perspective on problem definition, also pointing out cultural differences in defining terms such as conventional dwelling.


34. As noted by Rossi in his comments on this paper, the Oxford English Dictionary’s definition of home also is saturated with references to family, intimacy, acceptance, and other feelings of social support. In contrast, current definitions of homelessness contain no such allusions. This is a shift from definitions held by social scientists prior to 1980, in which homelessness often was equated with familylessness and spouselessness.


40. A small-scale study that also supports the ways in which definitional criteria affect the number of individuals viewed as homeless was conducted by James Morrison, “Correlations Between Definitions of the Homeless Mentally Ill Population,” *Hospital and Community Psychiatry* 40 (1989): 952-54.


44. Rossi, Fisher, and Willis, “The Condition of the Homeless.”


50. Hombs and Snyder, *Homelessness in America*.

51. Rossi, *Down and Out in America*.


53. Dennis, “Changing the Conventional Rules.”

54. Taeuber and Siegel, “Counting the Nation’s Homeless.”


58. This estimate is based on information in Burt and Cohen’s *Feeding the Homeless*. Additional information is contained in a memorandum from the Urban Institute to the U.S. Department of Agriculture, and information obtained in a telephone conversation with an Urban Institute researcher.