Session 4:

Next Steps in Homelessness Knowledge Acquisition and Social Policy

Strategies for Homeless Research in the 1990s
Peter H. Rossi

Federal Homeless Social Policies for the 1990s
Cushing N. Dolbeare

Strategies for Homeless Research in the 1990s

Peter H. Rossi
University of Massachusetts —Amherst

Abstract

Existing empirically based knowledge about homelessness is neither extensive nor well grounded. Proper policy formation and program design that address homelessness are badly served by knowledge deficiencies. This paper lays out a strategy that will develop appropriate research methodology and generate better empirically based knowledge. It proposes a program of technical research, as well as moditications to several national data collection programs, such as the Current Population Survey, to collect data on the literal homeless and the precariously housed. It also presents a rationale for starting longitudinal studies that focus on the processes that lead into and out of the homeless state. Finally, it proposes evaluation research to measure the effectiveness and efficiency of promising programs.

Introduction

Public policy information needs and the development of descriptive statistics are intimately related both in history and in current times. The development of each new problem in public policy is accompanied sooner or later by the development of one or more data collection efforts designed to provide firm data on the problem's size, location, and trends. Homelessness is no exception. It has been identified as a major social problem of national scope. As this and other conferences testify, there is some support for adding data on homelessness to the existing list of national problems on which the national data system provides timely information.

In an ideal world, it might be possible to advocate that policy formation take place only when the best possible information on the specific problem is on hand. In the real world, however, the identification of social problems, the formation of social policies, and the design of social programs often proceed at a pace that quickly outstrips the extent and quality of empirically based knowledge. In most cases, the country would not have it otherwise. In the 1980s, when American society identified homelessness as a serious social problem, it would have been heartless to postpone providing help and services until sufficient knowledge about homelessness had been accumulated. Accordingly, policy makers and human services providers proceeded to put in place a somewhat haphazard structure of policies and programs addressing the issue.

Much was thereby accomplished to avert more serious distress: By the end of the 1980s, several hundred thousand beds were available in shelters, and perhaps as many as 500,000 meals were served daily in food kitchens. Further, medical clinics have been set up in major cities to meet some of the medical needs. Some portion of the obvious and most serious “emergency” needs of the homeless are being met. But it is equally clear that no one knows the size or the pattern of the total need presented by homelessness, nor is it known whether the programs and services being delivered are adequate, effective, and efficient.

As articles in this journal testify, there is no firm understanding of the nature of the homeless problem, the extent and distribution of homelessness, and trends in the growth of homelessness. This knowledge deficiency inhibits the ability to plan effective programs, to allocate adequate resources efficiently, and to find long-term solutions. Surely the time has come to create the knowledge base that can improve policy measures and strengthen programmatic efforts.

The policy-driven needs for information about homelessness center around five main issues:

1. Magnitude: How big is the problem of homelessness? How many persons are homeless?

2. Distribution: How is the problem distributed in physical and social space?

3. Trends: What are trends in homelessness? Is the problem increasing in size? Is the distribution changing?

4. Policy formation: What are appropriate changes in policy that would be responsive to the problem?

5. Effectiveness: How effective are policies that are being enacted? Are such policies producing intended effects at a level of effectiveness that justifies the interventions involved?

Although it cannot be said that public policy in this country (or in any other, for that matter) is driven by data, neither can it be said that existing credible data are ignored entirely in making policy. For example, the proper way to target a public policy requires knowing something about the target and its location. It is also necessary to know enough about the macro- and micro-level processes involved to design appropriate and effective programs. Whether homeless policy will be strongly influenced by data remains to be seen. Currently, the absence of credible descriptive data surely must be an obstacle to pursuing any policy direction. Without firm knowledge on the magnitude of the problem, it is difficult to decide what expenditures may be needed. Without being certain whether the problem is mainly an urban one or is found in all environments, a delivery system is hard to design. Without knowledge of trends, it is difficult to know whether efforts are reducing the problem overall.

This paper outlines the empirical studies needed to provide credible data and the plausible research strategies necessary to collect the needed data. As many of the articles in this journal show, homelessness is not an easy topic for empirical research. Conventional approaches that have worked for such problems as unemployment and access to medical care cannot be used without drastic modification. New approaches are still in the development stage. Other promising strategies are suggested in this paper that have never been tried.

Conceptual obstacles

However convenient it may be to regard social phenomena as simple, it is usually wrong to do so. A casual observer might be tempted to regard homelessness as a binary condition (i.e., an individual is either homeless or not), but this is not a productive view. The controversies over what should be regarded as homelessness at either the conceptual or the operational level are testimony to the fact that the boundaries between the housed and the homeless are much in dispute. It is unfortunate, though not atypical, that definitional differences coincide with political cleavages blocking the construction of definitions that can get widespread support.

There are no scientific ways to resolve definitional disputes. It appears likely, however, that cost considerations will narrow the field somewhat. Those definitions that are most inclusive may be too expensive to implement, leaving the field to those definitions with more reasonable cost implications. For example, those definitions of homelessness that essentially cover all instances in which persons are housed inadequately—that is, households that are doubled up or overcrowded and persons living in dilapidated housing, as well as persons living in shelters or in public places—imply data collection operations that are probably beyond available resources. The paper by Cordray and Pion in this journal provides a thorough review of the conceptual and operational issues involved in defining homelessness.

Although empirical research cannot solve definitional disputes, it can play an important role. The definitions that matter are those that guide actual research operations: namely, the rules researchers use to identify homeless persons. Nominal definitions, such as that used in the McKinney Act, can only be used in research if they are translated into specific research operations. For example, operational definitions that rely on observer judgments (e.g., interviewer judgments about the homelessness of individuals from their dress, behavior, and other aspects of their appearance) may produce estimates of homelessness that differ from those definitions that rely on screening interviews. Enumerations made during the dead of the night (12 midnight to 6 a.m.) may produce estimates that differ from those at other times.

Accordingly, technical research should be directed at testing the “sensitivity” of definitional variation. Some definitions that appear to be far apart when stated nominally may actually produce quite equivalent estimates, or at least estimates that are close enough to be considered equivalent for many purposes.

Technical research testing the sensitivities of estimates based on different definitions is under way by Michael Dennis and his colleagues at the Research Triangle Institute. This research is based on their National Institute on Drug Abuse–funded research on drug usage among homeless persons in the Washington, DC, metropolitan area. From this endeavor, a great deal more will be known about how alternative definitions of homelessness affect both the costs of research and the findings concerning magnitudes and composition. Existing research shows that shelters and food kitchens vary widely in size, admissions practices, and other factors.[1] Additional research is needed to assess how local variations affect size and composition estimates.

A strong, practical need exists for technical research like that under way at the Research Triangle Institute. Such research shows the gains and losses in precision and validity from the use of various approaches. Because costs vary from approach to approach, efficient strategies can be devised that provide data of the level of validity and precision needed for various policy concerns.

Technical research need not be conducted on a grand scale. The point is to explore the sensitivity of estimates to definitional variations, an aim that can be adequately fulfilled by local studies, provided that enough such studies are conducted to examine inter-locality variation.

Developing non-dwelling unit (NDU) sampling and census strategies

The main reason for the lack of good information about the homeless is that conventional research approaches are seriously deficient when it comes to the literal homeless,[2] defined as those persons who do not have customary daily access to a conventional dwelling unit. Conventional censuses and sample surveys are based on the assumption that almost all persons in society have customary access to a conventional dwelling unit. This is simply another way of saying that censuses and sample surveys are actually based on samples or complete enumerations of structures and dwelling units; thus, these counts largely miss those people who cannot be located within such residences. Although "group quarters," such as homeless shelters, are included among the structures visited by census takers and survey interviewers, they have not been identified as such in the resulting data, so it is virtually impossible to identify how many persons are enumerated within them.[3] And, of course, some portion of the literal homeless—exact proportions are unknown—may not reside in shelters at any point in time but may instead be sleeping in public places or in other places not normally used for dwelling purposes. These people, therefore, cannot be included in any sampling frames that are based on enumerations within dwelling units.

Conventional census and sample survey approaches allocate persons essentially to their nighttime locations (i.e., where they sleep). This is a reasonable strategy for most people, the overwhelming proportion of whom have a unique sleeping location. The study of homeless populations, however, may benefit from a strategy that uses locations such as food kitchens, medical clinics, or other places where the literal homeless may be found in significant numbers during the day.

The steps needed to correct conventional survey and census practices appear obvious. Conventional data collection efforts based on dwelling units need to be supplemented by samples and censuses of non-dwelling unit places, defined as locations that are not designed as dwellings; these include shelters,[4] food kitchens, accessible public places (e.g., bus stations, parks), places abandoned as dwellings, cars, and trucks. Michael Dennis's article in this journal reviews the strategies several researchers have taken to devise procedures for sampling NDU places.

But progress in devising efficient NDU survey procedures has not been entirely successful. Although it is clear that surveys based on samples of NDU places can be conducted, the procedures devised thus far are not efficient and hence are quite expensive. Furthermore, the approaches appear to be suited primarily for densely populated urban areas and have not been adapted to less densely settled areas.

Efficiency can be significantly increased by a strategy that concentrates the enumeration effort in places where the homeless are most likely to be located. Prior stratification—advance sorting by the size of the homeless populations to be found—can significantly lower costs and raise the precision of resulting estimates.

To rank which NDU places have high concentrations of homeless persons, firm knowledge is needed about where homeless persons are to be found. Typically, experts are queried to identify places of high concentration.[5] Stratification based on such information has not been optimal, however. If NDU samples are to be used extensively, either in stand-alone surveys or as supplements to an established data collection activity such as the decennial census, sources with better knowledge need to be developed. It may be too much to hope for, but strategies that generally apply to a wide variety of local conditions are highly desirable.

A further difficulty is that the NDU places where the homeless can be found are not unique locations. For example, a typical homeless person can be found in several NDU places because he or she uses food kitchens and shelters and also spends nights in public places such as mass transportation vehicles and stations. The probability of a homeless person being found in a given location is proportionate to the frequency with which that person uses that space. Accordingly, censuses and surveys have to be corrected for the overlap in users of various locations. This overlap is not an insurmountable problem if data are collected from identified homeless persons about how often they use food kitchens, shelters, and other NDU locations.

The need for further development of NDU research methods easily justifies the support of technical research oriented around the design and testing of ways to draw and implement NDU surveys efficiently. Although much has been learned from the experiences of the several studies Dennis reviews, “craftlore accretion” is a much slower learning process than is a series of well-prepared technical trials. For example, several alternative approaches to prior stratification can be tested comparatively. The research program might test the utility of basing prior stratification on using advance “windshield surveys”[6] versus using experts to spot locations where homeless persons congregate. Another strategy might be to test the effects of undertaking street sweeps between 8 p.m. and 12 midnight, 4 a.m. and 8 a.m., or 12 midnight and 4 a.m., to determine which approaches minimize screening burdens and maximize the probabilities of encountering homeless persons.

It is especially critical to understand the relationships between surveys that depend heavily on observations taken in NDU locations with known concentrations of homeless persons (e.g., shelters and food kitchens) and surveys that are based on all NDU locations. Additional variations include identifying all high-density NDU locations and confining observations to such places.[7] Although each of these modifications by definition misses some of the homeless population, their coverage of the total homeless may be good enough for many purposes, and their relatively low cost is an especially attractive feature.

It is hoped that such a technical research program will produce a set of strategies leading to valid estimates of the size and composition of the homeless population through affordable, tested procedures. It is anticipated that the variations will indicate that some strategies are better than others in covering all the homeless. The point of the technical research is to identify strategies that are affordable and that also yield estimates good enough for given purposes.

Some effort should also go into testing the applicability of the kinds of models exemplified in Anderton’s and Cowan’s articles, both of which are in this journal. These models may compensate for incomplete data collection resulting from the use of partial sampling strategies.

Modifying existing data collection efforts

Out of the need for policy-relevant information, the federal government supports a variety of data collection efforts. On the level of households and individuals, there are several well-known periodic surveys. Poverty and unemployment are tracked through the monthly Current Population Survey (CPS), crime victimization through the National Crime Survey, the size and composition of the housing stock through the American Housing Survey (AHS), the prevalence of drug use through the National Household Survey on Drug Abuse, and the extent of participation in government social programs through the Survey of Income and Program Participation (SIPP). Other surveys collect data on firms, schools, prisons and jails, and local and state governmental units. Typically, the household surveys provide periodic data that are useful for plotting trends and detecting shifts in the social and areal location of social problems.

These surveys do not now have the capability to enumerate the homeless directly. Nevertheless, they can be modified to add considerably to data concerning the homeless. In describing proposed modifications, this paper focuses primarily on the CPS to simplify the discussion. Most modifications can also be applied to other household- and individual-based surveys.

The CPS is based on a sample of about 2,000 counties selected, with probabilities proportionate to size, from the more than 3,000 counties in the 50 states. Within each sampled county, subsamples of local areas are drawn, and the households and individuals living within dwelling units are drawn from samples of successively smaller areas. The U.S. Bureau of the Census, which conducts the survey, maintains interviewers on locations close enough to the small areas sampled to facilitate personal interviews with each of the 57,000 households selected for monthly interviews. Although the survey's main purpose is to provide monthly estimates of the employment status distribution of people aged 15 and older, supplements provide additional demographic information and, on occasion, data on special topics such as child care, participation in adult education, and fertility and birth expectations. The demographic supplement administered to households every March is well known as the source of much data on poverty and its distribution in society.

A well-established national survey operation of this size is clearly an important national resource. To obtain the policy-relevant information needed for a better understanding of homelessness, careful consideration should be given to adding these components:

1. Adding a sample of NDU places

   Assuming that the technical research previously described leads to an efficient NDU strategy, it is believed that adding NDU places to the CPS would be a reasonable way to proceed. The suggestion is not that NDU samples be drawn from each of the 2,000 counties in the sample, but that smaller subsamples of counties be drawn. Of course, monthly estimates of the homeless are not now needed. An annual or biennial supplemental effort would likely suffice. As envisaged, the homeless supplement probably would be designed like this: A subsample of counties would be drawn from among the counties already in the CPS. Within each selected subsample county, small areas (tracts, blocks) would be selected after prior stratification, using the efficient strategies developed under the technical program described in the last section. Within each of the small areas, a search of NDU places would be undertaken to enumerate the unsheltered homeless. NDU places sampled can be a mixture of daytime congregating places, such as food kitchens, and of nighttime places of concentration, such as parks and bus stations. The important details—how many counties need to be selected and how many small areas should have NDU searches—are needed to yield reasonably precise estimates of the unsheltered homeless. These details should, of course, be left to the sampling statisticians to resolve.

2. Oversampling shelters

As shown in several articles in this journal, enumerations of the homeless housed in shelters can be conducted relatively easily. Because shelters are already included in the coverage of the CPS but are not clearly identified as such, a small change in procedures would simply be to identify more clearly when shelters are encountered. It is not possible to tell whether the group quarters designated in the current CPS are fraternity houses, convents, shelters for the homeless, or other group residences.

The number of shelters in any CPS appears to be very small. Thus, serious consideration should be given to assembling rosters of shelters in all the counties sampled and collecting from each shelter (or a sample) counts of the number of accommodations available and of the number of persons actually sheltered, either on a given night or over some short period of time. In addition, the shelter rosters can be used to support oversampling of shelter residents for interviewing purposes.

3. Adding food kitchens as sample units

Some consideration also should be given to adding food kitchens as sample units to the CPS. As many of the recent studies of the homeless have found, food kitchens are widely used by the literal homeless, and client surveys can be an inexpensive method of obtaining data on homeless persons. As noted earlier, a combination of shelter residents and food kitchen clients may cover enough of the total literal homeless population to serve many, if not most, purposes.

4. Adding variables to the CPS

In contrast to the first three proposals, those suggested in this section would entail relatively minor modifications to the CPS. By collecting a small number of additional items of information from households, a great deal can be learned about the homeless and the precariously housed. As existing studies of the homeless consistently show, a large proportion of homeless people either are temporarily in that condition or move back and forth between the housed and homeless state. If modifications were made to the CPS, those who have been or will be homeless would be identified within the CPS sample, as currently collected. Because the CPS is also a limited longitudinal survey, based on 12 monthly contacts with sample households over a period of about 16 months, it should be possible to detect movement of individuals from a housed state to an unhoused state.[8] Unfortunately, the CPS does not collect information about either the origins or the destinations of those who enter or leave the CPS sample, and hence, little can be learned in this regard. But asking a few questions about each person entering or leaving a sampled household can provide data on this important component of the homeless. Undoubtedly, information about origins may be more reliable than that about destinations, but some information is better than none at all.

It should also be recognized that the CPS does not do very well in consistently detecting and recording "peripheral" household members—that is, persons temporarily or episodically residing in the household or persons whose acknowledged presence in the household may jeopardize the householder, as in the case of "live-in" male companions in households receiving Aid to Families with Dependent Children (AFDC). Clearly, additional efforts should go into improving the enumeration of such peripheral members, especially because the episodically homeless and peripheral household members probably overlap considerably.

The remaining suggestions about additional data collection are more problematic. First, one or two items may be sufficient to obtain prevalence measures of lifetime homelessness by asking adult respondents whether they have ever been homeless.[9]

> Second, consideration might be given to using respondents as reporters about homelessness among their primary kin—for example, asking parents about their children, and/or asking siblings about each other. Properly weighted and adjusted for only children (in the case of siblings), estimates can be constructed of the number of homeless persons known to be homeless by their primary kin. Of course, measures of this sort rest heavily on the assumption that every homeless person has some living primary kin who reside in conventional dwelling units, and that primary kin have accurate knowledge about each other's living arrangements[10] and are willing to acknowledge that one or more of their primary kin are among the homeless. To some degree, these assumptions clearly do not hold. Persons who grew up in single-child households obviously do not have siblings to report on their location. And although, by definition, every person has parents, some parents will be dead, and some of the homeless grew up in foster homes and may have lost any contact with their natural parents.

For presentation purposes, the discussion concerning modifications to existing data collection efforts has centered around the monthly CPS. However, each of the major household-based continuing surveys should be examined to determine how changes in them could lead to significant increments of existing empirical knowledge about the homeless. The addition of data types suggested in component 4 added to the SIPP would provide an even finer grained view of the interchange between the homeless and the housed; this is because SIPP surveys the same households every quarter over a two-and-a-half-year period. The AHS might also prove to be usable, especially because it has dense samples of six major metropolitan areas.[11] Since it already obtains information on individual conditions associated with homelessness, such as physical disabilities, mental illness, and substance abuse, adding to the National Household Survey on Drug Abuse appears especially attractive.[12]

The need for longitudinal studies

With a few notable exceptions, almost all studies of the homeless are based on surveys that are cross sections in time. These studies have been extremely important in describing the conditions in which the homeless live, the various maladies with which they are inflicted, and their social origins. The studies have also led to the realization that, for many homeless persons, homelessness is a temporary or episodic state. Those who have one or more such episodes in a year far outnumber those who are homeless throughout a year.

The high rates of population turnover among the homeless hold several implications for cross-sectional surveys. First, because the probability of being selected in a cross-sectional survey is proportional to the amount of time spent in the homeless state, such surveys do not fairly represent all who become homeless in any month or longer period. Instead, existing surveys overrepresent the long-term homeless and are dominated by their characteristics. Second, the turnover reveals that there are many recoveries from homelessness, short-term episodes of homelessness, and a significant number of persons who appear to be episodically homeless. The circumstances that lead an individual or a family to cross the easily permeable border between the housed and the homeless state should be of considerable policy interest. If more were known about these circumstances, it might be possible to identify leverage points and strategies to prevent the occurrence of a homeless episode. The potential fruitfulness of this kind of information is amply illustrated in Eric Lindblom's article, in which he proposes preventive strategies based on fragmentary knowledge about transitions into homelessness. Cross-sectional surveys do not provide the necessary information with sufficient precision.

Accordingly, a priority in research ventures into homelessness should be given to longitudinal studies, in which persons are tracked over the periods of time during which they become homeless, while they are in that condition, and when they leave homelessness. Among the life-course circumstances and events to be studied, special attention should go to those that affect income and relations with primary kin: It is income that conditions a person's ability to obtain conventional dwelling units, and there is abundant evidence that the primary kin of poor persons play a critical role in providing them with housing. Of course, any such study should also pay close attention to the roles played by social programs in such transitions.

All that said, the obstacles that lie in the way of successfully completing longitudinal studies of homelessness are considerably greater than the pitfalls that plague the roads taken by ordinary longitudinal studies of households and individuals. Some encouragement can be taken from the fact that several longitudinal studies of the homeless have been conducted. Piliavin and Sosin tracked a small sample of homeless persons in Minneapolis with moderate success.[13] Knickman and Weitzman used AFDC administrative records and personal interviews to track AFDC families through episodes of homelessness.[14]

The major obstacle to longitudinal studies of the homeless is the same one that impedes all surveys of this population. If it is difficult to locate these people for a cross-sectional survey, that difficulty is compounded when the task involves repeated contacts. Such studies are not impossible, only very difficult and therefore very expensive. For example, a common interviewing practice to facilitate follow-up is to obtain the names, addresses, and phone numbers of persons with whom a respondent is in fairly enduring contact, and then to use the designated persons as sources of follow-up clues. Tracing a respondent who is not at the place where previously contacted requires contacting all those persons earlier designated as enduring contacts. Whether this interviewing tactic will work with homeless persons is problematic. The efficacy of offering respondents incentives to maintain contact with a research project needs to be explored. Such incentives could include a sliding scale of payments that increase with each successive contact.

All that said, there are several kinds of longitudinal studies, some more feasible to carry out than others:

1. Prospective "entry-to-exit" longitudinal studies

   The most elaborate of longitudinal studies would start with a sample of persons at high risk of becoming homeless. It would contact these persons periodically over at least two years, noting those who become homeless and, among them, those who remain homeless and those who leave the homeless condition after the completion of an episode. A major difficulty for this kind of study lies in identifying a high-risk population at the outset. Even among the very poor, the risk of homelessness is very low. For example, in Chicago in 1986, close to 200,000 single adults had incomes less than 51 percent of the poverty level, but only 5,000 to 6,000 persons became homeless during that year,[15] suggesting that the risk of becoming homeless among this tragically destitute population is less than 1 in 20. Were the risk levels higher—say, 1 in 5—the screening task would still be formidable; but even so, there is no firm knowledge that those living below 51 percent of the poverty level are the sole recruits into homelessness. Indeed, until we know more about the risks of becoming homeless among various social groups, it seems impractical to implement an entry-to-exit longitudinal study.[16]

   There is considerably greater promise in applying the entry-to-exit design to groups who are already being tracked through administrative records or an ongoing longitudinal study. Indeed, one of the few examples of an entry-to-exit study used the AFDC record system to identify families who become homeless in New York.[17] AFDC families are a group at relatively high risk and comprise most of the families who show up in family homeless shelters in that city. Knickman and Weitzman's study was further facilitated because the New York City welfare department runs an emergency housing assistance (EHA) center to which AFDC families can go for direct assignment to family shelters.[18] The existence of the EHA provided the researchers with administrative data that recorded entry into homelessness. For other components of the homeless population, especially single persons, no comparable gateways to homelessness exist.

Taking advantage of existing longitudinal studies is another strategy worth investigating. For example, over a l2-month period, close to one-third of the 1,800 extremely poor, unattached adults "disappeared" from the longitudinal SIPP.[19] These were persons whom the SIPP interviewers could not trace. Of course, not all became homeless; some simply moved to another household. However, given their poverty and accompanying long-term unemployment, many did become homeless. By adding a few questions to the SIPP interview schedule about persons leaving or entering the SIPP sample, researchers could gain information on those transitions.

2. Semiretrospective entry-to-exit studies

It may be difficult to locate persons who are going to become homeless; it is easier to locate persons who are now homeless and to follow them over time until they exit from homelessness, reconstructing their prehomeless experiences by asking for retrospective accounts. There are good reasons to suspect the validity of retrospective data, given what is known about selective recall and time "telescoping,"[20] and it is likely that only salient past events may be reliably reported. For example, whether a person has been evicted from an apartment is probably remembered more clearly than how diligently the person conducted job searches when unemployed "last year." Even if retrospective data were not held in suspicion, there is still the issue of what to compare the prehomeless state with, as revealed in retrospective accounts. If, for example, 10 percent of the homeless were evicted from their last apartments, it is uncertain whether that rate is higher than what is the "normal" experience for persons of comparable income in the same housing market.

3. Full retrospective longitudinal studies

   Least attractive but perhaps most feasible is to collect from persons who are currently housed retrospective accounts of becoming homeless and leaving homelessness. This approach consists of identifying housed persons who have been homeless and reconstructing the history of events leading to their becoming homeless and surrounding their leaving that state.

   There are several difficulties with this approach: First, there is the well-known problem of the validity of retrospective accounts. Second, although literally hundreds of thousands of persons have been homeless over any year in the 1980s, homelessness is still rare, occurring to perhaps one in a thousand adults over a year's time. Of course, lifetime prevalence proportions will be higher, but they are unlikely to be much higher than one in a hundred. Locating and identifying those persons who have been homeless is clearly a massive and difficult task, especially if concern is restricted to those who have been homeless recently. Assuming that about one in a thousand adults were homeless last year and are now housed, 100,000 adults would need to be screened to locate 1,000 eligible respondents. Clearly, this is not a task for a stand-alone survey but one that could be added to a very large, ongoing survey such as the CPS.[21]

   Third, those who have visited homelessness and returned to the housed condition may not be representative of all homeless persons. Accordingly, persons who remain in the homeless condition for long periods of time would not be properly represented. Finally, it is not clear how to interpret findings because they are not easy to calibrate. For example, if fewer than 10 percent of the recovered homeless had lived in foster homes as children, the significance of that finding could not be judged without knowing the prevalence of foster care either among the general adult population or among the homeless. Thus, this variety of longitudinal study should lay heavy stress on data types for which there are existing series to provide the needed comparisons.

Notwithstanding the difficulties of both designing and implementing longitudinal studies of entry into and exit from the homeless condition, the potential payoff for policy formation is considerable enough to justify at least exploratory efforts. For example, if it turns out that many homeless leave that state after successful reconciliations with their primary kin, this suggests programs designed to facilitate such reconciliations, perhaps including financial aid to the kin involved.

Searching for useful proxies

Most of the research efforts discussed earlier are difficult to carry out with great fidelity and thus are expensive, but some are both feasible and relatively inexpensive. Especially attractive is the idea that the easier approaches, properly adjusted, can reasonably approximate the more expensive ones. Surveys of clients of shelters and food kitchens appear to be the most promising. Indeed, the only existing national surveys of the homeless are based on such approaches.[22] As suggested by James in his article in this journal, if the empirical relationships between the total homeless population and those persons served by shelters and food kitchens were known, the relatively inexpensive service-based surveys could be used to estimate both the size of the total homeless population and its composition. The obvious difficulty is that there is only fragmentary knowledge about those empirical relationships, and existing studies cast some doubt on the likelihood of finding such relationships to be consistent over time and across localities.

Shelters and food kitchens are local enterprises. If the capacities of local shelter accommodations and food kitchens were driven solely by the number of homeless persons in a locality, it could be argued that necessary adjustments could be made and the total homeless populations could be projected. Unfortunately, considerable evidence casts doubt on this proposition. Nationwide, shelter accommodations do not appear to be filled to capacity although smaller shelters appear to be operating closer to capacity. Moreover, no comparable operations data exist on food kitchens or medical clinics although there is evidence that both serve a wider clientele than just the homeless. Shelter capacities are at least partially determined by law, as in the case of the New York consent decree that obligates the city to provide for all persons seeking shelter; by the availability of funding from local, state, and federal sources; and by the willingness of local community groups to provide the private funds and volunteer labor. It seems unlikely that such capacities would be sensitively responsive to the size of local homeless populations, especially, as noted earlier, when the size of the homeless population is so difficult to ascertain.

And, as discussed by James, the very vexing problem remains of those homeless who do not use any of the services.[23] No striking differences appear in the composition of this nonservice group when compared with those persons who use shelters or food kitchens, although there is some hint that the nonservice-using homeless have higher levels of substance abuse and chronic mental illness. Even though this judgment is based on fragmentary data, it would probably not be far off the mark to use the service-using homeless to describe the composition of all the homeless.

The main difficulty posed by the nonservice-using homeless is ignorance of the size of this group and whether it constitutes a roughly constant proportion of the total homeless population across localities. What is known is that some homeless individuals consistently live in public places, seemingly avoiding the shelters,[24] and also do not use the food kitchens. It is difficult to understand how homeless persons can continue to get by for any protracted period of time without recourse to such services; their obvious alternatives are to raise money and find food by petty thievery, scavenging, begging, and peddling—precarious sources at best. The point, however, is that the proportion of nonservice users cannot be very large.

Whether studies can rely on service client data as a proxy for more direct enumerations of the total homeless population should be the topic of another set of technical researches. Although a great deal will be learned from the ongoing work by Dennis and his colleagues, some firm data are needed on the local variability in the ratio of the nonserved to the served. It would be even more useful to learn the reasons for that variation, how much influence shelter policies exert, and the role that such disabilities as chronic mental illness may play in keeping some homeless persons from seeking shelter accommodations.

Other forms of proxy seem less promising. An approach that may appear, at first glance, to be fruitful is to develop a model that predicts the size of a locality's homeless population as a function of indicators of other seemingly related social problems, such as unemployment, AFDC loads, and housing vacancy rates. But as indicated by Martha R. Burt's research on intercity variation in local shelter capacity, reported in this journal, this will not be an easy task. First, many indicators are not available on a timely basis. The decennial census rapidly becomes out of date, and census data in appropriate detail are often only available several years after the census date. Second, many of the indicators that appear promising, such as the unemployment rate, are only weakly related to shelter capacity, suggesting that they may also be weakly related to the total size of the homeless population.

All said, the major obstacle to developing credible proxies is the absence of reasonably sound data on the actual size and composition of the total homeless population. Without estimates of the total homeless population, proxy measures that may be proposed cannot be assessed for adequacy. Accordingly, there is a catch-22 situation: Proxy measures are needed because the most promising research strategies are likely to be very expensive. However, the selection of proxies that are credible must await the full development of the expensive strategies in order for their credibility to be tested.

Evaluation research

Up to this point, this paper has been concerned with research that would broaden and deepen understanding of the homelessness problem. It is hoped that the findings of such research would provide the empirical knowledge base needed to devise appropriate preventive and palliative policies and programs. This section turns to two other important policy functions that social science research can perform: directly helping to construct programs and assessing their effectiveness. These are the provinces of a methodology summarized as evaluation research.

Evaluation research is distinguished from other social science research by the kinds of problems to which it is applied, with its primary focus being on policies and programs rather than on epidemiological issues and basic causal processes. Because evaluation research shares its data collection methods with other forms of social research, the difficulties that plague the application of sample surveys generally also affect the use of sample surveys in evaluations. This section will not repeat those concerns.

Conventionally, evaluation research is divided into two sets of activities: (1) formative evaluation, consisting of research in support of program development and fine-tuning; and (2) summative evaluation, consisting of research concerned with effectiveness—the extent to which a program is achieving the effects it was designed to produce—and efficiency—the extent to which program costs are justiflied by the effects achieved. Formative and summative evaluations are not equivalent. A program can be running quite smoothly and yet be ineffective. Although it is not likely that an administratively chaotic program will be effective, the fact that it is poorly administered is not invariably fatal to its effectiveness. The remainder of this section will treat the two types of evaluations.[25]

Formative evaluations

The researches in this category range widely across data collection methods and in technical sophistication. However, they share one aim: to provide program operators or other sponsors with information about how a program is functioning. Such information typically includes data on coverage and targeting (i.e., whether the program is reaching not only its intended clients but also a significant portion of that clientele), data on services delivered to clients (e.g., number of persons accommodated in a shelter), data on the quality of services (e.g., comparison of services with some standard), and a long list of other data useful for judging how well the program is functioning.

Complex social programs that deliver mixes of services tailored to clients are often difficult to control, describe, and replicate because of their complexity. In addition, when a program is first implemented, it is almost always necessary to modify its procedures to accommodate unanticipated features of clients and program ecology, and it is sometimes necessary to modify its goals as well. The activities that make up formative evaluations are intended to provide managers with information that can control and improve programs and aid in their dissemination.

A critical element of formative evaluation activities is the installation and operation of a management information system that can generate timely information on the kinds of clients in the program, the kinds of services delivered, and associated costs, as well as on how clients progress toward the end of their stay in the program. Especially vital to proper management is the potential to generate information on changes occurring in a program in time to adjust the program accordingly.

Beyond the "head counting" provided by management information systems, it is often useful to have more qualitative information to determine whether services are being delivered at the proper dosages to clients who need and/or want them, and to detect needs that may be going unmet or services that are unneeded. Various approaches may be used to gather such data, ranging from formal surveys, to focus groups composed of clients and frontline workers, to ethnographic studies conducted by trained qualitative researchers.

In principle, the development phase of each new program should be accompanied by formative evaluation activities. In practice, however, because development is not ordinarily well funded, many programs are developed with only primitive forms of management information systems in place, mainly to serve fiscal accountability purposes. Accordingly, it is typical for programs to be evaluated formatively at some stage beyond early development. Indeed, it is likely that most of the so-called evaluations are descriptions of how the program is operated plus a brief history of the program.

This author advocates that every substantial[26] program develop formative evaluation activities. Program managers can benefit considerably from having a precise understanding of how the program functions. Almost all programs can benefit from fine-tuning, and managers could profit from advance warning of program defects. Good descriptive materials are needed to disseminate a program to other sites. Finally, data from formative evaluations can answer critical accountability issues regarding proper targeting and coverage.

Summative evaluations

The ultimate evaluation question concerns the extent to which a program achieves its stated goals. If humans were inert materials, it would be much easier to answer this question. However, when dealing with people who change spontaneously over time, the question has to be restated: Did the clients of a program improve more than they would had they not come in contact with the program? This formulation is especially critical in studying human services programs that deal with people in troublesome crises. Many people who are in trouble at one point in time extricate themselves from that condition with the passage of time and without services. In addition, social programs in this country do not exist in a social welfare vacuum: Other programs provide some services to persons not served by a given homelessness program. The seemingly spontaneous mutability of people means that the evaluation of social programs can be accomplished only by comparing what happens to persons who have gone through a program with what happens to a comparable group of persons who have not had that experience. In addition, for most social programs, clients either seek out or are selectively assigned to a program. The fact that someone seeks services may indicate a willingness to improve and suggest that improvement would probably occur with or without the aid of the program.

In sum, a program is effective to the extent that it improves its clients to a greater degree than comparably motivated nonclients of comparable psychosocial composition improve without the help of the program.[27]

Accordingly, all credible effectiveness assessments are comparative, the degree of credibility varying with the kinds of comparisons made. Because change can only be defined in longitudinal terms, the implication is that evaluations are longitudinal observations. For human services programs, the needed comparisons can be made in a limited number of ways.[28]

1. Reflexive comparisons

   The conditions of clients before entering a program are compared with the conditions of the same clients after they have experienced the program. Improvements thus detected seem to be due to the program. However, such conclusions are usually unwarranted. Reflexive comparisons produce the least credible results because spontaneous changes and those produced by forces or events not part of the program cannot be separated from program-induced changes. Clients change over time for many reasons, and the effectiveness of the experienced program is only one among many.

2. Comparable nonclients

   The changes detected in a client group after completing a program are compared with changes in a "comparable" group of nonclients observed over the same period.[29] The worth of this comparison depends heavily on how comparable the comparison group may be. For voluntary social programs to which clients come seeking services, a critical difference that usually emerges between clients and nonclients is the level of motivation to change, comprising the so-called self-selection bias. Although this kind of comparison produces results that are arguably more credible than those produced by reflexive comparisons, the possibility that self-selection factors may produce observed differences means that findings are not totally credible.

3. Randomized assignment

   Using the model of the randomized controlled experiment, a pool of clients is randomly[30] divided into those served (the experimental group) and those unserved (the control group). Assuming sufficient numbers in both groups, the two groups[31] are

> comparable in every possible way. Changes in the two groups are compared, with the observed differences being regarded as the effects of the program. Because clients are randomly assigned to a program, self-selection cannot be a factor in the observed differences. And because both the experimental and control groups are subject to the same processes of spontaneous change, that, too, cannot play a role in generating differences between them. Thus, randomized experiments produce the most credible effectiveness estimates.

This author is not aware of any completed randomized experiments estimating the effectiveness of any programs for the homeless. However, at least six experiments are now under way, sponsored by the National Institute of Alcohol Abuse and Alcoholism, that are testing the effectiveness of alcoholism treatments on homeless patients. Another experiment, funded by the National Institute of Mental Health, is treating substance-abusing single parents. These sponsors should be applauded for insisting on randomized designs.

Because of the superior credibility of effectiveness estimates derived from randomized experiments, it might be difficult to understand why so few such studies are conducted. The main obstacles to the use of randomized experiments are their cost and the widespread reluctance of human services providers to sanction random (read arbitrary) withholding of services to otherwise deserving and eligible clients.

Although little can be done to reduce the costs of randomized experiments, much can be done to use research funds efficiently, reserving such expensive strategies for circumstances in which they will be most productive. To begin with, randomized experiments should be undertaken only when there is some chance that a significantly effective program can be detected. Accordingly, randomized trials should not be used to test programs that are known to be effective. A shelter is clearly effective at providing indoor sleeping arrangements; thus, no experiments are needed to show that, compared with persons randomly turned away from shelters, those who are admitted are much more likely to have beds to sleep in. A justifiable shelter experiment is one that would compare the effectiveness of one kind of shelter with another; in such a case, both the experimental and control groups would receive sleeping accommodations, and the effect being measured would concern outcomes other than the provision of shelter. Similarly, randomized trials also should not be used to test trivial programs or ones that everyone agrees are likely to be ineffective. For example, a program to help raise the morale of homeless persons by instituting meditation sessions in shelters is so likely to fail that it would be a waste of time and money to mount an experimental evaluation.[32] It also makes little sense to test the effectiveness of programs that are so small in scale that no randomized experiment could have the power to detect effectiveness if it was present. Randomized experiments are for testing the effectiveness of programs that are promising, about which there is some doubt, and for which a randomized experiment of appropriate power can be designed.

As for opposition to randomized experiments on ethical grounds, considerable progress is being made toward specifying the circumstances under which randomization is proper. For example, for programs that add to existing services and are not intended to cover all potentially eligible clients, randomization is an equitable way to ration services, and placement in a control group would not lower the condition of persons so placed. And, of course, as argued before, randomized trials are to be undertaken only when it is uncertain that the program is effective but it is clear that the program has some promise.

A program of evaluation of homeless programs cannot be laid out in any great detail. Literally hundreds of recognizably different programs may now be in place throughout the country, and it would be an Augean task to set up formal evaluations of every one. What is needed is some way of identifying those programs that look promising and can potentially be replicated in other locations. These seemingly successful programs should be evaluated with considerable precision.

Conclusion

An ambitious program of research on homelessness has been laid out here, ranging from technical research aimed at improving data collection efforts to evaluations that test the efficacy of programs for the homeless. Priority should be given to exploring the data potentials of modifications to existing data collection efforts and to developing improved methods for locating and enumerating the homeless. The successful accomplishment of this latter goal would materially help researchers to achieve the important goal of understanding both the composition of the homeless population and the events and circumstances that lead individuals to experience one or more episodes of homelessness. Without much better evidence-based understanding, progress in dealing with homelessness can be made only by trial and error, a costly and wasteful strategy. Without such knowledge, it is difficult to design programs with potentially high payoffs.

Because the technical issues for longitudinal studies are similar to those for cross-sectional studies, the former will also be advanced by assigning the highest priority to a technical research program designed to improve data collection methods for locating and enumerating the homeless. There should be two parts to this technical research: First, existing data collection operations should be expanded to bring in better data within a few years; adding variables would also be relatively inexpensive. Second, a program should develop a better understanding of how to undertake valid censuses and surveys of the homeless population and of what the empirical differences are among alternative definitions of homelessness.

Although evaluation research should play a critical role in the development and testing of programs for the homeless, at present it is not possible to lay out a clear program. In part this is because there appear to be no special technical problems in using current evaluation models. The main obstacle to developing an overall strategy, however, is the lack of knowledge concerning the range of homeless programs and identifying those that appear promising enough to evaluate for effectiveness.

Author

Peter H. Rossi is the Stuart A. Rice professor of sociology and the acting director of the Social and Demographic Research Institute at the University of Massachusetts at Amherst.

Endnotes

1. M. R. Burt and B. E. Cohen, *America's Homeless: Numbers, Characteristics, and Programs That Serve Them,* Report 89–3 (Washington, DC:The Urban Institute, 1989); M. Sosin, P. Colson, and S. Grossman, *Homelessness in Chicago: Poverty and Pathology, Social Institutions and Social Change* (Chicago: Chicago Community Trust, 1988); U.S. Department of Housing and Urban Development, *The 1988 National Survey of Shelters for the Homeless* (Washington, DC: U.S. Department of Housing and Urban Development, 1989).

2. The literal homeless constitute the component of the total homeless population that is the most difficult to accommodate within conventional survey and census approaches. Hence, they are at the center of attention in this section of the paper.

3. This suggests that a slight alteration in data collection practices of such national surveys as the monthly Current Population Survey (CPS) would support estimates of the shelter population. All that is needed is for enumerators to record shelters as a special category of living quarters. Of course, shelters are relatively rare among all dwelling units, and estimates derived from this modification may have very wide confidence bands. (This recommendation is taken up in some detail in a later portion of this paper.)

4. As noted, group quarters—places shared by several unrelated persons—are recognized as conventional dwelling units in most censuses and surveys. For convenience, however, this discussion will consider shelters as NDUs.

5. P. H. Rossi, *Down and Out in America: The Origins of Homelessness* (Chicago: University of Chicago Press, 1989); G. Vernez et al., *Review of California's Program for the Homeless Mentally Disabled* (Santa Monica: RAND Corporation, 1988); Hamilton, Rabinovitz, and Alschuler, Inc., *The Changing Face of Misery: Los Angeles' Skid Row Area in Transition: Housing and Social Services Needs of Central City East* (Los Angeles: Community Redevelopment Agency, 1987).

6. "Windshield surveys" are made by driving through areas and making systematic observations. In this application, observers in the cars would note locations where seemingly homeless persons were congregating. An example of the use of windshield surveys can be found in F. G. Robinson, "Homeless People in the Nation's Capital" (Washington, DC: University of the District of Columbia, Center for Applied Research and Urban Policy, November 1985).

7. G. Vernez et al., *Review of California's Program for the Homeless Mentally Disabled.*

8. In research now being undertaken, this author found that close to 25 percent of extremely poor, unattached individuals (unmarried with incomes less than 51 percent of the poverty level) left the CPS sample over the period of a year and that a similar number entered the sample. Of course, some of this movement is between households, but some is undoubtedly between literal homelessness and being precariously housed.

9. Obviously, considerable thought and pretesting would have to be undertaken to find the proper wording that elicits the information wanted in acceptable ways. For example, it might be more acceptable to imbed such a question in an item asking about a range of lifetime experiences.

10. There is abundant evidence from studies of the homeless that most homeless individuals know the whereabouts of their primary kin, but there is also evidence that significant proportions of the homeless have been reared in foster homes or do not have any still-living primary kin.

11. It is understood that negotiations are currently under way to add a few items to the AHS.

12. Indeed, the technical research currently being conducted at the Research Triangle Institute (and reported in Dennis's paper) is intended to test the feasibility of adding shelter, food kitchen, and NDU components to the National Household Survey on Drug Abuse.

13. I. Piliavin and M. Sosin, "Tracking the Homeless," *Focus* 10, no. 4 (1987): 20–24. Piliavin is currently conducting an additional longitudinal study of the homeless in the San Francisco area.

14. J. R. Knickman and Beth C. Weitzman, *A Study of Homeless Families in New York City: Risk Assessment Models and Strategies for Prevention* (New York: New York University, Health Research Program, 1989).

15. Calculations from the special tabulations from the CPSs for 1985, 1986, and 1987 and from Rossi, *Down and Out in America.*

16. In his paper, Lindblom proposes a short list of "risk factors" that can identify persons at high risk of becoming homeless. His proposal appears very promising as a device for screening potential candidates for longitudinal studies. Of course, because it cannot be known at this point how strongly predictive his suggested high-risk characteristics are, his proposal only appears promising.

17. Knickman and Weitzman, *A Study of Homeless Families in New York City.*

18. Ibid.

19. These are special tabulations prepared by the author from the first three waves of the 1984 SIPP panel. The respondents in question were unmarried adults whose income was below 51 percent of the poverty level.

20. *Telescoping* is the term used to characterize the tendency of respondents to locate inaccurately in time events that occurred in the past. For example, in the National Crime Survey, respondents may claim to have been victimized within the past six months when, in fact, their victimization occurred much earlier.

21. About 117,000 persons aged 15 or older were included in the 1990 March supplement of the CPS.

22. Burt and Cohen, *America's Homeless;* U.S. Department of Housing and Urban Development, *The 1988 National Survey of Shelters for the Homeless.*

23. In the author's own mid-1980s study of the homeless in Chicago, a significant portion (23 percent) of the nonsheltered literal homeless claimed no knowledge of the availability of shelters in Chicago. This finding is of questionable validity, especially because most of the street homeless who claimed no knowledge of the shelters were found in places occupied by other homeless persons who had such knowledge.

24. Almost all shelters have some policies restricting access to their accommodations. Many shelters serve only one sex, and many exclude adolescents. Frequently, mentally ill and/or persons under the influence of alcohol or drugs are also excluded, as well as those who have caused trouble in the past. It is

hard to judge how much influence these exclusions have had in producing the less-than-capacity occupancy or in forming a nonservice-using group of homeless people.

25. P. H. Rossi and H. E. Freeman, *Evaluation: A Systematic Approach* (Newbury Park, CA: Sage Publications, 1989); R. A. Berk and P. H. Rossi, *Thinking about Evaluation* (Newbury Park, CA: Sage Publications, 1990); J. L. Cronback, *Designing Evaluations of Educational and Social Programs* (San Francisco: Jossey-Bass, 1982); T. D. Cook and D. T. Campbell, *Quasi-Experimental Design and Analysis Issues for Field Settings* (Skokie, IL: Rand McNally, 1979).

26. Criteria that lead a program to be called substantial are listed in Burk and Rossi, *Thinking About Evaluation.* However, it is doubtful whether very small programs (e.g., a homeless shelter with accommodations for six persons) should be the subject of a fully configured formative evaluation. In such programs, the program manager should know as much about the program as could be learned through more formal means.

27. In practice, this means that program clients improve more than they would if they experienced other "conventional" programs.

28. There are several additional ways to make comparisons, most of which are not ordinarily applicable to human services programs. See Rossi and Freeman, *Evaluation.*

29. There are several ways of either constructing such comparable groups or "locating" them as naturally formed groups. In testing educational programs, schools may be compared with other schools serving a neighborhood comparable in socioeconomic status and ethnic background. Nonclients may be made comparable through the use of statistical standardizations such as multivariate techniques.

30. Random in this context does not mean arbitrary or haphazard; rather, it means the use of a device that ensures that the probability of being assigned to one or the other group is known, nonzero, and identical within each group.

31. Experiments are not restricted to two group designs. Indeed, some of the largest field experiments contained more than ten experimental groups, each receiving some variant of a program. See R. J. Struyk and M. Bendick, *Housing Vouchers for the Poor: Lessons from a National Experiment* (Washington, DC: The Urban Institute, 1981).

32. An actual example may be more informative. An enacted change in food stamp regulations authorizing food kitchens to accept food stamps was bound to fail to gain acceptance by either food kitchens or homeless persons because it was not accompanied by measures motivating both parties to accept that form of use. Thus, it could not be implemented and was clearly not a candidate for a randomized trial. See M. R. Burt and B. E. Cohen, *Feeding the Homeless: Does the Prepared Means Provision Help?* 2 vols. (Washington, DC: The Urban Institute, 1988), for a description of this program and its failures.