Where to from Here? A Policy Research Agenda Based on the Analysis of Administrative Data

Dennis P. Culhane and Stephen Metraux
University of Pennsylvania

Abstract

This article outlines a policy research agenda based on the analysis of administrative data. Computerized records of client characteristics and their related shelter utilization patterns offer researchers a rich source of longitudinal data that makes possible a wide range of investigations and can be analyzed by using an array of multivariate statistical tools. Specifically, this article discusses the contributions administrative data can make to (1) enumerating and determining the characteristics of the homeless population, (2) understanding the effect of homelessness on related public systems, (3) gauging the effect of policy interventions on the use of homeless services, (4) evaluating the effectiveness of system-level delivery of homeless services, and (5) measuring the performance of individual homeless service providers.

The article concludes by commenting on several issues that policy makers might consider regarding the implementation of automated information systems among homeless service providers.

Keywords: Homelessness; Policy; Methods

Introduction

As data collection on users of homeless services in many U.S. localities becomes automated, the kind of research being conducted on homeless populations will likely change. Both individual homeless service providers and local homeless service systems have increasingly been turning to management information systems (MIS), also referred to as client tracking programs, to standardize their administrative records just as many public agencies have done. Using information systems to manage homeless programs is nothing new; some individual provider organizations have been using off-the-shelf applications or have had customized programs for these purposes for over a decade. There have even been several longstanding municipal and statewide systems (e.g., Columbus, OH; New York City; Philadelphia; Phoenix; St. Louis; and Rhode Island) that have collected uniform data on homeless service use. What is changing is that such systems are becoming both more prevalent and more

standardized, thus making a broader range of research efforts possible on both local and multilocal levels. This research will greatly inform the planning and evaluation of legislation and programs affecting this population.

This article outlines a research agenda to accompany this expanded data capacity. Much of this agenda draws on research conducted at the University of Pennsylvania, where access to several years of administrative data from New York City and Philadelphia has created an opportunity for researchers to try a variety of analytical approaches. The agenda proposes ways to expand our understanding not only of how homeless persons and shelter systems interact, but also of how homelessness affects other related systems such as human services, health care, and criminal justice whose clientele frequently overlap. Specifically, the areas of research considered are (1) population enumeration and composition, (2) integrated database research, (3) time series analysis, (4) program evaluation, and (5) system management and administration. How the analysis of administrative data compares with or can supplement other information-gathering approaches is also discussed. The article concludes by considering measures both to facilitate the establishment of data infrastructures and to realize the potential for policy research that they can provide.

Population enumeration and composition

Perhaps no other research issues have consumed more resources or engendered greater debate than those that address the most basic policy and planning questions in this field: How many people experience homelessness, and what are their characteristics? Over the past decade, cross-sectional methods, also known as point-in-time methods, have become the predominant means of collecting data to describe and enumerate the homeless population. More recently, administrative data have been used to calculate period prevalence rates and to ascertain the characteristics of shelter users over time (see Culhane et al. 1997a).

Cross-sectional surveys of the homeless population elicit information over a very short period, usually between one day and one week. Their results offer point-prevalent "snapshots" and include a variety of specific methods, ranging from straightforward counts and basic sampling (Lee 1989; Metraux 1994; Robertson 1987) to more sophisticated sampling procedures (Burnam and Koegel 1988; Dennis 1993; Rossi et al. 1987). These methods have proved replicable across many localities, as

demonstrated by the Urban Institute (Burt and Cohen 1989) and the Census Bureau (Burt and Taeuber 1991) in estimating or enumerating the population across the nation. Not only can these case-identification methods be applied over a broad variety of jurisdictions for purposes of enumeration, but they can also be used to gather in-depth data on individuals, for example, through the use of standardized diagnostic instruments. Specifically, administering such instruments to large, population-based samples has dramatically improved the accuracy of estimates of the prevalence of mental disorders and substance abuse in the homeless population (Fischer and Breakey 1991; Koegel, Burnam, and Farr 1988).

Cross-sectional studies have been the preferred approach to population enumeration and compositional analyses because they minimize a number of difficulties involved in surveying the homeless population, particularly the risk of double counting and of overrepresenting those who are easiest to find. The approach has been invaluable in producing defensible counts, acceptable to public officials and planning agencies. From the perspective of the information demands of public policy, however, the approach produces very basic data. Unless studies are conducted periodically, the primary use of these data is to inform the immediate planning of program or shelter capacity. Only indirectly can the number of persons homeless over longer periods be estimated on the basis of such data. Moreover, given the study frame (a single day), such data are unsuitable for tracking the service use patterns of subpopulations, for monitoring trends in shelter utilization over time, or for measuring the effect of various programs.

Automated MIS technology provides a promising alternative source of data on population size and characteristics. Homeless service providers are increasingly using MIS for two operational functions: program administration (managing the daily demand for shelter beds, complying with reporting or billing requirements, etc.) and service planning (providing information on individual homeless persons or families for case managers). If reliably maintained, such a system can record the dates and duration of service utilization, as well as a range of additional data on personal characteristics. These systems can thus be structured to produce a client-specific, provider-specific, and/or period-specific longitudinal data archive.

Such longitudinal data are well suited for answering a number of research questions on population enumeration and composition and for using an array of statistical procedures. From an epidemiological perspective, basic demographic data can be used to

calculate prevalence rates by age, race, and sex, and the corresponding relative risk for shelter admission, for any selected time frame in a given jurisdiction (see Culhane et al. 1997a; Culhane and Metraux 1996). Multivariate regression approaches, including ordinary least squares regression, logistic regression, and event history analysis, can be used to estimate predictors of length of stay, repeat shelter use, or housing outcomes (Culhane and Kuhn 1998; Metraux and Culhane 1997; Wong, Culhane, and Kuhn 1997). One can also apply cluster analysis to define distinct patterns of shelter use and to measure their association with demographic characteristics (Kuhn and Culhane 1998). The effectiveness of past or proposed homeless policies on shelter utilization rates (prevalence) or on the service use patterns of subpopulations can also be estimated by analyzing census trends (see the section on time series analysis and Culhane, Metraux, and Wachter 1998).

In describing the advantages of administrative data, it is also necessary to point out certain limitations. Perhaps the most conspicuous disadvantage from an enumeration perspective is that administrative data can provide information only for persons who use the services being tracked. Undercounting can also result from service providers' declining to participate in the data collection system for various reasons (see Culhane et al. 1997a; Rossi 1994). Gaps in coverage and undercounting caused by those who do not use services can be mitigated through the expanded collection of data from tracking street outreach contacts or using survey methods to estimate the number or the proportional size of the non-service-using homeless population.

Another potential disadvantage, particularly for research on population characteristics, is the quality of administrative data. In many cases, data collection is part of an intake procedure in which many individuals and households must be processed quickly and in which interviewers are unlikely to have formal training in interviewing procedures. Given those circumstances, in-depth clinical interviews and lengthy survey instruments prove impractical, so self-reports are usually the only way to collect data on such areas as health status, substance abuse history, or mental health. But self-report questions can be misinterpreted or not answered truthfully when personally identifying data are collected along with data on these potentially sensitive issues. When combined with data from other service systems, however, administrative records can still be used to obtain such data, as discussed in the next section.

Integrated database research

The integration of administrative data for homeless services with data from other service systems offers a means of checking the reliability of self-reported individual characteristics and can also be used to investigate a range of policy-relevant issues. More specifically, integrated database research can examine how the policies and behaviors of other public systems influence the risk of shelter admission and the duration of episodes of homelessness, how homelessness affects individuals' interaction with other public systems, and what the costs that homeless persons incur for other systems are. Thus, data integration can expand tremendously the number and type of research questions that can be pursued and the research that may be directly relevant to the policies of the homeless service system, as well as to those of other public agencies.

A vast potential source of data exists for conducting such research. Because administrative data systems are usually organized by unique identifiers—Social Security number or some combination of name, date of birth, race, and sex—homeless services data can often be merged with a variety of other administrative record systems. A few of the more significant data sources and the research issues they could address are considered here.

Welfare

Perhaps no other policy area has aroused as much concern for its potential impact on homelessness than the recent passage of the Clinton administration's welfare reform legislation. Senator Daniel Moynihan's comment that "we will have children sleeping on grates" indicates the belief that recent changes in the welfare system will lead to an increase in homelessness, so the relationship between welfare and homelessness will undoubtedly figure prominently in future research evaluating the impact of policy changes on welfare recipients.

By matching the identifiers of people terminated from specific welfare programs against shelter registries, a change in the rate of public shelter use among those affected can be directly measured. For example, the shelter admission rate before and after termination of benefits for a particular cohort could be compared, or a group whose benefits were terminated could be compared with a control group of persons who remain on the rolls during a specified period. Culhane et al. (1997c) used both methods to

evaluate the impact on homelessness of two laws recently enacted by the Pennsylvania Legislature that severely restricted access to General Assistance (GA) welfare benefits. Two data sets were merged: one of persons from Philadelphia whose GA benefits were terminated as a result of the two laws and the other of persons using Philadelphia homeless shelters. Using survival curves (Allison 1995), the study found that within a year of termination, the rate of shelter admission among the terminated cases had increased markedly, compared with the persons themselves a year earlier and with a control group from a year earlier.

If demographic or other characteristic variables had been made available for all terminated cases, a multivariate model could have been developed to identify the risk factors associated with an increased probability of shelter admission among terminated cases. Such information would be very useful for targeting assistance to those terminated cases at greatest risk of shelter admission. Future research could similarly compare the impact of caseload terminations in Temporary Assistance for Needy Families, cuts in food stamps, eligibility restrictions in Supplemental Security Income, and other changes on specific cohorts of recipients.

For this type of research to be conducted, public assistance agencies must either do the research themselves or be willing to contribute the identifiers of affected cases to external research organizations. Such data-sharing arrangements will require an officially approved protocol for transferring data and protecting confidentiality. If agencies are unwilling to contribute identifiers, the state or county welfare authority may be willing to conduct the match, mask the identifiers, and then permit the researchers to conduct the analysis on de-identified data. However it is done, research of this nature will be invaluable to policy analysts attempting to measure the impact of these new laws and regulations on homelessness and should be seriously considered in any jurisdiction where automated shelter data are available.

Public health

Public health is another major area of research that can benefit from the integration of administrative data sources. Many public health departments already conduct regular studies of specific population groups by matching their identifiers to vital statistics

or other surveillance databases, and homeless persons are an important population to include in such research. For example, merging homeless data with vital statistics records would enable researchers to study the birth outcomes of pregnant women admitted to public shelters. Such studies could include analyzing how pregnancy outcomes vary as a function of the sequencing of the homelessness episode and the pregnancy, or how pregnant homeless women fare relative to nonhomeless women with similar characteristics (Culhane and Webb 1995). Mortality studies of infants as well as adults could be undertaken to examine any associations between admission to a shelter, duration of stay, or street outreach contact and mortality risk. Merges with other surveillance databases could examine other special populations, including children (vaccination and lead poisoning registries, poison or accidental injury databases), people with HIV/AIDS (Culhane et al. 1997b), or people with tuberculosis.

Research of this nature could affect how public health programs are funded, targeted, and evaluated. Results would be most directly useful for needs assessments and program design, as high-risk groups, including the homeless, are targeted for interventions. Specific cohorts of program participants could be tracked to determine whether existing or new interventions have a mediating effect on the health of homeless persons or on the rate of shelter admission for at-risk populations. A registry of public shelter users at local health departments could be used to identify residentially unstable persons who are consequently eligible for special programs. A prior history of homelessness could also be used as a risk indicator for vulnerable or at-risk populations in a variety of public health planning and research projects.

A prerequisite for conducting such research would be the establishment of protocols for maintaining the confidentiality of the service users, comparable to those used for databases on specific diseases that health departments already maintain. Confidentiality and informed consent are familiar to health officials who deal with this sensitive or legally protected health information. Again, external research organizations would likely require special approval to conduct such research, or else they may have to develop protocols whereby empowered health officials conduct the matching routines and then encrypt, strip, or mask the identifiers before investigators are permitted to perform their analyses.

Health services

Similarly, research can be conducted on the relationship between use of health services and shelter admission. Merges can be conducted with records from such sources as Medicaid claims, Medicare records, specific hospital or HMO files, Health Care for the Homeless databases, and public health center data to examine significant relationships between use of health services and homelessness. Data on diagnoses can be used to estimate the treatment rates for specific disorders in the homeless population and related trends over time. For example, two studies (Culhane, Averyt, and Hadley 1997, 1998) merged three years of Philadelphia shelter registry records with nine years of data on users of mental health services to assess the diagnostic distribution for mental disorders and substance abuse. These studies also used the merged data to examine the interrelationship and sequencing of hospitalizations, emergency room use, and episodes of homelessness.

In another study, Salit et al. (1996) examined records of homeless persons who used Health and Hospital Corporation facilities in New York City to compare the diagnoses of homeless and nonhomeless patients, the differential use of hospital days by diagnosis, and the estimated differences in costs of care. Such a method could also be used to ascertain the proportion of these costs, for both homeless and nonhomeless patients, that are paid by federal, state, or local sources or that are not compensated at all.

Research of this nature is valuable for health departments, state Medicaid offices, managed care organizations, hospital administrators, and other entities that study or plan health services for poor persons. The appropriateness of hospital discharge policies, the adequacy of access to primary health care in shelters, and the need for alternative residential programs are issues that can be studied or informed by this type of research. Patterns of health service use by homeless persons can also inform risk-adjustment procedures for this population, particularly since Medicaid recipients are increasingly required to enroll in privately managed health plans that face cost pressures to restrict hospital stays.

The health services data can also take the place of self-reports or diagnostic survey instruments in identifying subgroups of homeless persons with specific morbidities for more refined analyses. Through the creation of health status risk factors, then, it would be possible to gain a better understanding of how health issues

affect the service system involvement of different types of homeless persons. For example, the shelter stay history of persons with severe mental illness could be tracked to determine the pattern, frequency, and length of stay and the extent to which shelters have become an ancillary feature of, or a surrogate for, other types of treatment.

Again, given the sensitivity of health data and the laws protecting such information, research of this nature must be conducted under carefully designed conditions, with the appropriate authorization from or participation of public agencies.

Housing

Housing is another major public policy area that can benefit from this type of research. In this case, merging by individual identifiers may not be as useful as merging by unique *geographic* identifiers. Homeless services data systems will often collect prior address information or a residential history as part of the eligibility determination or assessment process. This information can be used to identify the characteristics of a client's previous housing or neighborhood. For example, most cities have automated records on property taxes and delinquencies, tax foreclosures, and ownership; others may have automated data on building type, age, history of code violations, and so on. Databases of public safety records, such as crime and fire data, social service and welfare program utilization data, public health data, court-ordered eviction data, and public or private utility termination data, are all likely organized by address or include an address field, as well as the dates of various actions.

By using the digital geography of a particular area (by block or parcel) and an address-matching routine (usually built into a geographic information system program, such as ArcView), researchers can match these addresses and create geography-specific attribute tables that can then be used for further analysis by property or other geographic units, including blocks, block groups, census tracts, or administrative districts and boundaries (health districts, police precincts, city council districts, etc.). For example, Culhane, Lee, and Wachter (1997) used multivariate analyses to examine whether neighborhood variables could account for the distribution of homeless families' prior addresses by census tract in Philadelphia and New York City.

Such research would be valuable for designing, targeting, and siting homelessness prevention programs. Multivariate studies

of neighborhood-level effects could inform the allocation of housing development resources or suggest the potential benefits of other policy strategies, such as those targeting crime prevention, mobility, antiabandonment, or fair housing enforcement, as a means of reducing residential instability and homelessness. Analyzing household migration patterns prior to shelter admission can help shape public policy interventions toward the residential instability process, of which homelessness is but one facet (Lee and Culhane 1996). Dynamic models of neighborhood change, including abandonment trends, gentrification, and neighborhood deterioration, could also be studied for their impact on shelter admissions (Culhane, Lee, and Song 1997).

One of the benefits of using a spatial database is that new area variables (e.g., code violations, tax delinquencies, crime) can take on geographic configurations different from those that have traditionally been available through census data. Customized areas of analysis can replace such standard and somewhat arbitrary units as census tracts, can be created to conform to underlying parcel-level data, and can also provide time intervals that are far more frequent (even continuous) than decennial census intervals. This capacity enables researchers to develop potentially rich, textured, and dynamic models of how housing and neighborhood conditions influence the frequency of housing emergencies and shelter admissions over time.

Criminal justice

People being discharged from jails and prisons are at high risk of shelter admission. Conversely, homelessness, particularly street homelessness, is likely a significant risk factor for arrest and incarceration. Some shelter operators have also reported an increased incidence of persons' being discharged from prisons directly to shelter facilities (Walker 1997). The nature of these pathways and the frequency with which they are traversed are potentially discernible through the integration of homeless services databases with arrest, conviction, and incarceration records.

Much like research on health services, this research could examine how people move between these systems, the costs of their movement, and the characteristics of persons at risk of homelessness or incarceration. The use of shelters as halfway houses by people coming from jails and prisons is of significant concern to agencies that view their mission as serving primarily the indigent poor and do not want to assume the responsibilities

of serving ex-convicts. Municipal jails have also increasingly been forced to provide medication and mental health services to prisoners, possibly reflecting a heightened visibility of homeless mentally ill persons on the streets. Combining administrative data from homeless, health care, and criminal justice services could document the financial and human costs of such shifts in service delivery and provide a basis for designing more efficient alternatives.

Additionally, administrative data could help in assessing the relationships between homelessness, law enforcement, and crime. On one hand, homelessness is commonly associated with increases in both nuisance crimes committed by homeless persons attempting to fashion a private life in public spaces and more serious crimes committed against people or property. On the other hand, there is concern that homeless persons receive unfairly harsh treatment from police and the courts. Little empirical evidence currently exists to substantiate either position, but a better understanding of such links between homelessness, crime, and the criminal justice system would clarify both public safety issues and the need for providing better, more appropriate services in this area.

Among all the areas discussed here, integrating databases in criminal justice has perhaps the most potential pitfalls. There is undoubtedly a risk that an established association between crime and homelessness could further stigmatize the population and perhaps even lead to punitive measures. Establishing protocols that protect the confidentiality of the records used also takes on an added urgency, as merging shelter and police records could reveal information that might lead to the arrest of particular persons who use shelters. These risks have to be weighed against the potential benefits and considered in light of the legal and political context in which this work is conducted.

Limitations on integrated database research

As shown above, integrating databases offers a promising way to inform both research and policy in areas that are undoubtedly affected by homelessness. Gaining access to data from other agencies is often fraught with difficulties, however, and in many cases will be impossible. Legal restrictions present one obstacle that may prevent access to a particular data set. Another source of difficulty, of a more political nature, is obtaining the cooperation of agency heads, who will often decide whether to participate in data sharing on the basis of perceived self-interest for

the agency or the current political administration. Technically, sharing databases often requires compatibility between different computer systems as well as the availability of information system personnel with the requisite time and technical skills. Finally, integrating data systems frequently necessitates the concurrence of shelter system administrators, directors of homeless programs, and services consumers that this research benefits them. Failure to meet any one of these conditions can stymie a request or introduce complex delays.

Time series analysis

Integrated database research, while perhaps the most powerful policy analysis tool associated with administrative data on homeless services, cannot address all the critical policy questions in this field. Many of the people affected by public policies, particularly those that restrict eligibility for welfare or health services, will not be registered in the corresponding administrative data systems. Their eligibility may have been denied, or they may never have applied for social benefits because they perceived that they would be denied them or punished for receiving them. Over time, this population of persons would likely grow larger than the population of terminated cases for which identifiers exist. Many social policies also have indirect or secondary effects that would not be measured or discernible by studying the direct effects on individuals in administrative record systems. For these reasons, other policy analysis tools that assess the aggregate impact of policy changes will need to be used. Again, administrative data are ideal for this purpose.

Time series analysis is one tool that can enable researchers to model the trend in a particular phenomenon, such as shelter admission rates, and to test the impact of a policy change or other events occurring along a time continuum, such as welfare caseload terminations, on those rates. Time series analysis takes into account both seasonality (common with shelter admissions) and preintervention trends in the data (e.g., increases in shelter admission rates may precede any effect from welfare cuts) and can specify the lag before the impact of a policy might be felt. Dependent variables of interest in such research might be the monthly rate of shelter admission, the readmission rate, the average length of stay and daily census, and the proportion of shelter users with certain characteristics.

Time series methods would enable a researcher to measure not only the aggregate impact of other system changes, such as

welfare policy, but also the impact of *shelter system* policy changes. For example, if eligibility rules for shelter admission are tightened, this would likely lead to a decrease in admissions, but it could also reduce pressure on providers to discharge or place clients in housing and thereby increase the average length of stay. It could also change the case mix of persons admitted to shelters, increasing the population at greater risk for a long shelter stay and increasing the systemwide average length of stay. Administrative data, as opposed to data collected through traditional survey research methods, provide multiple measurements of variables over time and thus allow for the possibility of using time series analysis techniques to test such hypotheses.

New York City provides one example of how municipal homeless policy has been affected by a variety of noteworthy events, including a series of court cases, the availability of different types of subsidized housing programs and shelter accommodations, and three different mayoral administrations (Culhane, Metraux, and Wachter 1998). These dynamics continue, with recent changes including a diversion policy under which a family is denied shelter accommodations if on intake it is determined to have alternative housing options (Dehavenon 1996). Time series analysis would make it possible to estimate the significance of any association between these events and various indicators of shelter utilization.

Program evaluation

Program evaluations in the area of homeless services, as in most social service areas, are few and far between. Resources for services are usually slim and demand is great, so funding for program evaluations is typically scarce. Ideally, the evaluations involve experimental study designs to determine the outcome differences between two groups of randomly selected and assigned persons, where the only difference between the groups is the effect of the program in question. Unfortunately, experimental study designs are not always feasible, particularly if the intervention, such as a subsidized or supported housing placement, is costly to provide or has too little turnover to enroll enough subjects, or if random assignment is viewed as unethical or impractical. The controlled implementation of a particular intervention is also not always feasible in an applied setting, where funding constraints may require delayed, partial, or staged implementation or where sudden shifts in policy may mandate a shift in program capacity or activities. These factors typically mean that a relatively small number of persons

are enrolled and followed in homelessness program evaluation research, if and when it is conducted.

Alternatively, even if an experimental design is not in place, the ongoing activities of a *system* of service providers, including the creation of special programs, can be evaluated if administrative data on services used and persons served are routinely collected. Administrative data systems can reduce the collection burden by using the operational infrastructure to record service activity and track client progress. Merges with other administrative data systems can examine the collateral impact on medical costs, incarceration rates, public assistance receipt, and so on. The lack of random assignment can be compensated for through quasi-experiments using the demographic characteristics of large numbers of subjects enrolled systemwide to control statistically for any preexisting differences in subject selection. Moreover, the large number of subjects potentially enrolled can afford more statistical power for determining the characteristics associated with a particular outcome by program type. Selected interviews or small-sample surveys can be used for supplementary information or validation studies. While such a method is not ideal for evaluating a program, the trade-off with statistical power, practicality, cost, and reduced data-gathering burden may make it the only feasible approach in many circumstances.

Such a research approach can be used to evaluate a broad set of program activities, including the effectiveness of "aftercare" case management programs in reducing shelter recidivism, the cost and benefit of supported housing programs relative to continued homelessness, the relative effects of transitional housing versus subsidized housing placements, and the effectiveness of various efforts to prevent homelessness. Hence, the availability of an administrative data system can help make such program evaluations a routine part of policy evaluation and program planning.

System management and administration

Managers of municipal shelter systems or networks of providers may wish to understand not only systemwide patterns of utilization and the effect of various policies and programs, but also the performance of providers in meeting program objectives. Administrative data can be used to compare various provider organizations and administrative units within them on performance objectives. For example, a system manager may wish to examine how well a set of shelter providers is moving families through its programs and into independent housing. Data on average length

of stay can be used as a proxy for resource inputs and rates of housing placement as a proxy for program outputs. Other outcome measures (e.g., client satisfaction, residential stability, and income) could be used as well. Both average length of stay and readmission rate could also be used as dependent variables in a "best practice frontier" analysis to compare providers on these performance measures while controlling for differences in case mix and service provision.

Other management research techniques could be applied to study staffing patterns across the service system, to assess how funding changes affect staff qualifications or turnover. Data gathered about provider revenues and expenditures could also be matched to service utilization data to observe how differences or changes in funding patterns affect service delivery. For example, if federal funding drops, does this correspondingly increase dependence on other funding sources, thereby reducing the provision of certain types of services and affecting client outcomes?

Finally, these administrative data systems could be used to plan, implement, and monitor changes in reimbursement mechanisms or to create performance-based contracting systems. For example, some service systems may want to create incentives for providers to focus on timely discharge and resettlement assistance, as opposed to receiving reimbursement based solely on a client's continued stay. Administrative data can be used both to model the appropriate reimbursement rate and to devise case-mix adjustments. Once developed, administrative data can be used to monitor how providers and clients fare under such a system and to support modifications.

The chief benefit of these management research and planning approaches is that service system managers can increase the accountability of provider organizations and truly move the system toward agreed-on program objectives. Continuous shelter stays for which a provider can be continuously reimbursed could lead to the creation of dependence that is contrary to clients' best interests. Also, inappropriately long shelter stays deprive other clients of system resources, and failing to provide oversight and accountability and to measure outcomes could compromise the political viability of funding. If shelters are perceived as wasteful warehouses that accomplish little or nothing at great expense, they will lose their political appeal and could suffer funding cutbacks. Alternatively, well-managed programs that efficiently help people obtain permanent housing may be able to use those

demonstrated outcomes to request increased funding and expanded programs.

Conclusions

The use of administrative data has great potential for shaping new research opportunities and public policy analysis in the area of homeless services. The research approaches described here, which take advantage of advances in MIS technology that enable quick compilation and efficient manipulation of large data sets on service use, have already shown that they can provide detailed, timely, policy-relevant information unavailable through traditional research methods. As homeless service agencies increasingly use MIS for record keeping, the data are, in many cases, already available, so performing the analyses described here would entail little additional cost.

Implementing an administrative information system for homeless services can be more difficult, however. One of the perceived barriers is the hardware, software, and system administration costs. While these systems can be expensive, their cost can be kept at reasonable levels, as many agencies have shown by their use of innovative, inexpensive tracking systems. In some cases, these systems are as simple as maintaining paper intake and census records at the provider level and sending them once a month to a central data entry site running standard database software on a single PC. With some modification, this is much like the system that has operated for more than six years in Philadelphia and has permitted a broad set of research and policy analysis projects. Of course, more complicated network configurations are also possible.[1]

Another frequently encountered problem is that homeless service providers, system managers, and clients are skeptical that benefits would outweigh the costs and additional work demands of implementing and maintaining an automated data collection system. This article has outlined the benefits of such a system mainly for researchers, policy analysts, system planners, and administrators. But a well-designed system can also make the day-to-day tasks of client and program management easier as it

[1] Dennis Culhane has been funded by the U.S. Department of Housing and Urban Development, the U.S. Department of Health and Human Services, the Fannie Mae Foundation, the University of Pennsylvania, and PRWT Services, Inc., to lead the development of a standardized homeless services software application known as the ANCHoR System (Culhane and Smith 1997).

helps with previously gathered information and produces useful output and reports. Clients benefit from fewer redundant services and forms to complete, a better-organized service system, better classification of their own personal information for case planning, and better accountability on the part of those intended to serve them.

From a research perspective, while the implementation of an information system will yield larger amounts of richer data, it will supplement, not supplant, other methods for gathering primary data. Primary data collection will continue to be the most effective method of carefully exploring client opinions, beliefs, characteristics, and so on. Administrative records also fail to capture what happens to people when they are not in contact with the service system, thus leaving major gaps in our understanding of informal support systems and resettlement patterns. Ideally, more in-depth interviews and follow-up studies on cases randomly selected by an administrative data system should be conducted. In this way, both basic research and more in-depth issues could be explored systematically, and results could be imputed for the larger population from which the sample is drawn.

In conclusion, administrative data systems have potential to inform public policy and research in the area of homeless services at a level of detail, cost-effectiveness, and timeliness not available with other methods. Federal, state, and local governments, which stand to gain an effective tool for evaluating both policy and program performance, should encourage the creation and implementation of such systems among local homeless service providers. As systems proliferate, government agencies will also have to consider the need for standards so that data from different sites are comparable and can be meaningfully aggregated across sites or jurisdictions. Also, setting up such systems will require establishing and adhering to protocols for maintaining the confidentiality and proper use of client records across agencies and research entities. Conducting research based on administrative data also invites collaboration between universities and public agencies to create dialogue on the policy implications such research would hold. Given the rapid change in the social welfare policy environment of the past two years, the creation of an information and analysis infrastructure would greatly enhance our understanding of the nature and causes of homelessness and possible solutions for the future.

Authors

Dennis P. Culhane is an Associate Professor of Social Work and Stephen Metraux is a Doctoral Candidate in Sociology at the University of Pennsylvania.

References

Allison, Paul D. 1995. *Survival Analysis Using the SAS System: A Practical Guide.* Cary, NC: SAS Institute.

Burnam, M. Audrey, and Paul M. Koegel. 1988. Methodology for Obtaining a Representative Sample of Homeless Persons: The Los Angeles Skid Row Study. *Evaluation Review* 12(2):117–52.

Burt, Martha, and Barbara Cohen. 1989. *America's Homeless: Numbers, Characteristics, and Programs That Serve Them.* Washington, DC: Urban Institute Press.

Burt, Martha, and Cynthia M. Taeuber. 1991. Overview of Seven Studies That Counted or Estimated Homeless Populations. In *Conference Proceedings for Enumerating Homeless Persons: Methods and Data Needs,* ed. Cynthia M. Taeuber, 31–75. Washington, DC: U.S. Department of Commerce, U.S. Bureau of the Census.

Culhane, Dennis P., June M. Averyt, and Trevor R. Hadley. 1997. The Rate of Public Shelter Admission among Medicaid-Reimbursed Users of Behavioral Health Services. *Psychiatric Services* 48(3):390–92.

Culhane, Dennis P., June M. Averyt, and Trevor R. Hadley. 1998. The Treated Prevalence of Behavioral Health Disorders among Adults Admitted to the Philadelphia Shelter System: Results from the Integration of Longitudinal Data. *American Journal of Orthopsychiatry,* forthcoming.

Culhane, Dennis P., Edmund F. Dejowski, Julie Ibañez, Elizabeth Needham, and Irene Macchia. 1997a. Public Shelter Admission Rates in Philadelphia and New York City: The Implications of Turnover for Sheltered Population Counts. In *Understanding Homelessness: New Policy and Research Perspectives,* ed. Dennis P. Culhane and Steven P. Hornburg, 101–34. First published in 1994 as *Housing Policy Debate* 5(2):107–40.

Culhane, Dennis P., Erica Gollub, Mark Shpaner, and Randall S. Kuhn. 1997b. The Co-Occurrence of AIDS and Homelessness in Philadelphia: Results from the Integration of Administrative Data. Unpublished paper. University of Pennsylvania.

Culhane, Dennis P., Meg Koppel, Stephen Metraux, and Irene Wong. 1997c. *Mitigating the Impact of State Welfare Cuts for Single Adults: The Implementation and Utilization of the Homelessness Prevention Pilot Project in the City of Philadelphia.* Philadelphia: University of Pennsylvania, School of Social Work.

Culhane, Dennis P., and Randall S. Kuhn. 1998. Patterns and Determinants of Shelter Utilization among Single Adults in New York City and Philadelphia. *Journal of Policy Analysis and Management,* forthcoming.

Culhane, Dennis P., Chang-Moo Lee, and Dayoung Song. 1997. Spatio-Temporal Analysis of Neighborhood Dynamics and Homelessness Incidence: The Effect of Neighborhood Deterioration, Gentrification, and Minority Segregation. Unpublished paper. University of Pennsylvania.

Culhane, Dennis P., Chang-Moo Lee, and Susan M. Wachter. 1997. Where the Homeless Come From: A Study of the Prior Address Distribution of Families Admitted to Public Shelters in New York City and Philadelphia. In *Understanding Homelessness: New Policy and Research Perspectives,* ed. Dennis P. Culhane and Steven P. Hornburg 225–63. First published in 1996 as *Housing Policy Debate* 7(2):327–65.

Culhane, Dennis P., and Stephen Metraux. 1996. Assessing Relative Risk for Homeless Shelter Usage in New York City and Philadelphia. Paper presented at Housing Is Not Enough: Helping Families Achieve Self-Sufficiency, Family Impact Seminar, Washington, DC, September 13.

Culhane, Dennis P., Stephen Metraux, and Susan Wachter. 1998. Homelessness and Public Shelter Provision in New York City. In *Housing and Community Policy in New York City: Facing the Future,* ed. Michael Schill. Albany, NY: SUNY Press, forthcoming.

Culhane, Dennis P., and Kirby Smith. 1997. *The ANCHoR System for Homeless Services: An Information System for the Continuum of Care (Resource Guide).* Philadelphia: University of Pennsylvania and PRWT Services, Inc.

Culhane, Dennis P., and David Webb. 1995. Homelessness and Birth Outcomes in Philadelphia. Paper presented at the Annual Meeting of the American Public Health Association, October 25, San Diego, CA.

Dehavenon, Anna Lou. 1996. *From Bad to Worse at the Emergency Assistance Unit: How New York City Tried to Stop Sheltering Homeless Families in 1996.* New York: Action Research Project on Hunger, Homelessness, and Family Health.

Dennis, Michael L. 1993. Coverage of a Service-Based Methodology: Findings from the DC*MADS Homelessness Study. Paper presented at Towards Census 2000: Research Issues for Improving Coverage of the Homeless Population, September 28–29, Arlington, VA.

Fischer, Pamela J., and William Breakey. 1991. The Epidemiology of Alcohol, Drug, and Mental Disorders among Homeless Persons. *American Psychologist* 46(11):1115–28.

Koegel, Paul, M. Audrey Burnam, and Rodger K. Farr. 1988. The Prevalence of Specific Psychiatric Disorders among Homeless Individuals in the Inner City of Los Angeles. *Archives of General Psychiatry* 45(12):1085–93.

Kuhn, Randall, and Dennis P. Culhane. 1998. Applying Cluster Analysis to Test a Typology of Homelessness by Pattern of Shelter Utilization: Results from the Analysis of Administrative Data. *American Journal of Community Psychology,* forthcoming.

Lee, Barrett. 1989. Stability and Change in an Urban Homeless Population. *Demography* 26(2):323–34.

Lee, Chang-Moo, and Dennis P. Culhane. 1996. Locating the Homeless: A Philadelphia Case Study. *GeoInfo Systems* 5(7):31–34.

Metraux, Stephen. 1994. *Enumerating the Homeless in Tarrant County, Texas: Practical Methods and Reliable Results.* Unpublished master's thesis. University of Texas at Arlington.

Metraux, Stephen, and Dennis P. Culhane. 1997. Family Issues and Recurring Homelessness among Women in New York City Homeless Shelters. Paper presented at the Eastern Sociological Association Annual Meeting, April 9–13, Baltimore, MD.

Robertson, Michael O. 1987. *Homelessness in Albuquerque.* Albuquerque, NM: City of Albuquerque, Department of Human Services.

Rossi, Peter H. 1994. Comment on Dennis P. Culhane et al.'s "Public Shelter Admission Rates in Philadelphia and New York City: The Implications of Turnover for Sheltered Population Counts." *Housing Policy Debate* 5(2): 163–76.

Rossi, Peter H., James D. Wright, Gene A. Fisher, and Georgianna Willis. 1987. The Urban Homeless: Estimating Composition and Size. *Science* 235:1336–41.

Salit, Sharon A., Evelyn M. Kuhn, Arthur J. Hartz, Jade M. Vu, and Andrew L. Mosso. 1996. The Healthcare Costs of Homelessness: Hospitalization Rates and Length of Stay among Homeless Persons in New York City. Paper presented at the annual meeting of the American Public Health Association, November 17–21, New York.

Walker, Adrian. 1997. Summer Fails to Thin Shelter Crowds: Youths, Freed Prisoners Said to Increase Shelter Populace. *Boston Globe,* June 18, p. B1.

Wong, Yin-Ling I., Dennis P. Culhane, and Randall S. Kuhn. 1997. Predictors of Exit and Reentry among Family Shelter Users in New York City. *Social Service Review* 71(3):441–62.