

Research on Homelessness : Sources and Implications of Uncertainty

SCOTT T. FITZGERALD, MACK C. SHELLEY II and PAULA W. DAIL

American Behavioral Scientist 2001 45: 121

DOI: 10.1177/00027640121957051

The online version of this article can be found at:

<http://abs.sagepub.com/content/45/1/121>

Published by:



<http://www.sagepublications.com>

Additional services and information for *American Behavioral Scientist* can be found at:

Email Alerts: <http://abs.sagepub.com/cgi/alerts>

Subscriptions: <http://abs.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Citations: <http://abs.sagepub.com/content/45/1/121.refs.html>

Research on Homelessness

Sources and Implications of Uncertainty

SCOTT T. FITZGERALD

University of Iowa

MACK C. SHELLEY, II

PAULA W. DAIL

Iowa State University

The difficulties surrounding research on homelessness are numerous and substantial. Using a statewide census of homelessness, this article analyzes and critiques common methodological techniques employed in that study. In addition, using cross-tabulation and loglinear modeling, the relationship between key demographic variables and the cited primary cause of homelessness, as well as the types of housing needed, are assessed. Through a multiphased process, which isolates the effects of the operationalized definition of homelessness, significant changes emerge. The operational definition of homelessness is found to affect the demographic composition of the sample, the estimation of annual incidents of homelessness, and the estimation of the statewide number of homeless individuals. The impact of the uncertainty surrounding research on homelessness on policy formation is addressed.

Difficulties with enumerating the homeless population accurately have plagued efforts to describe homelessness effectively. Concerns have included the operational definition of homelessness, determining an appropriate sampling frame, the accuracy of the numbers resulting from any counting effort, and debate over the best counting methodology (e.g., point-in-time or annual). In addition to being a very difficult personal circumstance, homelessness also is a very fluid social problem (i.e., most homeless individuals move into and out of homelessness more or less at random as part of a lifestyle of chronic poverty and/or family abuse). This gives rise to major difficulties in recording incidents of homelessness. The number of reporting agencies changes appreciably over time, and agency reporting techniques and the record-keeping abilities of agencies and shelters are variable and subject to change. In addition, there is good reason to believe that the incidence of homelessness is underreported by some agencies, owing perhaps to inadequate record keeping, and overreported by other agencies that have reasons to prefer that larger numbers be reported. As a

result, researchers face difficult methodological issues that are not easy to resolve.

There is a substantial research literature on homelessness. The major issues include competing household-level and structural explanations of the causes of homelessness (Burt 1992a; Jencks, 1994; Shinn & Weitzman, 1990; Snow, Anderson, & Koegel, 1994; Wright, 1989), the emergence of the “new homeless” (Axelson & Dail, 1988; Hopper & Baumohl, 1996; Wright & Lam, 1987); the effects of housing policy (Dattalo, 1991; Dolbeare, 1996; Shinn & Gillespie, 1994; Stone, 1993; Wright & Lam, 1987); descriptions of homeless assistance programs (Burt et al., 1999), and public policy responses to homelessness (Department of Housing and Urban Development 1994; Dolbeare, 1992, 1996; Foscaninis, 1996; Jencks, 1994; Lazere, 1995; O’Flaherty, 1996;). Despite the wide range of approaches to studying homelessness, most research efforts in this area face substantial methodological challenges.

As discussed throughout this article, the difficulties surrounding research on homelessness are numerous and substantial. Furthermore, these difficulties are often exacerbated due to constraints imposed by limited time and available research funds. This article documents the methodological choices made by one group of researchers struggling with these difficulties and constraints. Our goal is to begin to identify precisely how methodological techniques (with particular emphasis on the operationalized definition of homelessness) affect the research findings on homelessness using the findings from a 1997 statewide study of homelessness as a starting point.

First, we examine whether the operational definition of homelessness significantly affects the demographic characteristics of the sample of the homeless population. Next, we turn to the question of whether the operational definition of homelessness significantly affects the estimated number of homeless individuals and the estimated number of incidents of homelessness. Third, using loglinear modeling techniques, we determine if there are relationships between selected demographic variables, the primary causes of homelessness, and the types of housing needed by the homeless. Fourth, we identify the extent to which the operational definition affects these relationships. Fifth, we discuss these findings in the context of their implications for specific data management and estimation techniques. Finally, the ramifications of the uncertainty surrounding homeless research on policy formation and evaluation are addressed.

ENUMERATION METHODS

Any systematic effort to count the homeless must begin by attempting to define the problem in precise, operational terms. However, a widely acceptable and uniformly interpreted definition of homelessness has yet to emerge among either researchers or homeless advocates. Generally, the most common

definition of homelessness (and the one used in this study) is the one proposed in Section 103 of the Stewart B. McKinney Homeless Assistance Act (1987), and codified as Title 42—The Public Health and Welfare, Chapter 119, Homeless Assistance, Subchapter I (General Provisions 113023/4—general definition of a homeless individual). This amendment states that, for purposes of this act, the terms *homeless* or *homeless individual* include an individual who lacks a fixed, regular, and adequate nighttime residence, and an individual who has a primary nighttime residence that is (a) a supervised publicly or privately operated shelter designed to provide temporary living accommodations (including welfare hotels, congregate shelters, and transitional housing for the mentally ill); (b) an institution that provides temporary residence for individuals intending to be institutionalized; or (c) a public or private place not designed for or ordinarily used as a regular sleeping accommodation for human beings. Excluded is any individual imprisoned or otherwise detained pursuant to an act of Congress or a state law (PL 100-77; July 22, 1987). This definition is supplemented by U.S. Department of Education (1987) guidelines regarding homeless children. Other complications arising in the process of operationally defining the incidence of homelessness include whether to include those who are doubled up (or “precariously housed”), especially as a response to poverty and/or domestic violence. Furthermore, agencies that administer homeless assistance programs sometimes broaden this definition to include individuals who are residing in transitional or supportive housing.

The McKinney Act definition has been criticized for being vague and difficult to operationalize, leading to a host of disparate studies. Martha Burt proposed a working definition of homelessness that has developed out of studies to count the homeless funded by the Department of Education and the Department of Housing and Urban Development. The categories of homeless are taken from the McKinney Act: adults, children, and youth sleeping in places not meant for human habitation and adults, children, and youth in shelters. Burt also described a category of adults, children, and youth at imminent risk of residing on the streets or in shelters. This category includes children in institutions; adults in institutions; and adults, children, and youth living doubled up. Burt recommended that researchers carefully design their studies around a working definition of homelessness, clearly defining the categories used and the methods employed to enumerate the categories. Results can vary dramatically depending on the definition used. Thus, definitions are essential to public policy and in lieu of consensus on definitions it is important for researchers to demonstrate clearly the definition used in their research (Burt, 1992b, 1995, 1996; Hopper, 1995).

An accurate, reliable count of the homeless has yet to be designed or obtained. This is due, in large part, to the inherent methodological difficulties associated with counting a population that lacks a fixed residence and, in many cases, does not want to be identified (Burt, 1992b). Furthermore, definitional issues make comparing studies difficult and often misleading (Breakey & Fischer, 1990; Cordray & Pion, 1991). Most researchers, legislators, and

advocates agree that people literally living on the street or in emergency shelters are homeless. However, there is no consensus as to the definitional status of, for example, people doubled up with family or friends for an undetermined length of time, people living in transitional housing, people in substandard housing, and runaway children (Burt, 1996).

Three major approaches have been used to produce local and regional estimates on homelessness: indirect estimates, direct estimates, and capture-recapture-based estimates. Indirect estimates (e.g., Smith & McDaid's [1987] estimate of homeless in Pennsylvania) are generally based on data collected from service providers such as soup kitchens, shelters, and government agencies. Indirect estimates are liable to be inflated by duplication. Duplication most likely can arise in two circumstances: first, when services rather than individuals are counted and, second, when individuals are counted but service providers pool data. This problem can be overcome by using systems designed to collect individual-level data for which each individual is assigned a unique identifier to control for duplication when the data are pooled. Indirect estimates of soup kitchens and other generic service providers can produce an inflated count by including a substantial number of individuals who use the service but are not homeless. It is also possible for indirect counts to produce underestimates because the count is, by definition, limited to those who use the services (Breakey & Fischer, 1990; Burt, 1992b).

Direct estimates (e.g. Rossi, Fischer, & Willis's [1987] study in Chicago) take the form of direct counts of people in shelters, at meal sites, or on the street. The strengths of this approach are that it can be highly cost-effective and is likely to include members of the homeless population who are generally missed by research that samples sleeping places. Problems with this approach include the cost of the field team to enumerate the population of a locale or an entire city, difficulty ascertaining who is homeless, and finding those homeless individuals who avoid contact (Breakey & Fischer, 1990; James, 1991).

Capture-recapture methods (e.g. Breakey, Fischer, & Cowan's [1986] study in Baltimore) make multiple counts at given locations and derive an estimate of the size of the total homeless population using dilution equations. This approach assumes that within two or more counting periods, every individual in the population has some chance of being included in the sample. The strengths of this approach include the potential for increased population coverage and the ability to evaluate the quality of the sampling work done (Cowan, 1991). There are, however, significant problems associated with applying this methodology to homeless research. For example, this approach makes two important assumptions. First, it assumes that the number of individuals entering the population is equal to the number leaving. Second, it assumes freedom of movement (i.e., there is nothing to prevent movement from any one place to any other place). These assumptions are difficult, if not impossible, to defend in the context of homelessness (Breakey & Fischer, 1990; Burt, 1992b, 1996).

There are strengths and weaknesses associated with each methodological approach of enumeration. National estimates generally are extrapolations based on data obtained in local counts. Due to differing definitions and methodologies, comparing different studies is problematic. In addition to the difficulties previously mentioned, national estimates must attempt to account for differences in population, economies, resources, and programs and other variations across the nation (Breakey & Fischer, 1990; Burt, 1996; Kondratas, 1991; Walker, 1991).

Further exacerbating the methodological difficulties already mentioned is the issue of point-in-time estimates versus annual estimates. The three estimation approaches discussed in this section (indirect, direct, and capture-recapture) are each designed to produce a point-in-time estimate (e.g., an estimate for 1 night, 1 week, or 1 month, depending on the length of time the counting was employed). However, extrapolating from data obtained using these approaches to produce an annual estimate is highly problematic due to the lack of reliable information on the rate at which people enter and leave the state of homelessness. Despite these limitations, annual estimates (period prevalence) are often calculated from cross-sectional estimates (point prevalence) for policy purposes (Breakey & Fischer, 1990; Burt, 1992b, 1996).

In light of the difficulties ascribed to a national estimate, some researchers and bureaucrats have questioned the cogency of national estimates. Christopher Walker (1991) argued that evaluations of treatment and program delivery are more useful than estimates of the size of the population. Breakey and Fischer (1990) argued that national estimates are less useful for public policy and planning than are numbers of specific subgroups of the homeless population. Eleanor Chelimsky (1991) suggested that researchers and analysts should focus their attention on relative rather than absolute measurements of homelessness.

DATA AND METHOD

The methodology used to create and manage the data reported in this article is consistent with the methodology employed in a 1997 statewide census of homelessness—hereafter referred to as “1997 statewide study” (author citation, 1998; for a complete methodological discussion see author citation, 2000). As such, discussion of these techniques will set the stage for both an analysis of the results of the 1997 statewide study as well as provide the database used for further analysis. Figure 1. provides a visual representation of the data management and estimation procedure.

The 1997 statewide study represents the first statewide assessment of homelessness in the state where the study was carried out, based on an unduplicated count of homeless individuals. The analysis includes an estimation of incidents of homelessness (both at the state level and county level) and the number of homeless individuals statewide. The study used a survey research design whereby surveys were mailed to all public schools in Iowa and all known

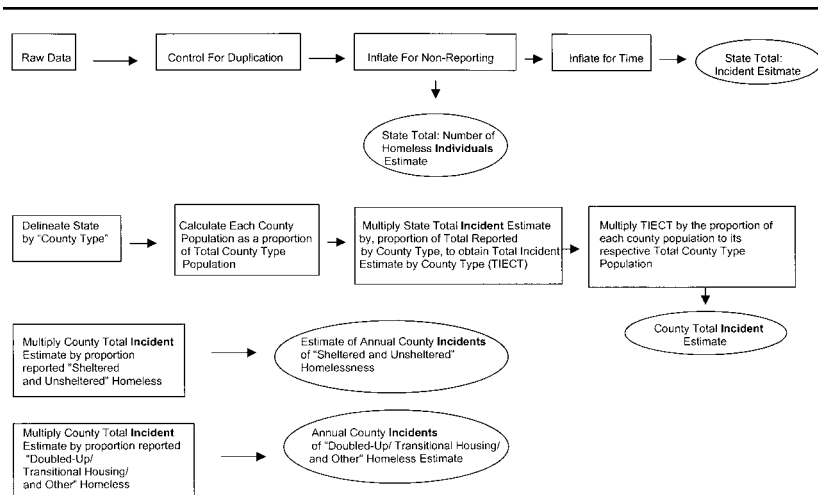


Figure 1: Flowchart Depicting the Procedures Used to Obtain Annual Statewide Estimate of Homeless Individuals and Statewide and County Estimates of Incidents of Homelessness

homeless shelters in the state, community action program agencies, county general relief offices, transitional housing programs, county department of human services offices, and miscellaneous programs such as medical outreach services serving the homeless population.

Respondents for the schools were asked to identify all homeless children known to them during the current academic year to date. Respondents for the various shelters and agencies were asked to provide information about each homeless person served during a 1-month period between March 15 and April 15 of 1997. After an extensive follow-up with nonrespondents was attempted to increase the survey response rate, all returned data were entered into an SPSS data file. The response rate for schools was 55.2%. The response rate for all agencies combined was 49.2%. The total overall response rate was 53.8%.

Elimination of reporting duplications. Following a check for data entry errors, the initial step in data analysis was elimination of the duplicate data lines, followed by calculations designed to inflate the raw data to control for the nonreporting schools and agencies. Each agency and school participating in the study was asked to provide the following information for each reported homeless person: (a) the first four letters of the individual's last name and (b) the last four digits of the individual's Social Security number (SSN). A unique identifier was created for each reported homeless person by combining these two variables and was used to locate and remove multiple data lines representing a single individual. Data lines with missing information for name and/or SSN were

coded through the use of a weighted algorithmic coding process. The process of controlling for duplication took place (a) within the data provided by schools, (b) within the data provided by agencies, and (c) between the data provided by schools and the data provided by agencies. The midrange data set was used for this study, and approximately 9% of the total reported number of cases were duplications.

DATA FROM SCHOOLS

Data from the 861 responding schools among the 1,561 schools in the school sample was entered in the file labeled *schools*. Each data line was checked against similar data lines. When a unique identifier appeared more than once, the first data line was coded as 0 (unduplicated data line) and the other(s) was (were) coded as 99 (duplicate data line). An algorithm was created to facilitate assessment of probable duplication status for the data lines that were missing one or both components of the unique identifier.

Scoring algorithm:	Name	5 points
	SSN	5 points
	Age	3 points
	Gender	1 point
	Race	1 point
	County	1 point
	District	1 point
	Building	1 point

The eight variables used in the sort were ascribed individual weights to facilitate coding of data lines with missing elements of the unique identifier. When a data line was missing either name or SSN, the available variable was checked against similar data lines to assess duplication status and was assigned a code number from 5 to 18. The code number was produced by adding the weighted values of each matching variable. When a data line was missing both name and SSN, it was coded 88 (unknown).

As an example, Case 1 (see Table 1) shows what was done in a hypothetical case where part of the unique identifier was missing. Because both data lines have missing values in the SSN column, a complete unique identifier cannot be created. Entry lines 100 and 101 are compared to one another on each of the remaining seven variables. A score of 13 is obtained (name = 5, SSN = 0, age = 3, gender = 1, race = 1, county = 1, district = 1, building = 1). The SSN is unknown; consequently, it is not considered a match and does not receive a weighted score.

In example Case 2 (see Table 2), again part of the unique identifier is missing. Again both data lines have missing values in the SSN column so the algorithm is employed. Entry lines 200 and 201 are compared, to one another, on each of the remaining seven variables. A score of 6 is obtained (name = 5, SSN = 0, age = 0, gender = 0, race = 0, county = 1, district = 0, building = 0). Using this process, all

TABLE 1: Example Case 1

<i>Entry Number</i>	<i>Name</i>	<i>Social Security Number</i>	<i>Age</i>	<i>Gender</i>	<i>Race</i>	<i>County</i>	<i>District</i>	<i>Building</i>
100	aaaa	000-00-0000	16	2	1	57	1111	109
101	aaaa	000-00-0000	16	2	1	57	1111	109

TABLE 2: Example Case 2

<i>Entry Number</i>	<i>Name</i>	<i>Social Security Number</i>	<i>Age</i>	<i>Gender</i>	<i>Race</i>	<i>County</i>	<i>District</i>	<i>Building</i>
200	bbbb	000-00-0000	11	1	2	57	2222	209
201	bbbb	000-00-0000	16	2	1	57	1111	109

data lines in the schools data set were ascribed a number representing probable duplication status within the schools data set.

DATA FROM AGENCIES

Data from 223 responding agencies among the 453 included in the sample were entered in the file called *agencies*. Each data line was checked against similar data lines. When a unique identifier appeared more than once, the first data line was coded 0 (unduplicated data line) and the other(s) was (were) coded 99 (duplicate data line). An algorithm was created to facilitate assessment of probable duplication status for the data lines that were missing one or both components of the unique identifier. This algorithm is the same as the one used for the school data except that district and building are omitted and agency (1 point) is added. The same process was used to identify probable duplication status within agencies as was employed for the schools data. However, in the agencies data, the maximum weighted score was 17 (seven variables), compared to 18 (eight variables) for the schools data.

As described above, the data sets were scanned separately for duplications (i.e., duplications within agencies and duplications within schools). Upon completion of these tasks, the two data sets were merged and scanned for duplications between the agencies and schools data sets. The above duplication removal process was repeated; however, this time the algorithm for cases with missing elements of the unique identifier included six weighted variables (name, SSN, age, gender, race, and county, for a total of 16 points).

Upon conclusion of this process, the data lines that had been coded via the scoring algorithm (i.e., those cases with missing elements of the unique

TABLE 3: Unduplicated Reported Number of Homeless in All Categories

<i>Data Source</i>	<i>Low</i>	<i>Midrange</i>	<i>High</i>
Shelters	1,435	1,481	1,672
Agencies	1,667	1,697	1,774
Schools	1,726	1,805	1,845
Total	4,828	4,983	5,291

identifier) were reintroduced to the data set. This process involved producing low-, middle-, and high-range data sets and estimates based on assumptions made regarding the probability of duplication.

The low-range data set is the most conservative unduplicated estimate. It assumes that all weighted coded items are duplicates; therefore half of all such paired entries were recoded 0 (nonduplicate) and half were recoded 99 (duplicate). All items coded 99 were then deleted. All items coded 88 (unknown) also were deleted. The high-range data set contains the least conservative unduplicated estimate. It assumes that all items coded 88 (unknown) and 5 through 18 were nonduplicates and therefore retained in the data set. Items coded 99 (duplicate) were deleted.

The midrange data set was produced and subsequently used for all estimates and analysis in this article. Items coded from 5 to 10 were assumed to be nonduplicative and then were recoded 0 (nonduplicate). Items coded from 11 to 18 were assumed to be duplicates, so half of all such pairs were recoded 0 (nonduplicate) and half were recoded 99. All items coded 99 (duplicate) were deleted. All items coded 88 (unknown) also were deleted. From the 1,881 homeless identified by the schools, 53 were found to be duplicates and were removed from the data set, leaving 1,828 unduplicated cases in the school data. From the 3,665 homeless identified by agencies and shelters, 479 duplicates/unknowns were discovered and removed from the agency and shelter data sets, leaving 3,186 unduplicated cases. When the data sets were merged, 31 additional duplicates were eliminated, leaving a total of 4,983 unduplicated cases. Approximately 9% of the total reported number were duplications (see Table 3).

Inflating for nonreporting and annualizing the data. Due to the low response rate (54% overall), it was necessary to make adjustments for nonreporting. Response rate adjustments were calculated using the response rate of the schools, the response rate of the shelters, and the response rate of the remaining agencies (general relief, department of human services, community action programs, transitional housing providers, and miscellaneous). The response rate adjustment for shelters was refined further by using shelter-bed capacity information. The data provided by agencies covered a 1-month period (March 15 to April 15, 1997), whereas the data provided by schools was for the 1996-1997 school year.

To produce an annualized estimate of incidents of homelessness, an inflation equation was applied to the agency and shelter data. The shelter bed capacity ratio is a ratio of the number of reported clients for 1 month to the number of available beds on any given night. For the middle-range data, the proportion was 1,481/1,236, producing a shelter-bed capacity rate of 1.201. The middle-range estimate assumes that one half of the nonreporting agencies had, on average, the same number of homeless as the reporting agencies during the reporting period, whereas the other one half of the nonreporting agencies had zero homeless to report.

Estimating incidents of homelessness. In addition to estimating the number of homeless individuals, an annual incident estimate was calculated. An incident of homelessness refers to one episode, of indeterminate length between 1 and 30 days, of homelessness for one individual. Each incident, by definition, is mutually exclusive of all other incidents of homelessness for the individual in question. For example, if an individual is homeless for an entire year, he or she is said to represent 12 incidents of homelessness.

To systematically arrive at an incidents estimation, the 99 counties in the state were divided into three categories (Bruner, 1993): large metro counties (largest population center is 50,000 or more), small metro counties (largest population center between 5,000 and 49,999), and rural counties (largest population center less than 5,000). There are eight large metro counties, which together contain 42% of the total state population; 45 small metro counties, representing 40% of the state population, and 46 rural counties, which account for 18% of the total state population. For each of the 99 counties, the individual county population was calculated as a proportion of the total county-type population to provide a basis for allocating the estimated incidents of homelessness across counties. For example, the total county-type population for the large metro counties is 1,183,275. A proportion of the state total population was calculated for each of the three county types. This proportion was multiplied by the state total estimate to produce a number for each of the three county types. The total number for each county type then was multiplied by the proportion of each county to its respective county-type population to obtain a total county incidents estimate for each of the 99 counties across the state.

The incident estimation procedure produced a statewide estimate as well as county-level estimates of the total number of incidents of homelessness statewide as well as an estimate of the total number of homeless individuals statewide. The statewide estimation of incidents of homeless was disaggregated to provide individual county-level estimates of incidents of homelessness. The estimate of homeless individuals that followed is an annualized estimate of individuals who experienced at least one incident of homelessness during the year. The formula used to estimate the number of homeless individuals statewide assumes that one third of those reported by the agencies are chronically homeless (i.e., 12 incidents), one third are episodically homeless (i.e., 6 incidents),

and one third were homeless only once (i.e., 1 incident) during the year of this study. These assumptions were based on national research findings on the multiple patterns of homelessness (Hopper, 1995; Link et al., 1995; Piliavin, Wright, Mare, & Westerfelt, 1996; Rossi, 1991; Sosin, Piliavin, & Westerfelt, 1990; Wright & Devine, 1995).

The current research. Using the data generated for the 1997 statewide study, this article examines some of the methodological difficulties encountered by that research. First, the operational definition of homelessness dramatically affects the findings. Because there is still no common agreement about what actually constitutes homelessness, we created an expanded definition of homelessness that refers to individuals who are categorized as currently living (a) on the streets and/or abandoned buildings, (b) in public or private shelters, (c) in transitional housing for the mentally ill, (d) in single-room-occupancy facilities, (e) doubled up with family or friends, (f) in transitional housing, (g) in youth group homes, (h) in their own home or apartment, or (i) in an other/unknown situation. We further determined that the core definition of homelessness would refer to only Categories a through d, above.

Loglinear analysis. We augmented the population and incident estimates by modeling the relationship between key demographic variables (e.g., gender, race, and age) and the cited primary cause of homelessness and the type of housing needed through the use of multinomial logit models. Before this procedure was implemented, the database (i.e., the nonduplicative data lines $n = 4,983$) was further refined through a listwise deletion process of all data lines that have missing values on any of these five variables. This resulted in an expanded definition population of 2,585 homeless individuals, with 1,300 fitting the core definition. This reduced data set was only used for the loglinear analysis; therefore, this process does not affect the estimates of homeless individuals or incidents of homelessness.

Using SPSS Version 8.1, loglinear models were produced for a series of variable combinations. Gender, age, and race were treated as independent variables, whereas the primary causal factor of homelessness and the type of housing needed were treated as dependent variables. The reference categories for the analysis were gender (female), race (biracial), age (adult), primary factor of homelessness (relocation from another area), and type of housing needed (supervised independent living). Each parameter estimate is a logarithm of an odds ratio. The actual interpretation of each estimate is somewhat convoluted because the odds ratios are calculated in terms of reference groups. For example, a parameter estimate of -1.01 for domestic violence and gender is interpreted as the difference between the log odds (i.e., the likelihood) of a male being homeless because of domestic violence (as opposed to relocation from another area) and the log odds of a female being homeless because of domestic violence (as opposed to relocation from another area). To ease interpretation,

cross-tabulation tables are provided in the body of the text. The complete log odds ratios are located in Appendixes A and B.

For our purposes, it suffices to focus on the direction of the estimate and whether the estimate is statistically significant. This simplification is appropriate for two reasons. First, the direction of the estimate (i.e., whether the value is a positive or negative integer) is relevant to the task at hand because it specifies which category of the independent variable is more likely to be homeless due to each of the categories specified in the primary factor of homelessness variable. Second, a statistically nonsignificant result indicates that there is not enough evidence to reject the null hypothesis that there is no relationship between the dependent and independent variable. Therefore, the narrative of the findings will only discuss relationships that are statistically significant.

The overall fit of a saturated multinomial logit model can be assessed through measures of dispersion and association. Shannon's entropy measure and Gini's concentration measure can be interpreted as measures of association similar to (although not identical to) an R^2 in regression. Multiplying Shannon's entropy measure by 2 produces a value that has an asymptotic chi-square distribution, which facilitates the use of a change in chi-square measure (F statistic) to determine whether the operational definition of homeless has statistically significant effects (Norušis, 1994).

The model of race (a variable with five categories) and the dependent variables (each eight categories, respectively) produces a relatively large number of estimates. The results of these models are discussed; however, to conserve space and because of the many nonsignificant parameter estimates with these two variables, we do not present these estimates in the tables.

FINDINGS

In the 1997 statewide study, low-, mid-, and high-range values were calculated for the estimate of homeless individuals and annual incidents of homeless. The figures used in subsequent calculations and reported in all charts, tables, and discussions were the midrange numbers and estimates. Based on the reported midrange raw data, 71.8% of the homeless in all definitional categories are in the large metropolitan counties, 23.6% are in the small metropolitan counties, and 4.6% are in the rural counties. The midrange estimate of homeless individuals is 26,298. The midrange estimate of incidents of homelessness is 59,558.

Table 4 shows the relative frequencies of the three demographic variables used in the analysis for both the expanded definition and core definition of homelessness. The sample population, regardless of definition used, is overwhelmingly composed of White adults. There are, however, a few changes as the result of the operationalized definition of homelessness. When the expanded definition is used, there are slightly more women in the sample, but when the core definition is used, there are more men. It is also worth noting that the ratio of

TABLE 4: Demographic Frequencies: Expanded and Core Definition of Homelessness

	<i>Expanded Definition</i>	<i>Core Definition</i>
Gender		
Male	1,270	742
Female	1,315	558
Age		
Youth (1-17)	870	312
Adult (18+)	1,715	988
Race		
White	1,901	980
Black	483	245
Asian/Pacific Islander	16	10
Native American/American Indian	38	12
Biracial	147	53

adults to youth is 2:1 with the expanded definition, whereas it rises to 3:1 when the core definition is employed.

Tables 5 and 6 present cross-tabulations of gender, age, and race with the primary causal factor of homelessness for each of the operationalized definitions of homelessness. When the expanded definition of homelessness is used, the entropy value for gender and the primary factor of homelessness is .013 and the concentration value is .0086. When only the core definition is used, the values are .036 and .023, respectively. These extremely small values indicate that regardless of definition used, there is only a small relationship between gender and the primary factor of homelessness. Despite the relative lack of relationship between the two variables, there are significant relationships within categories. For the expanded definition of homelessness, there are 3 statistically significant log odds ratios ($\alpha = .05$) for the categories of gender and the primary factor of homelessness when the expanded definition of homelessness is employed. In short, males are less likely than females to be homeless due to domestic violence, males are less likely than females to be homeless due to family disruption/conflict or breakup, and males are more likely than females to be homeless due to drug- or alcohol-related issues. When the core definition is used, the same these results are repeated, with the addition of the greater likelihood of females being homeless due to eviction than males. Finally, comparing the computed and critical F values, $F_{(7,7)} = 2.76 < F_{(\alpha=.05)} = 3.79$, indicates that the effect of the operational definition of homelessness is not statistically significant for this combination of variables.

The relationship between age and the primary causal factor of homelessness follows a similar pattern. When the expanded definition of homelessness is used, the entropy value is .016 and the concentration value is .01. When the core definition of homelessness is used, the values are .029 and .018, respectively. These values indicate that age does not do much to differentiate the primary

TABLE 5: Cross-Tabulation of Demographic Variables and the Primary Causal Factor of Homelessness: Expanded Definition (N = 2,585)

	<i>Domestic Violence</i>	<i>Eviction</i>	<i>Economic/ Employment Problems</i>	<i>Family Disruption/ Conflict/Breakup</i>	<i>Drug/ Alcohol Issues</i>	<i>Deinstitutionalized (jail or treatment)</i>	<i>Mental Health Issues</i>	<i>Relocation From Another Area</i>	<i>Total</i>
Male	120 (29.7)	188 (46.8)	340 (58.2)	168 (41.2)	175 (64.6)	33 (61.1)	72 (52.2)	174 (53.7)	1,270 (49.0)
Female	284 (70.3)	214 (53.2)	244 (41.8)	240 (41.8)	96 (35.4)	21 (38.9)	66 (47.3)	150 (46.3)	1,315 (50.9)
Youth (1-17)	187 (46.3)	189 (47.0)	136 (23.3)	165 (40.4)	44 (16.2)	7 (13.0)	31 (22.5)	111 (34.3)	870 (33.7)
Adult (18+)	217 (53.7)	213 (53.0)	448 (76.7)	243 (59.6)	227 (83.8)	47 (87.0)	107 (77.5)	213 (65.7)	1,715 (66.3)
White	286 (70.8)	301 (74.9)	433 (74.1)	308 (75.5)	215 (79.3)	35 (64.8)	109 (79.0)	214 (66.0)	1,901 (73.5)
Black	76 (18.8)	65 (16.2)	106 (18.2)	74 (18.1)	45 (16.6)	17 (31.5)	22 (15.9)	78 (24.1)	483 (18.7)
Asian	0	3 (.7)	4 (.7)	0	1 (.4)	0	1 (.7)	7 (2.2)	16 (.6)
Native American	13 (3.2)	5 (1.2)	7 (1.2)	3 (.7)	2 (.7)	1 (1.9)	3 (2.2)	4 (1.2)	38 (1.5)
Biracial	29 (7.2)	28 (7.0)	34 (5.8)	23 (5.8)	8 (3.0)	1 (1.9)	3 (2.2)	21 (6.5)	147 (5.7)

NOTE: Values are frequencies, with percentages in parentheses.

TABLE 6: Cross-Tabulation of Demographic Variables and the Primary Causal Factor of Homelessness: Core Definition (N = 1,300)

	<i>Domestic Violence</i>	<i>Eviction</i>	<i>Economic/ Employment Problems</i>	<i>Family Disruption/ Conflict/Breakup</i>	<i>Drug/ Alcohol Issues</i>	<i>Deinstitutionalized (jail or treatment)</i>	<i>Mental Health Issues</i>	<i>Relocation From Another Area</i>	<i>Total</i>
Male	49 (25.0)	104 (48.4)	224 (69.6)	82 (45.3)	111 (84.7)	22 (78.6)	44 (65.7)	160 (66.3)	742 (57.1)
Female	147 (75.0)	111 (51.6)	98 (30.4)	99 (54.7)	20 (15.3)	6 (21.4)	23 (34.3)	54 (33.8)	558 (42.9)
Youth (1-17)	75 (38.3)	96 (44.7)	46 (14.3)	50 (27.6)	5 (3.8)	2 (7.1)	4 (6.0)	34 (21.3)	312 (24.0)
Adult (18+)	121 (61.7)	119 (55.3)	276 (85.7)	131 (72.4)	126 (96.2)	26 (92.9)	63 (94.0)	126 (78.8)	988 (76.0)
White	147 (75.0)	168 (78.1)	257 (79.8)	121 (66.9)	111 (84.7)	16 (57.1)	52 (77.6)	108 (67.5)	980 (75.4)
Black	43 (21.9)	34 (15.8)	39 (12.1)	51 (28.2)	19 (14.5)	12 (42.9)	14 (20.9)	33 (20.5)	245 (18.8)
Asian	0	0	3 (.9)	0	0	0	1 (1.5)	6 (3.8)	10 (.8)
Native American	2 (1.0)	1 (.5)	4 (1.2)	2 (1.1)	1 (.8)	0	0	2 (1.3)	12 (.9)
Biracial	4 (2.0)	12 (5.6)	19 (5.9)	7 (3.9)	0	0	0	11 (6.9)	53 (4.1)

NOTE: Values are frequencies, with percentages in parentheses.

factor of homelessness. However, there are statistically significant log odds ratios. Youths are more likely than adults to be homeless due to domestic violence, eviction, economic/employment problems, drug-/alcohol-related issues, deinstitutionalization, and mental health issues. It is important to note that the data collection methods did not differentiate between individual youths and youths within families. Therefore, these results could well reflect the number of youths in families who are homeless. Take, for example, a single-parent household with three children. If the primary cause of homelessness cited by the parent is domestic violence, then this is the primary cause of homelessness for each of the three children. The results when the core definition is used are quite similar. The only change is that the estimates for economic/employment problems and deinstitutionalization are no longer statistically significant. The F -test result, $F_{(7,7)} = 1.81 < F_{(\alpha=.05)} = 3.79$, indicates that the effect of the operational definition of homelessness is not statistically significant for this combination of variables.

The relationship between race and the primary causal factor of homelessness is the weakest of the three demographic variables. When the expanded definition of homelessness is used, the entropy value is .005 and the concentration value is .003. When the core definition is used, the values increase to .013 and .008, respectively. There are only two statistically significant estimates when the expanded definition of homelessness is used. The log odds ratio for Asian Americans and domestic violence is -3.03 , and the log odds ratio for Whites and drug-/alcohol-related issues is .93. When the core definition is used, there are three statistically significant log odds ratios. Specifically, two parameters emerge as statistically significant that were not before: The log odds ratio for Whites and domestic violence is 1.24, and for Blacks and domestic violence, it is 1.20. The third significant parameter estimate is a log odds ratio of -3.17 for Asian Americans and domestic violence (which is also statistically significant with the expanded definition of homelessness), whereas the log odds ratio for Whites and drug-/alcohol-related issues becomes statistically nonsignificant. Finally, the F -test result, $F_{(7,28)} = 2.6 < F_{(\alpha=.05)} = 2.36$, leads to the conclusion that the effect of the operational definition of homelessness is statistically significant for this combination of variables.

Tables 7 and 8 follow the same format as the previous tables with the exception that now the dependent variable is the cited type of housing needed for the homeless in the sample. Let us first examine the relationship between gender and the type of housing needed. When the expanded definition of homelessness is used, the entropy value is .048 and the concentration value is .043. When the core definition of homeless is used, the values are .066 and .059, respectively. In the expanded-definition sample, there are three significant parameter estimates. In short, males are more likely than females to need permanent one-bedroom, single-resident-occupancy, or adult-group-home types of housing (as opposed to supervised independent living). When the core definition is used, these three log odds ratios remain statistically significant. The F test, $F_{(7,7)} = 1.38 < F_{(\alpha=.05)} =$

TABLE 7: Cross-Tabulation of Demographic Variables and the Cited Type of Housing Needed: Expanded Definition (N = 2,585)

	<i>Permanent One Bedroom</i>	<i>Permanent Two or More Bedroom</i>	<i>Transitional Housing</i>	<i>Juvenile Group Home</i>	<i>Single-Room Occupancy</i>	<i>Mental Health Facility</i>	<i>Adult Group Home</i>	<i>Supervised Individual Living</i>	<i>Total</i>
Male	508 (64.4)	445 (35.8)	113 (40.6)	15 (35.7)	157 (91.8)	3 (50.0)	16 (76.2)	13 (38.2)	1,270 (49.1)
Female	281 (35.6)	799 (64.2)	165 (59.4)	27 (64.3)	14 (8.2)	3 (50.0)	5 (23.8)	21 (61.8)	1,315 (50.9)
Youth (1-17)	15 (1.9)	656 (52.7)	147 (52.9)	39 (92.9)	3 (1.8)	2 (33.3)	0	8 (23.5)	870 (33.7)
Adult (18+)	774 (98.1)	588 (47.3)	131 (47.1)	3 (7.1)	168 (98.2)	4 (66.7)	21 (100)	26 (76.5)	1,715 (66.3)
White	643 (81.5)	849 (68.2)	197 (70.9)	25 (59.5)	136 (79.5)	5 (83.3)	19 (90.5)	27 (79.4)	1,901 (73.5)
Black	116 (14.7)	290 (23.3)	39 (14.0)	6 (14.3)	24 (14.0)	0	1 (4.8)	7 (20.6)	483 (18.7)
Asian	7 (.9)	6 (.5)	2 (.7)	0	0	1 (16.7)	0	0	16 (.6)
Native American	9 (1.1)	20 (1.6)	2 (.7)	3 (7.1)	3 (1.8)	0	1 (4.8)	0	38 (1.5)
Biracial	14 (1.8)	79 (6.4)	38 (13.7)	8 (19.0)	8 (4.7)	0	0	0	147 (5.7)

NOTE: Values are frequencies, with percentages in parentheses.

TABLE 8: Cross-Tabulation of Demographic Variables and the Cited Type of Housing Needed: Core Definition (N = 2,585)

	<i>Permanent One Bedroom</i>	<i>Permanent Two or More Bedroom</i>	<i>Transitional Housing</i>	<i>Juvenile Group Home</i>	<i>Single-Room Occupancy</i>	<i>Mental Health Facility</i>	<i>Adult Group Home</i>	<i>Supervised Individual Living</i>	<i>Total</i>
Male	359 (70.0)	161 (35.5)	48 (39.7)	6 (40.0)	141 (93.4)	1 (33.3)	15 (78.9)	11 (45.8)	742 (57.1)
Female	154 (30.0)	293 (64.5)	73 (60.3)	9 (60.0)	10 (6.6)	2 (66.7)	4 (21.1)	13 (54.2)	558 (42.9)
Youth (1-17)	3 (.6)	223 (49.1)	65 (53.7)	12 (80.0)	3 (2.0)	2 (66.7)	0	4 (16.7)	312 (24.0)
Adult (18+)	510 (99.4)	231 (50.9)	56 (46.3)	3 (20.0)	148 (98.0)	1 (33.3)	19 (100)	20 (83.3)	988 (76.0)
White	411 (80.1)	316 (69.6)	86 (71.1)	9 (60.0)	119 (78.8)	2 (66.7)	19 (100)	18 (75.0)	980 (75.4)
Black	83 (16.2)	109 (24.0)	25 (20.7)	0 22	(14.6)	0	0	6 (25.0)	245 (18.8)
Asian	6 (1.2)	3 (.7)	0	0	0	1 (33.3)	0	0	10 (.8)
Native American	4 (.8)	5 (1.1)	1 (.8)	0	2 (1.3)	0	0	0	12 (.9)
Biracial	9 (1.8)	21 (4.6)	9 (7.4)	6 (40.0)	8 (5.3)	0	0	0	53 (4.1)

NOTE: Values are frequencies, with percentages in parentheses.

3.79, indicates that the effect of the operational definition of homelessness is not statistically significant for this combination of variables.

The relationship between age and the type of housing need is the strongest relationship found in this analysis. When the expanded definition of homelessness is used, the entropy measure is .138 and the concentration value is .126. When the core definition of homelessness is used, the values are .141 and .136, respectively. The first five categories in Table 7 are statistically significant regardless of definition. These relatively high associations are primarily due to the types of categories contained within the type of housing variable. Juvenile group homes should, by definition, be composed of only youth. Meanwhile, single-room occupancies (SROs) are generally populated by adults. The F -test result, $F_{(7,7)} = 1.02 < F_{(\alpha=.05)} = 3.79$, indicates that the effect of the operational definition of homelessness is not statistically significant for this combination of variables.

The relationship between race and the cited type of housing needed is very weak. There are, however, a few things worth noting. First, when the expanded definition is used the entropy value is .018 and the concentration value is .015. When the core definition is used, the values are .018 and .01, respectively. These values are quite similar to what was found with the other two demographic variables. However, in contrast to the findings of gender and age, there were no statistically significant parameter estimates when the expanded definition of homelessness is used. The core definition produces two statistically significant estimates. The log odds ratio for Whites and juvenile group homes is -3.23 , and the log odds ratio for Blacks and juvenile group homes is -5.13 . Finally, the F test, $F_{(7,28)} = 1.0 < F_{(\alpha=.05)} = 2.36$, indicates that the effect of the operational definition of homelessness is not statistically significant for this combination of variables.

DISCUSSION

This study examined the effects of the operational definition of homelessness on certain variables included in a census of homelessness and found that the demographic characteristics of the sample are affected by the operationalized definition of homelessness in several ways. First, the majority of homeless individuals in the sample, based on the expanded definition, are males, whereas females are in the majority when the core definition is employed. Also, the ratio of adults to youths is 2:1 with the expanded definition and 3:1 with the core definition. In contrast, the racial composition of the sample remains almost identical. Second, overall the demographic variables are not very effective in differentiating the categories of the primary causal factor of homelessness or the type of housing needed. There are, however, significant relationships between the independent variables within the definitional categories of the dependent variables. Domestic violence and family conflict/breakup is more likely to be the primary

cause of homelessness for females than males, whereas males are more likely to be homeless due to drug-/alcohol-related issues. Homeless males are more likely than females to need individual types of housing (e.g., permanent one bedroom, SROs, and adult group homes). Third, the operational definition of homelessness does not have a statistically significant effect for five of the six models. The only exception is found for the relationship between race and type of housing needed.

The overall fit of each of the six models tested is quite weak—regardless of the operational definition. Each of the models is characterized by relatively high levels of covariance among the parameter estimates. This covariance suggests that the models would likely be improved by providing a more parsimonious set of categories in the dependent variable. These findings suggest that the operational definition of homelessness has a small effect on both the demographic composition of the sample and the associational relationships between variables. These effects, coupled with the much larger effects of the operational definition on the number of homeless individuals, warrants further attention by both producers and consumers of research on homelessness.

Lessons from the 1997 statewide study. Commonly, statewide studies of homelessness in the region of concern for this study have relied on key informant survey information (Molseed, 1995; Wright & Wright, 1993). Informants were asked to estimate the number of homeless clients served by their respective shelter, agency, or school. The reported numbers for schools and shelters were assumed to be additive, and therefore, each set of reported numbers from shelters and schools was added with all other reported sets of numbers from shelters and schools for that county to produce a county composite.

To collect data from social service agencies, Homeless Coordinating Boards were contacted. In situations where social service agencies within a county met and produced a composite figure for the county, these reported figures were not adjusted. However, when it appeared to the researchers that separate agencies within a county supplied only their own figures, these data were assumed to be duplicative, and only the highest reported number from the social service agencies for a given county were used in the study.

The 1997 statewide study employed a sophisticated methodology to control for duplication. This advancement provides greater confidence in the findings produced by the study by using an unduplicated count of the number of homeless as a solid starting point for three important aspects of the study. First, the unduplicated raw data provide the basis for the estimation of incidents of homelessness. In response to the low response rate, a series of calculations and formulas was employed to produce the estimation of incidents. The study produced low-, mid-, and high-range estimates. If only the raw data calculation within the midrange estimate is manipulated, then by assuming that there are duplicate data lines, the estimates change considerably. The midrange estimated number of incidents reported in the 1997 statewide study is 59,558. If the identical

estimation process is followed, however this time without removing the 10% of the data determined to be duplicate data lines, the estimate changes noticeably to 63,309.

Second, in the 1997 statewide study the midrange estimated number of homeless individuals is 26,298. Again, if the identical estimation process is followed, this time without removing the 10% of the data determined to be duplicate lines, the estimate of individuals also changes noticeably, to 27,870. As is apparent in both of these conservatively manipulated examples, the estimates of incidents and individuals are quite sensitive to the size of the raw data set used in the calculations. Therefore, the greater the confidence we have in the accuracy of the raw data set, the greater confidence we can have in the estimates derived from that data set.

Third, unduplicated data are an essential component of accurate presentation and analysis of demographic information. Each line in the data set represents a distinct individual; therefore, there can be a greater level of confidence in information about the gender, race, and age of the sample. When the data management techniques employed are not designed to produce a count of distinct individuals, the demographic information runs the risk of being skewed in terms of the reported characteristics of the sample.

The estimate of the number of incidents of homelessness used in the 1997 statewide study marks another change from previous studies. Although the incidence estimation does not translate directly into an estimate of the number of homeless individuals, this information may prove to be valuable in the political arena. The incident estimation provides information about areas where no homeless were reported. Without this (or a similar) estimation technique, it would be necessary to assume that each county that did not report any homeless in fact does not have any homeless living in the county.

Due to the imprecision of homeless counting measures, it is assumed that counties that do not report may have failed to do so for a variety of reasons other than there being no homeless in the county. Furthermore, the incident estimation is especially cogent for rural areas where individuals are likely to live doubled up with family or friends and therefore escape usual counting measures (Dail, 1997). In addition, the use of an incident estimation allows for an aggregate statewide estimate of incidents and a disaggregation of this statewide estimate to the individual county level. This is important, because the statewide estimated number of individuals cannot be disaggregated to the county level. To disaggregate the statewide estimate of individuals would be to place individuals in counties where no individuals were reported. If an individual were placed in a given county, then that individual, by definition, represents a discrete individual (an individual with a gender, an age, and a race). The result would be a drastic assumption that there are, for example, x number of men of y age and z race in County Q despite the reported number of zero individuals for County Q. In contrast, when the incident estimation is disaggregated, the process only assumes that there are x number of incidents of homelessness in the county during one

year. This incident estimation process does not force researchers to presume to know how many individuals are homeless and therefore is appreciably less presumptive than an attempt to disaggregate the estimate of homeless individuals.

The low response rate of the 1997 statewide study can affect the results in a variety of ways. First, with a response rate of only 53.8%, by definition there is nothing known about 46.2% of the contacted agencies, schools, and shelters. For example, the percentage of nonresponding units that had zero homeless to report for the reporting time is unknown. Nor is it known what percentage of the nonresponding units had knowledge of homeless individuals during the reporting period but failed to respond to the survey. The dearth of information about 46.2% of the units contacted was controlled for by creating low-, mid-, and high-range estimates. These estimates make various assumptions about the nonresponding agencies and how many, if any, homeless individuals they should have reported on average. With a greater response rate, the amount of estimation and the related assumptions could be markedly reduced.

Second, the low response rate may affect the composition of the sample. Homeless shelters in particular often serve specific subpopulations within the homeless population. For instance, most domestic violence shelters serve women and children exclusively, and there are shelters that serve single males only as well as shelters that serve single females only. As a result, a greater response rate by certain types of shelters or agencies may skew the demographic composition of the complete sample in favor of the subpopulation served by that type of agency or shelter.

The issue of overrepresentation is compounded in the 1997 statewide study by the combination of school data with the data collected from agencies and shelters. By definition, data collected from schools ($n = 1,805$) was for children only, whereas data from agencies ($n = 1,697$) and shelters ($n = 1,481$) was for children and adults. As a result, when these data sets are combined, children may be overrepresented.

As mentioned earlier, the 1997 statewide study developed an extensive data management process. Many assumptions were made to produce the estimation of the number of incidents of homelessness and the estimation of the number of homeless individuals despite the copious amounts of missing data. For both estimates (the number of incidents and the number of individuals), these assumptions included those about the likelihood of duplication for data points without complete unique identifiers and assumptions about the homeless client load of nonreporting agencies and schools. Furthermore, for the incident estimation it was assumed that the number of reported homeless for the cross-sectional time period was, on average, representative of the other months of the year. It was also assumed that estimates made at the level of county type (i.e., large metro, small metro, and rural) may be accurately further disaggregated to county-level estimates based solely on county populations.

Finally, for the estimate of the number of homeless individuals, it was assumed that the homeless population in the geographic region being examined

is composed of three distinct groups. To this end, it is assumed that one third of the homeless population is homeless during each of the 12 months of the year, one third is homeless during 6 months of the year, and one third is homeless during only 1 month in the year. The implication of an increasing number of assumptions is often greater uncertainty in the findings. In turn, uncertainty often fuels confusion and debate. If our goal is to produce accurate, interpretable research, then reducing the number of assumptions used in deriving estimates is desirable.

In sum, the problems associated with the low response rate, and the subsequent plethora of assumptions necessary in producing statistical estimates, are prime examples of the need for caution in interpreting results of research on homelessness. Despite the sophisticated methodology and advances made as a result of creating an unduplicated data set, studies using methodologies similar to the 1997 statewide study should be viewed as having generated new baseline data on the number of homeless and not as the definitive representation of homelessness in the state where the study was carried out.

CONCLUSION: HOMELESSNESS, UNCERTAINTY, AND PUBLIC POLICY

The process of policy development can be conceptualized as taking place in four stages. First is prepolicy development. This stage includes defining the problem, assessing the policy demands, and formulating the agenda. The second stage is policy adoption, the third stage is policy implementation, and the fourth stage is policy evaluation. The policy evaluation stage assesses the results of the policy that was created and implemented (Bonser, McGregor, & Oster 1996). The use of data collection and analysis is essential at all stages but most especially for both policy formation and policy evaluation. In the case of data on homelessness, the first stage of the policy development process is hampered due to the lack of definitional consensus. As a result, it is necessary for researchers to choose an operational definition of homelessness that is project specific so that the evaluation stage of policy development can be sufficiently met by using an identical operational definition.

In addition to the specific issue of defining the problem, accurate data are essential to the policy development process. This article has detailed many of the problems associated with homeless research in general and, specifically, homeless research in nonurban states such as Iowa represents. Three conclusions drawn from these findings directly relate to the issue of data and policy formation. First, it has been determined that simple enumeration techniques are not sufficient. The estimates are greatly affected by the operational definition of homelessness employed. The low response rate necessitates a multiphased estimation technique that includes a great deal of assumptions that may or may not be valid. This conclusion has strong ramifications for homeless policy

formation because the most straightforward and popular technique to assess the extent of the problem is to ask the question, How many homeless are there? Furthermore, comparisons between estimates made from different studies are not justified when different operational definitions are used. Therefore, unless the identical operational definition is employed, trend analysis, a very useful policy evaluation technique, is generally not an option.

Second, it has been determined that the operational definition of homelessness employed does affect (albeit weakly) the demographic composition of the sample. Because definitional consensus has not been reached in the research or political communities, it is up to individual researchers to select the operational definition used in their research. Studies that employ different operational definitions not only will arrive at different conclusions as to size of the homeless population but are also very likely to arrive at different conclusions regarding the demographic composition of the homeless population. Third, although the effects of the operational definition on associational relationship was small, further modeling of this type involving other data sets would determine whether this particular effect should be given greater attention by researchers.

It is important that future statewide research on homelessness build on the methodological advances developed for the 1997 statewide study. As detailed throughout this article, the creation of an unduplicated data set is an essential first step in producing accurate data. An unduplicated data set, coupled with a greater response rate, will bolster the accuracy of the numbers reporting the number of homeless individuals and the number of incidents of homelessness. Furthermore, an unduplicated data set also facilitates a greater number of analytical techniques that can help broaden our understanding of homelessness. In short, many of the assumptions that were made in the 1997 statewide study can be eliminated in future research by creating an unduplicated data set and an increased response rate. The task of creating an unduplicated data set can be simplified by assigning a unique identifier to each member of the sample. If this goal is met, then the complex scoring algorithms used in this study can be abandoned. Focusing attention on obtaining only the information needed to create a unique identifier may encourage the quality of information provided by individuals and service providers wary of releasing personal information, thus greatly simplifying the process of identifying duplicates.

The difficulties associated with making policy decisions about homelessness are compounded by the lack of high response rates to institutional surveys generally and more specifically by inconsistent response rates across types of institutions (shelters, schools, transitional housing programs, relief and welfare agencies, and the like) such as were experienced in this research. Papers presented at the International Conference on Survey Nonresponse, held in Portland, Oregon, from October 28 to 31, 1999, addressed many aspects of nonresponse in

institutional and personal/household surveys. Among the recommendations to increase response rates are use of the Dillman (1978, 2000) total (or tailored) design method when costs and time permit, computerized or online administration of the data collection instrument, appropriate incentives (e.g., money or shared data), better coordination among responding institutions facilitated by definitive instructions and clear motivations in cover letters, access to administrative records to fill in missing information, consistent data compilation procedures by responding institutions, careful consideration of the "micro-to-macro" factors associated with who within an organization actually supplies and records the information, a clearer understanding of how gatekeeper functions within the organization may work at cross-purposes with the goal of data reporting, consideration of the extent to which size differences among the solicited organizations affect ability to respond fully or at all, the degree to which staff are used to responding to informational requests, the optimal formats for recontacting nonrespondents, and the readability of the form and appropriateness of the information requested. Of these, to enhance response rates for future studies of homelessness, it seems essential to focus particularly on the modalities for Dillman-style follow-ups, incentive structures to make it more attractive for often overworked and understaffed schools and agencies to supply fully the information that is solicited, and how best to ensure a high degree of similarity across institutions in data compilation and reporting.

Finally, definitional consensus must be pursued at a national level. Agreement on the definition of homelessness, both within and between academic and political communities, would have many positive effects. For example, research done at local, state, and national levels could be compared and contrasted much more readily. Also, different research projects could then be woven together to create a broader picture of homelessness in the United States. However, because of the lack of definitional consensus in the research and political communities, it is particularly important for researchers to clearly define the operational definition and methods employed in their research.

Perhaps at some future time, the research problems identified in the preceding pages will be solved and perfect knowledge about the extent of the homeless problem will be obtained. However, many policy decisions will take in the interim, in the trenches of imperfect human knowledge. Policy makers respond to the immediate demands of their constituents and the demands of interest groups. Thus, policy makers, faced with the immediacy of the political process, will continue to make decisions based on the available information. With this reality in mind, researchers of homelessness need to be vigilant in developing and advancing methodological techniques that produce interpretable and accurate data on this complex social phenomenon.

APPENDIX A
Multinomial Logit Parameter Estimates: Primary Factor of Homelessness With Gender and Age

	<i>Domestic Violence</i>	<i>Eviction</i>	<i>Economic/ Employment Problems</i>	<i>Family Disruption/ Conflict/Breakup</i>	<i>Drug-/Alcohol- Related Issues</i>	<i>Deinstitutionalized</i>	<i>Mental Health Issues</i>
Gender							
Expanded definition	-1.01* (.16)	-.28 (.15)	.18 (.14)	-.50* (.15)	.45* (.17)	.30 (.30)	-.06 (.20)
Core definition	-1.76* (.23)	-.73* (.22)	.15 (.20)	-.86* (.22)	1.02* (.29)	-.57 (.48)	-.03 (.30)
Age							
Expanded definition	.50* (.16)	.53* (.16)	-.54* (.15)	.26 (.15)	-.98* (-.48)	-1.2* (0.41)	-.58* (.23)
Core definition	.82* (.24)	1.09* (.23)	-.48 (.25)	.34 (.25)	-1.83* (.47)	-1.06 (.69)	-1.35* (.52)

NOTE: Standard errors given in parentheses.

* $p < .05$.

APPENDIX B
**Multinomial Logit Parameter Estimates:
Type of Housing Needed With Gender and Age**

	<i>Permanent One Bedroom</i>	<i>Permanent Two or More Bedroom</i>	<i>Transitional Housing</i>	<i>Juvenile Group Home</i>	<i>Single-Room Occupancy</i>	<i>Mental Health Facility</i>	<i>Adult Group Home</i>
Gender							
Expanded definition	1.06* (.35)	-.12 (.35)	-.08 (.36)	-.11 (.47)	2.85* (.44)	.46 (.83)	1.56* (.60)
Core definition	1.01* (.41)	-.44 (.41)	-.26 (.44)	-.22 (.64)	2.76* (.51)	.36 (1.1)	1.39* (.66)
Age							
Expanded definition	-2.77* (.47)	1.25* (.40)	1.25* (.41)	3.56* (.68)	-2.73* (.67)	.55 (.88)	-2.62 (1.48)
Core definition	-3.46* (.74)	1.48* (.52)	1.66* (.55)	2.79* (.79)	-2.23* (.75)	2.02 (1.15)	-2.14 (1.52)

NOTE: Standard errors given in parentheses.

* $p < .05$.

REFERENCES

- Axelson, L., & Dail, P. W. (1988). The changing character of homelessness in the United States. *Family Relations, 37*, 463-469.
- Bonsor, C. F., McGragor, E. B., & Oster, C. V. (1996). *Policy choices and public action*. Englewood Cliffs, NJ: Prentice Hall.
- Breakey, W. R., & Fischer, P. J. (1990). Homelessness: The extent of the problem. *Journal of Social Issues, 46*(4) 31-47.
- Breakey, W. R., Fischer, P. J., & Cowan, C. D. (1986, November). *Homeless people in Baltimore: Demographic profile and enumeration*. Paper presented at the American Public Health Association meeting, Las Vegas.
- Bruner, C. (1993). Reinventing common sense: Indicators of well-being for Iowa's children. Des Moines, IA: Child and Family Policy Center.
- Burt, M. R. (1992a). *Over the edge: The growth of homelessness in the 1980's*. New York: Russell Sage Foundation.
- Burt, M. R. (1992b). *Practical methods for counting homeless people*. Washington, DC: Department of Housing and Urban Development, Interagency Council on the Homeless.
- Burt, M. R. (1995). Critical factors in counting the homeless: An invited commentary. *American Journal of Orthopsychiatry, 65*(3), 334-339.
- Burt, M. R. (1996). Homelessness: Definitions and causes. In J. Baumohl (Ed.), *Homelessness in America* (pp. 15-23). Phoenix, AZ: Oryx Press.
- Burt, M. R., & Aron, L. Y. (1999, December). In T. Douglas, J. Valente, E. Lee, & B. Iwen (Eds.), *Homelessness: Programs and the people they serve*. Washington, DC: Urban Institute Press.
- Chelmsky, E. (1991). Politics, policy making, data, and the homeless. *Housing Policy Debate, 2*(3), 683-697.
- Cordray, D. S., & Pion, G. M. (1991). What's behind the numbers? Definitional issues in counting the homeless. *Housing Policy Debate, 2*(3), 587-617.
- Cowan, C., (1991). Estimating census and survey undercounts through multiple service contacts. *Housing Policy Debate, 2*(3), 869-883.
- Dail, P. W. (1997). *Final report, Year II counting homeless Iowans project—Rural homelessness*. Submitted to the Iowa Department of Economic Development, Des Moines.
- Dattlalo, P. (1991). Moving beyond emergency shelter: Who should fund low-income housing? *Social Work, 36*(4), 297-301.
- Department of Housing and Urban Development. (1994). *Priority: Home! The federal plan to break the cycle of homelessness*. Washington, DC: Author.
- Dillman, D. (1978). *Mail and telephone surveys: The total design method*. New York: Wiley.
- Dillman, D. (2000). *Mail and Internet surveys: The tailored design method*. New York: Wiley.
- Dolbear, C. N. (1992). Federal homeless social policies for the 1990's. *Housing Policy Debate, 2*(3), 1057-1094.
- Dolbear, C. N. (1996). Housing policy: A general conclusion. In J. Baumohl (Ed.), *Homelessness in America* (pp. 34-45). Phoenix, AZ: Oryx Press.
- Foscarinis, M. (1996). The federal response: The Stewart B. McKinney Homeless Assistance Act. In J. Baumohl (Ed.), *Homelessness in America* (pp. 160-171). Phoenix, AZ: Oryx Press.
- Hopper, K. (1995). Definitional quandaries and other hazards in counting the homeless: An invited commentary. *American Journal of Orthopsychiatry, 65*(3), 340-345.
- Hopper, K., & Baumohl, J. (1996). Redefining the cursed word: A historical interpretation of American homelessness. In J. Baumohl (Ed.), *Homelessness in America* (pp. 1-14). Phoenix, AZ: Oryx Press.
- James, F. (1991). Counting homeless persons with surveys of users of services for the homeless. *Housing Policy Debate, 2*(3), 733-752.
- Jencks, C. (1994). *The homeless*. Cambridge, MA: Harvard University Press.

- Kondratas, A. (1991). Estimates and public policy: The politics of numbers. *Housing Policy Debate*, 2(3), 631-647.
- Lazere, E. B. (1995). *In short supply: The growing affordable housing gap*. Washington, DC: The Center on Budget and Policy Priorities.
- Link, B., Phelan, J., Bresnahan, M., Stueve, A., Moore, R., & Susser, E. (1995). Lifetime and five year prevalence of homelessness in the United States: New evidence on an old debate. *American Journal of Orthopsychiatry*, 65(3), 347-354.
- Molseed, M. J. (1995) *Homeless children and families in Iowa: The 1994 summary*. Des Moines: Iowa Department of Education.
- Norusis, M. J. (1994) *SPSS Advanced Statistics 6.1*. Chicago: SPSS Inc.
- O'Flaherty, B. (1996). *Making room: The economics of homelessness*. Cambridge, MA: Harvard University Press.
- Piliavin, I., Wright, B.R.E., Mare, R. D., & Westerfelt, A. H. (1996). Exits from and returns to homelessness. *Social Service Review*.
- Rossi, P. (1991). Strategies for homeless research in the 1990's. *Housing Policy Debate*, 2(3), 1029-1055.
- Rossi, P. H., Fischer, G. A., & Willis, G. (1987). The urban homeless: Estimating composition and size. *Science*, 235, 1336-1341.
- Shinn, M., & Gillespie, C. (1994). The roles of housing and poverty in the origins of homelessness. *American Behavioral Scientist*, 37, 505-521.
- Shinn, M., & Weitzman, B. (1990). Research on homelessness: an introduction. *Journal of Social Issues*, 46(4), 1-11.
- Smith, T. K., & McDaid, E. W. (1987). *Homelessness in Pennsylvania: Numbers, needs, and services*. A report to the Pennsylvania Department of Public Welfare. Philadelphia: The Conservation Company.
- Snow, D., Anderson, L., & Koegel, P. (1994) Distorting tendencies in research on homelessness. *American Behavioral Scientist*, 37, 461-475.
- Sosin, M., Piliavin, I., & Westerfelt, H. (1990). Toward a longitudinal analysis of homelessness. *Journal of Social Issues*, 46(4), 157-175.
- Stewart B. McKinney Homeless Assistance Act of 1987. Pub. L. number 100-77, 101 Stat. (July 22, 1987).
- Stone, M. E. (1993). *Shelter poverty: New ideas on housing affordability*. Philadelphia: Temple University Press.
- U.S. Department of Education. (1987). *Memo concerning guidelines for defining and counting homeless children*. Washington, DC: Author.
- Walker, C. (1991). Federal homeless information needs and local practice. *Housing Policy Debate*, 2(3), 617-627.
- Wright, J. D. (1989). *Address unknown: The homeless in America*. New York: Aldine de Gruyter.
- Wright, J. D., & Devine, J. A. (1995). Housing dynamics of the homeless: Implications for a count. *American Journal of Orthopsychiatry*, 65(3), 320-329.
- Wright, J. D., & Lam, J. A. (1987). Homelessness and the low-income housing supply. *Social Policy*, 17, 48-53.
- Wright, R. D., & Wright, S. E. (1993). *Homeless children and adults in Iowa: Addressing issues and options in education, services and community*. Des Moines: Iowa Department of Education.