Subsidized Housing, Emergency Shelters, and Homelessness: An Empirical Investigation Using Data from the 1990 Census

Dirk W. Early
Department of Economics and Business
Southwestern University
Georgetown, TX 78627
Phone: (512) 863-1592
Fax:(512) 863-1535
earlyd@southwestern.edu

Edgar O. Olsen
Department of Economics
University of Virginia
Charlottesville, VA 22901
Phone: (804) 924-3443
Fax: (804) 924-7659
eoo@virginia.edu

July 2001

JEL Classification: I3
Keywords: Homelessness

The authors would like to thank Brendan O'Flaherty, Steve Raphael, Horst Tomann, and participants in the Public Economics Workshop at the University of Virginia and sessions at the Nineteenth Annual APPAM Research Conference, the Seventieth Annual Conference of the Southern Economic Association, and the Tenth Annual AREUEA International Real Estate Conference for helpful comments on earlier drafts of this paper. Olsen gratefully acknowledges financial support from the Bankard Fund for Political Economy at the University of Virginia. We also thank Kathryn Kelly for her research assistance.

Subsidized Housing, Emergency Shelters, and Homelessness: An Empirical Investigation Using Data from the 1990 Census

Abstract

This paper uses data on the only systematic count of the homeless throughout the United States to estimate the effect on the rate of homelessness of a wide variety of potentially important determinants, including several major policy responses to homelessness that have not been included in previous studies. It improves upon estimates of the effect of previously studied determinants by using measures that correspond more closely to underlying theoretical constructs, especially by accounting for geographical price differences. It also conducts numerous sensitivity analyses and analyzes the consequences of the undercount of the homeless for point estimates and hypothesis tests. The paper's most important finding from a policy perspective is that targeting the current budget authority for housing assistance on the poorest eligible households will essentially eliminate homelessness among those who apply for assistance. Achieving this goal without concentrating the poorest households in housing projects and without spending more money requires vouchering out project-based assistance. The primary methodological finding of the paper is that the 1990 Decennial Census did not produce sufficiently accurate counts, especially of the street homeless, to permit very precise estimates of the effects of many factors which surely affect the rate of homelessness. The main exceptions are the price of housing and average March temperature. Plausible models of the undercount imply that in regressions with a rate of homelessness as the dependent variable estimators of the coefficients of explanatory variables will be biased towards zero. In regressions with the logarithm of a rate of homelessness as the dependent variable, only the estimator of the constant term will be biased downwards. The unknown magnitude of the undercount precludes predicting the effects of policy interventions on the number of homeless based on the results in this paper and previous studies.

1. Introduction

Many studies have attempted to explain differences in rates of homelessness across metropolitan areas in the United States. This paper contributes to this literature in several ways. First, we include variables to represent the influence of several major policy responses to homelessness that have not been included in previous studies, and we include better proxies for others. Specifically, we use data on the only systematic count of the homeless throughout the United States to estimate the effect on the rate of homelessness of the number of shelter beds, the quality of homeless shelters, the number of households subsidized by federal housing programs, and the targeting of housing assistance on the poorest of the poor. Second, we estimate more precisely the effects of other determinants of homelessness that have been included in previous studies by using measures of these determinants that correspond more closely to the underlying theoretical constructs, especially by accounting for geographical price differences. Finally, we give considerable attention to the consequences of the undercount of the homeless for point estimates and hypothesis tests.

2. Literature

More than a dozen studies have attempted to determine the factors that explain the incidence of homelessness and to estimate the importance of these factors. The main impediment to estimation and hypothesis testing has been the absence of reasonably accurate estimates of the rate of homelessness in particular areas. One consequence of the large errors in measuring the rate of homelessness is that the coefficients of many variables with strong theoretical reasons for inclusion in the regression equations are statistically insignificant at the usual levels of significance. The deficiencies of all existing measures of rates of homelessness are so widely

recognized that two recent studies (Troutman, Jackson, and Ekelund, 1999; Quigley, Raphael, and Smolensky, 2001) estimate statistical relationships based on a number of different measures and search for qualitative results that apply across all regressions.

The overwhelming majority of studies have attempted to explain differences in some measure of the rate of homelessness across different localities. Many of the earlier studies (Appelbaum, Dolny, Dreier,and Gilderbloom, 1992; Bohanon, 1991; Elliot and Krivo, 1991; Honig and Filer, 1993; Quigley, 1990; Tucker, 1987, 1990) relied on HUD's highly suspect 1984 estimates of the rates of homelessness in 60 cities (HUD, 1984).[1] This measure has also been used as one of two measures in a recent study (Troutman, Jackson, and Ekelund, 1999). These estimates are based on the opinions of local officials involved with the problem of homelessness. It is safe to say that few, if any, had a good basis for their estimates, and some of these officials might have been tempted to overstate the problem in their localities to obtain additional funds for coping with it. A carefully designed study of the number of homeless in Chicago a year later estimated about a ninth as many homeless as HUD's best estimate based on the opinions of local experts (Rossi, 1989, Chapter 3).

Similar remarks apply with equal force to one of the four measures of the rate of homelessness used by Quigley, Raphael, and Smolensky (hereafter QRS). To justify its request for grants to assist the homeless under HUD's Super Notice of Funding Availability, each local government must provide an estimate of the number of homeless persons in its area. QRS use this estimate for 49 counties in California divided by the population of the county as a measure of the rate of homelessness. However, they point out that these estimates are not produced by a

[1] See Rossi (1989, pp. 52-54) for a description and critical evaluation of this attempt to estimate the number of homeless persons. Rossi (1989, Chapter 3) and Taeuber (1991) provide descriptions and assessments of the major attempts to count the homeless.

consistent method and imply fancifully that there were more that 360,000 homeless persons in California at one point in time in 1996.

Recognizing the unreliability and likely bias in the stated opinions of those involved with the homeless and the severe difficulties in counting the street homeless, Burt (1992) and QRS used the ratio of shelter beds to population in more than 100 metropolitan areas as an index of the rate of homelessness in these areas. If the number of shelter beds per homeless person were the same everywhere, linear regressions explaining the rate of homelessness based on this index would provide the same results of hypothesis tests as regressions based on the true rate of homelessness, and the ratio of the estimated coefficients of any two variables would be the same. If the dependent variable in the regression is the logarithm of the rate of homelessness, only the estimated constant term is affected by using the index rather than the true rate of homelessness. Unfortunately, data on the number of homeless from the 1990 Decennial Census and the number of shelter beds from Burt suggests enormous variation in the number of shelter beds per homeless person across the metropolitan areas used in the aforementioned studies. Using the sample of metropolitan areas where the number of shelter beds was observed, this ratio ranged from .23 to 7.7 with a mean of 2.29 and a standard deviation of 1.35.[2] Using shelter beds as a proxy for the number of homeless has another serious drawback. It precludes estimating the effect of additional shelter beds on the size of the homeless population and the number of homeless on the streets.

The majority of the remaining studies have used data from the 1990 Decennial Census. During the night of March 20-21, the Census Bureau conducted their Shelter and Street Night

[2] These numbers are almost surely too high everywhere due to the substantial Census undercount of the homeless, and an unknown part of the variation in the estimated number of shelter beds per homeless person may be due to variation in the extent of the undercount.

operation.[3] This was an effort to enumerate two components of the homeless population, namely, persons sleeping overnight in a wide range of facilities used to temporarily house the homeless and persons visible at street locations thought to be frequented by homeless individuals. Known as the "S-Night" count, this operation found 178,638 persons in temporary housing and 49,734 persons living on the streets (U.S. Department of Commerce, 1991). Although this is easily the best source of aggregate cross-sectional data available to study the determinants of homelessness, it has serious shortcomings for this purpose.

The enormous effort devoted to compiling a list of all major types of facilities used to temporarily house the homeless probably kept the errors in counting the sheltered homeless to relatively low levels.[4] Assessments of the enumeration in selected locations suggest that the count of the sheltered homeless in the types of facilities surveyed was very accurate. The primary problems with the shelter count were that some sheltered homeless slept in other types of facilities and the fraction of the sheltered homeless in these other facilities varied from place to place.

The count of the street homeless was much more problematic. Between 2 A.M. and 4 A.M. on the morning of March 21, census takers attempted to count people at street locations designated by local officials as known congregating areas for the homeless. They counted everyone they saw out in the open at the assigned locations except persons in uniform or those

[3] See Taeuber and Siegel (1991) for a detailed description.

[4] See Taeuber and Siegel (1991, p. 98) for a description of how this list was constructed. It was surely the most comprehensive list of places providing beds for persons who have no usual residence elsewhere that had been compiled up to that time. It included emergency shelters for homeless persons, hotels and motels used entirely to house homeless persons, all hotels and motels charging $12 or less per night regardless of length of stay and whether people regarded themselves as homeless, temporary overnight shelters for runaway and homeless youth, churches that house homeless persons overnight, sites ordinarily classified as noninstitutional group quarters which the city said housed mostly homeless persons (this included YMCA's in some cities), and any site set up as a temporary shelter, even if open only on the night of the count of the sheltered homeless. Although the Bureau of the Census did a thorough job of enumerating the most important facilities used to house the homeless, it is clear that some homeless persons stayed overnight at other types of facilities such as hotels that were used occasionally, but not regularly, to house homeless people for a night or two (U.S. General Accounting Office, 1991, p. 4).

engaged in obvious money-making activities other than panhandling. Between 4:00 A.M. and 8:00 A.M., census takers attempted to count all individuals exiting predesignated abandoned buildings. In neither case were the people counted asked whether they were homeless. [5] Although this count was conducted at about 24,000 street sites suggested by knowledgeable local officials as places where homeless people frequently spent the night, it is reasonable to believe that a substantial fraction of the street homeless were not counted because they were not seen by the census observers at these locations or they slept at other locations that night.

In an attempt to identify the shortcomings of their S-Night count, the Census Bureau funded independent assessments of the enumeration. These reviews provide evidence that the street homeless were much less likely than the sheltered homeless to be counted. In Phoenix, for example, interviews with twelve of the sheltered homeless found all had been counted, but only two of ten street homeless believed they were enumerated. In Chicago, only five of 18 homeless persons interviewed believed they were included in the S-Night count. Of those five, four were in shelters. In three cities, Los Angeles, New York, and Philadelphia, observation teams were sent to sites that were to be visited by a Census enumerator. In those cities, no enumerator was observed in more than half of the street locations identified by the Census as potential gathering places for the non-shelter-using homeless (U.S. General Accounting Office, 1991, pp. 9 -11). Unfortunately, these assessments do not provide credible estimates of the extent of the undercount at a random sample of locations that could be used to adjust the Census estimates of the rate of homelessness in each locality. The consequences of the undercount and its variability for the regression results will be discussed later.

[5] Although some people with regular housing were undoubtedly included in the count, errors of omission were surely much greater.

Three recent studies (Grimes and Chressanthis, 1997; Troutman, Jackson, and Ekelund, 1999; QRS, 2001) use as the dependent variable in a regression model the rate of homelessness (or the logarithm of this rate) based on the 1990 Census count in each metropolitan area for which the values of their explanatory variables could be easily obtained. Two other recent studies (Early, 1998; Early and Olsen, 1998) use these counts in the process of choice-based estimation of a logit model based on household data, and Early (1999) uses them in the process of choice-based estimation of a structural model of homelessness. Specifically, they combine data on individual households from a survey of the homeless conducted by the Urban Institute in 1987 and surveys of the housed conducted by the Bureau of the Census in 1985 through 1988.

Choice-based estimation requires knowledge of the probability of selection of each observation. The major weakness of the three studies based on household data is that the probability of selection is not known with certainty for the sample of the homeless and the probabilities used in the estimation almost surely differ greatly from the truth in many cases. Strictly speaking, the Urban Institute sample is not a random sample of any type from the homeless population. It is a stratified clustered random sample of persons who used shelters or soup kitchens during a week.[6] In estimating the models with household data, the three studies assumed that it was a simple random sample of the homeless in each locality with probability of selection for each element of the sample in that locality equal to the ratio of the number of homeless in the sample to the number of homeless counted in the 1990 Decennial Census. Since the Census homeless counts were surely too low in all areas, the probabilities used were too high.

[6] For all practical purposes, it is impossible to obtain a random sample of the homeless at any point in time because this requires knowledge of the location of all homeless persons at that time. See Burt and Taeuber (1991) for a description of the method used to obtain the Urban Institute sample of homeless persons. This method is based on the plausible belief that many homeless persons who do not sleep in a shelter on a given night obtain a meal from a shelter or soup kitchen during the day. Interviews with a nonrandom sample of 445 homeless persons revealed that 68 percent had used a shelter or soup kitchen during the previous week.

The aforementioned studies based on household data have some advantages; those based on aggregate data have others. Some factors that affect homelessness such as income and gender are specific to the household; others such as the temperature and the price per unit of housing service are characteristics of the localities in which households live. The three studies based on household data provide more precise estimates of the effect of variables specific to the household because there is considerably more variability in the values of these variables across households than in their means across metropolitan areas. The studies based on aggregate data for urban areas provide more precise estimates of the effects of characteristics of the area because they are based on a much larger sample of these areas. The household data on the homeless and housed are available for only 15 urban areas. The studies based on aggregate data from the 1990 Decennial Census use information for at least 200 urban areas.

A few other important studies do not fit neatly into the preceding categories. Cragg and O'Flaherty (1999) use time-series data for New York City between 1986 and 1993 to study why families enter and leave homeless shelters. They are especially interested in the effect of giving households in shelters priority for receipt of long-term housing assistance, though they consider many other potentially important causes of changes in the number of households entering and leaving shelters. Culhane and Metraux (1999), Culhane and Kuhn (1998), and Culhane, Lee, and Wachter (1996) study the determinants of shelter usage and exit rates using data on individuals and families who slept in public shelters in New York and Philadelphia. Allgood, Moore, and Warren (1997) analyze the determinants of the duration of sheltered homelessness.

In one series of regressions, QRS use data on the number of households served by California's Homeless Assistance Program in each of California's 58 counties for each year between 1989 and 1996 as an index of the number of categorically eligible households who are

homeless sometime during the year. This program provides payments to AFDC-eligible (or apparently eligible) families needing shelter due to homelessness. Under the program's rules, homelessness exists when a family lacks a fixed and regular nighttime residence or is living either in a supervised temporary shelter or in a public or private place not designed for, or ordinarily used as, a regular sleeping accommodation by human beings. In practice, a family with children literally has to be on the street with nowhere to go before qualifying for aid. This is the only study to focus on the number of households who are homeless sometime during a year rather than the number homeless at a point in time. The disadvantage of this data is that it deals with homelessness among a subset of the population in one state. Its advantages are that it is not based on opinions and it is panel data, which allows the authors to control for unobserved differences in factors explaining homelessness between counties that are reasonably constant over time and unobserved determinants of homelessness that are about the same in all counties but differ over time.

3. Model, Data, and Estimation Methods

The rate of homeless in a metropolitan area is assumed to depend upon variables reflecting the budget constraints of households in the absence of government assistance, others reflecting policy interventions that affect budget spaces of poor households and vary across metropolitan areas, and still others that reflect differences in tastes. Some explanatory variables that reflect the budget constraints of households in the absence of government assistance are the means of variables that differ across households in an area (poverty status, unemployment); others are characteristics of localities (price of housing service, rental vacancy rate). Some of the policy interventions are focused on the homeless (shelter availability, shelter quality, McKinney Act

funding). Others are housing policies that might reduce homelessness (availability of subsidized housing, targeting of subsidized housing on the very poor), non-housing policies that help those at the greatest risk of homelessness (mental health spending), or proxies for noncash assistance (general assistance, AFDC).[7] Some of the remaining variables are the means of characteristics of individuals (race and gender) and others are characteristics of localities (temperature, precipitation, crime rate, and regional dummies). They capture differences in tastes as well as differences in circumstances that are not captured by other explanatory variables.

Like other recent studies, we estimate separate regressions for the sheltered, street, and total rate of homelessness because an explanatory variable may have qualitatively different effects on these rates and these differences often matter for policy evaluation. Some government policies that increase the overall rate of homelessness decrease the number of street homeless. They draw people from the streets and the worst regular housing situations into shelters. These policies might be judged successful despite their effect on the overall rate of homelessness. Other policies that increase the overall rate of homelessness increase the number of sheltered and street homeless. To distinguish between these policies requires separate regressions for the sheltered, street, and total homeless.

We also estimate a relationship explaining the fraction of the homeless in shelters. Since it might reasonably be argued that the decision to be homeless is a decision with a somewhat longer time horizon than the decision of a homeless person to be in a shelter on a particular night, we estimate the relationship explaining the fraction of homeless in shelters on the night of the homeless count with a short as well as a full set of explanatory variables. For example, even

[7] These should be considered proxies for noncash assistance because cash assistance is included in cash income, which is already incorporated in the analysis through the explanatory variable measuring the percent of households with incomes less than half the poverty line. It seems plausible that localities that provide generous real cash assistance to the poor will provide more noncash assistance of all types.

though the price per unit of housing services affects the fraction of all poor households that are homeless, it may have little effect on the fraction of the homeless who are in shelters on a particular night.[8] We also use slightly different definitions of some variables in these regressions. For example, in explaining the fraction of poor households who are homeless, our measure of shelter availability is the number of shelter beds per poor person. In explaining the fraction of the homeless who were in shelters on the night of the homeless count, our measure is the number of shelter beds per homeless person.

We explore the sensitivity of our results to a number of alternative specifications. First, although almost all of our results account for the likely endogeneity of shelter availability in explaining rates of homelessness, we also estimate relationships for total, sheltered and street homelessness treating this variable as exogenous. Second, we attempt to account for the effects of the likely undercount of the street homeless in the 1990 Decennial Census by estimating relationships explaining total homelessness based on the assumptions that a fifth and a tenth of the street homeless were counted as well as on the actual count. Third, we estimate several different functional forms of the relationship explaining the overall rate of homelessness. Finally, we account for one type of heteroskedasticity in the error term.

Table 1 defines each variable used in the regressions and provides the sources of all data. Table 2 contains our expectations about the qualitative effect of each explanatory variable on each dependent variable. Most of these expected effects require no explanation. However, a few words about our dependent variable, the policy variables of particular interest for this study, and the expected effects of a few other variables are in order.

[8] This is inconsistent with a discrete choice model involving the choice between being housed, being in a shelter, and being on the street. Even though the utility associated with the latter two alternatives does not depend on the price per unit of housing service, the probability that a person selected at random from the population will be in a shelter, conditional on being homeless, does depend on this price.

The source of data on the number of homeless across metropolitan areas is the 1990 Decennial Census. These homeless counts are reported for 284 metropolitan areas, which is the unit of observation used in this study.

In calculating the rate of homelessness all previous studies divide the number of homeless by the total population in the urban area. However, only very low-income households are homeless. The 1996 National Survey of Homeless Assistance Providers and Clients shows average income of homeless families (whose average size is three) is only $475 per month or 46 percent of the federal poverty level for families of this size (Interagency Council on the Homeless, 1999). Our calculations based on the household data from this survey indicate that less than 4 percent of the homeless had incomes above the official poverty line. We could not adjust for geographic price differences in this calculation because the location of the household is not revealed. Had it been possible to use a poverty line adjusted for geographical price differences in making this calculation, this percentage surely would have been smaller, perhaps zero. Those with the highest nominal incomes typically live in places with the highest price levels.

A more appropriate measure of the rate of homelessness is the number of homeless divided by the poverty population. Unlike the usual estimates of the poverty population, ours is based on a poverty line adjusted for variations in prices across urban areas. Our geographical consumer price index is based on Malpezzi, Chun, and Green's (1998) housing price indices, the assumption that the prices of other goods are the same everywhere, and BLS weights used to construct the national CPI (U.S. Bureau of Labor Statistics, 1999). In constructing this price index, we first rescale Malpezzi, Chun, and Green's rental and owner-occupied price indices so that the weighted average of each index across metropolitan areas is one, where the weight for a

particular metropolitan is the fraction of all renters or owner-occupants who live in that area. The price index for other goods is set equal to one in all areas.

The Census reports the number of persons in each metropolitan area in various ranges of incomes relative to the poverty line, for example, under 0.50 and 0.50 to 0.74. If a metropolitan area has an estimated price index above 1.00, our estimate of the number of poor persons includes all officially classified as poor plus some number of persons from the groups above the official poverty line. For example, if a metropolitan area has a price index of 1.1 it is assumed that the fraction (0.10/0.24) of the persons in the 1.00 to 1.24 range were poor. If the price index is below 1.0, the number of poor is less than the official count. If the price index were 0.8, all persons in the under 0.50, and 0.50 to 0.74 ranges would be counted as poor, but only (0.05/0.24) of the persons in the 0.75 to 0.99 range were included in our count of the poor in that metropolitan area. This proxy for the number of poor adjusting for variations in prices across urban areas assumes a uniform distribution of households within each range of income relative to the poverty line.

The difference between our specification of the dependent variable and those in other studies amounts to a difference in the assumed functional form of the stochastic relationship explaining the rate of homelessness as usually measured. Obviously, a specification with the usual measure of homelessness can be obtained from our model by multiplying both sides of our specification by the poverty rate adjusted for geographical price differences.

Some of the explanatory variables needed in the estimations were not available for all areas. We deleted some metropolitan areas on account of missing information about a key variable, and we handled the missing values of other variables as described below.[9]

Since a crucial step in the analysis is to control for variations in prices across areas, we restrict the sample to areas for which price indices are available. We also omitted Anchorage and Honolulu since these areas are so dissimilar from the other metropolitan areas in our sample. This reduces the sample to 224 metropolitan areas.

The source of the number of shelter beds, one of our important policy variables, used a somewhat different definition of a shelter bed than did the Census in attempting to count the sheltered homeless, collected data at a different time than the Census, measured shelter beds across cities instead of metropolitan areas, and did not collect data for a large fraction of these areas. Specifically, Burt used phone surveys to obtain shelter bed capacity in late 1989 in cities that are part of 109 of the 224 metropolitan areas in our sample.

Both Burt and the Census included in their definition publicly and privately run homeless shelters and hotels and motels used entirely to house homeless persons. Unlike the Census, Burt included in her count beds occupied by homeless persons in hotels that were used occasionally, but not regularly, to house homeless people (Burt, 1992, pp. 130-135). Unlike Burt, the Census included shelters for battered women and all beds in hotels and motels costing $12 or less per night regardless of whether the occupants considered themselves homeless (U.S. Census Bureau, 1999, p. D-3). So Burt does not have an exhaustive count of shelter beds and the Census undoubtedly undercounted the sheltered homeless to some extent.

[9] In earlier versions of this paper, we limited our sample to the 100 metropolitan areas for which values of all explanatory variables are available. The absence of data on the number of shelter beds accounted for the bulk of the deletions of observations. We expected the addition of 124 observations combined with our treatment of missing values of explanatory variables to increase the number of statistically significant coefficients with the expected signs. The gains in this regard were modest.

Although Burt conducted her survey in late 1989 and the Census on March 20, 1990, this is not a significant problem for the analysis. The number of shelter beds was surely about the same at the two times because the Census conducted its survey before some shelters close down at the end of the period of peak usage.

We transformed Burt's city level measures of shelter beds to metropolitan area measures by assuming that the number of shelter beds per capita was the same in the entire metropolitan area as in Burt's cities in the area. Since central cities probably have more beds per capita than do suburbs, this will probably overestimate the availability of shelter beds.

Instead of dropping observations lacking shelter bed counts, the fitted value from a regression explaining the availability of shelter beds using the 109 metropolitan areas for which data are available is used for areas where the number of shelter beds is unknown. Using the fitted values for availability of shelter beds adds measurement error to that variable. Since this has the potential of biasing the coefficient on shelter availability, we estimated all regressions restricting the sample to observations for which shelter availability is observed. The results using the smaller sample size are broadly similar to the results using all 224 observations. Since the larger data set allows for more precise estimates of other parameters, only these results are reported.

It is reasonably argued that the number of shelter beds in a locality is endogenous in a regression explaining the rate of homelessness across localities. That is, localities with a high rate of homelessness for reasons not captured by the observed explanatory variables will have more shelter beds. To account for this possibility, instrumental variables estimation is employed throughout. (Table 4 reports results treating shelter availability as exogenous to determine the effect of accounting for the potential endogeneity of shelter availability.) An instrument is

constructed by regressing the availability of shelters on the other explanatory variables and additional variables that might explain shelter availability. Descriptions of the additional variables included in the first stage regressions are included at the end of Table 1. These include a measure of the political climate in each city, measures of the cost of building and maintaining homeless shelters, the median income of the metropolitan area and the number of governments in total and per capita. The last variable is included on the plausible argument that the tax base upon which poor areas can draw is effectively smaller in a municipally fragmented metropolitan area, than the tax base of poor areas located in less fragmented cities.[10] The results of the first stage regressions are reported in Appendix A. The fitted values of shelter availability are then used as an instrument in the regression explaining the rate of homelessness.

A few other variables – general assistance, number of subsidized units, subsidized units targeted toward the poor, crime rate, and price of land – had missing values, but for a much smaller fraction of the sample. Two regressors represent the influence of each of these variables. One is the observed value if available and zero otherwise. The other is a dummy indicating whether the value is missing. The dummy variable takes the value of one if missing and zero otherwise. Unless otherwise noted, the descriptive statistics reported in Table 1 describe the observed values of these variables.

All previous studies have used nominal magnitudes such as median income or cash assistance to the poor to explain the incidence of homelessness when the conceptually correct measure requires an adjustment for geographical price differences. We refine these measures by using the geographical price index described above.

The unemployment rate is included as an explanatory variable in addition to a measure of the level of real income of the poor to allow for the possibility that a poor person with a given

[10] We thank Steve Raphael for suggesting this variable.

total income over a year is more likely to become homeless if he or she has a spell of unemployment during that period.

Many past studies have used median rent in a locality, or rent at some other percentile of the rent distribution, as an index of the price of rental housing. This amounts to explaining the fraction of households in a locality that spend nothing on a good by expenditure on that good at some percentile of the distribution of expenditure. Obviously, these two variables have many common determinants. Differences across metropolitan areas in rent at some percentile of the distribution of rents reflect differences in the quantity of housing service at that percentile as well as differences in the price per unit of housing service. To overcome this shortcoming of previous research, we use the rental housing price index developed by Malpezzi, Chun, and Green (1998).

Although our measure of the price per unit of housing service is surely a more satisfactory measure of the appropriate underlying theoretical construct, its inclusion is not unobjectionable. Some unobserved factors that affect the rate of homelessness affect the entire distribution of quantities demanded and hence the overall demand for housing. If the supply curve is upward sloping, markets with a greater demand for housing will have a higher price per unit of housing service. In other words, the error term in the equation explaining the rate of homelessness is surely correlated with our price index. Like all previous studies of homelessness, we ignore this potential endogeneity of the housing price index.

It has been argued that an increase in the number of federally subsidized housing units will reduce the number of homeless (Interagency Council on the Homeless, 1994, Burt, 1992, Appelbaum, Dolny, Dreier,and Gilderbloom, 1991). Indeed, HUD (1997) has argued for the renewal of 1.8 million units of Section 8 housing to avoid an "unprecedented explosion of homelessness" in the U.S. (HUD, 1997, p. 2). The existing empirical evidence does not support

these views. The results reported by Early (1998) indicate that current subsidized housing programs do not have much of an impact on the number of homeless since they are not targeted toward those households with a high probability of being homeless. Therefore, for additional units of subsidized housing to be successful in reducing homelessness, they need to be targeted toward high-risk households. The percentage of units going to very poor households (those with income less than five thousand dollars) is used as a proxy for the targeting of units toward households that are likely to be homeless without their subsidy. The number of subsidized households divided by the number of poor persons is included to measure the importance of the overall availability of subsidized housing in explaining variations in rates of homelessness.

4. Results

Table 3 contains the results of regressions explaining the incidence of total, sheltered, and street homelessness among the poor across 224 metropolitan areas, accounting for the likely endogeneity of shelter availability. The F statistics reveal that we can reject at the 1% level of significance the hypothesis that the coefficients of all variables in the total, sheltered, and street homeless regressions are zero. The fit of the regression explaining street homelessness is considerably poorer than that of the regression explaining sheltered homelessness. The most plausible explanation for this result is that the extent of the undercount of the street homeless varies more than the extent of the undercount of the sheltered homeless across metropolitan areas.

Beyond these overall tests of the models, the most obvious result of the regressions is that almost no variables are statistically significant at the usual levels. More importantly, we are able to reject the null hypothesis that the coefficient is zero against the alternative that it has the

expected sign at the 10% level of significance in only a few cases. We can be very confident that an increase in the price of rental housing will increase the number of sheltered homeless and the total homeless population. We can be moderately confident that it will increase the number of street homeless, that is, we just fail to reject the null hypothesis at the 10 percent significance level. We can also be fairly confident that targeting subsidized housing toward very poor households reduces the number of sheltered and total homeless and that higher temperature increases the number of street homeless. The other coefficients are not estimated precisely enough to be at all confident of the effects of the corresponding variables. Only 36 percent of the estimated coefficients have the expected sign.[11]

For some variables such as the percentage of the poor who are female and percentage whose income is less than 50 percent of the poverty line, the imprecision in estimating coefficients might be explained by the small variation in the value of the variable across metropolitan areas. For others, it might be largely attributable to measurement error in the explanatory variables. For example, our variable measuring shelter quality is the real cost per bed of operating the shelter. It does not reflect the real market value of the structure. Our data for McKinney Act funding per poor person and per-capita spending on mental hospitals are for the state rather than the specific metropolitan area. Still other variables are arguably endogenous. For example, it might reasonably be argued that mental health spending per capita is positively correlated with the error term in our regressions explaining homelessness because states with a high incidence of mental illness, a variable not included in our regressions, will have high mental health spending per capita and more homelessness. However, an important

[11] When the sample is limited to the 109 metropolitan areas where data on shelter beds are available, 62 percent of the coefficients had the expected sign. However, we are able to reject the null hypothesis that the coefficient is zero against the alternative that it has the expected sign at the 10% level of significance in only 4 percent of the cases as opposed to 11 percent of the cases when the larger sample is used.

explanation for the estimated coefficients with unexpected signs and the imprecision of the estimates is surely that the extent of the undercount of the homeless, especially the street homeless, varies enormously across metropolitan areas.

It is important to realize that the most plausible simple models of the extent of the undercount of the homeless imply that all of the estimates in Table 3 are biased towards zero. Since the undercount was surely greatest for the street homeless, the bias is greatest in the regression explaining street homelessness. If the Census count of the street (or sheltered) homeless were the same fraction F of the number of street (or sheltered) homeless in each locality, then the estimates of the coefficients of the explanatory variables in Table 3 would be a fraction F of the estimates based on the true rate of homelessness. Since this undercount would result in an estimate of the standard deviation of the error term in the regression that is a fraction F of the standard error based on the true rate of homelessness, it would have no effect on hypothesis tests involving the coefficients of the explanatory variables. The t-scores associated with the coefficients of explanatory variables are unaffected by this simple type of measurement error in the rate of homelessness.

If the dependent variable in a linear regression model is the logarithm of the rate of homelessness as is the case in a number of previous studies and in some of the regressions reported in Table 6, the OLS estimates of the coefficients of the explanatory variables are unaffected by the aforementioned undercount and the estimated constant term is reduced by the absolute value of the logarithm of F. So if one of the regressors in the linear regression with this dependent variable is the logarithm of a variable, the estimate of the percentage change in the number of homelessness in response to a given percentage change in the variable is unaffected by the proportional undercount. Since the undercount has no effect on the standard deviation of

the error term in the regression in this case, it has no effect on hypothesis tests involving coefficients of the explanatory variables.

Unfortunately, even for these simple stochastic models and in the simple case where the extent of the undercount is the same everywhere in percentage terms, correcting for the biases in the estimators of the parameters requires knowledge of the extent of the undercount, and no credible estimate of the extent of the undercount exists for either the street or sheltered homeless. So even if the same fraction of the homeless were missed everywhere, we have no basis for correcting for the bias in the estimated coefficients.

A more realistic model of the undercount would recognize that the extent of the undercount varied from place to place. For example, we might assume that the observed count in each locality is a fraction of the true count plus an error term with mean zero and constant variance and that this error term is uncorrelated with the error term reflecting the unobserved determinants of the number of homeless and the explanatory variables in the regression. As a result, the standard deviation of the error term in the regression explaining the number of homeless counted is greater than it would have been had the extent of the undercount been the same everywhere. This is surely a major explanation for the large number of insignificant coefficients reported in this paper and other papers based on the 1990 Census counts of the homeless.

Taken at face value, the results in Table 3 imply that a sufficiently large increase in the targeting of housing assistance on the poorest eligible households would eliminate homelessness. For example, a metropolitan area with mean values of all variables contains about 73 counted homeless per 10,000 poor persons and about 15 percent of households currently receiving housing assistance in the area have incomes less than $5,000. Increasing the percentage of

recipients of housing assistance who are in this group to 76 percent would reduce homelessness to zero.

This calculation is not affected by an undercount of the type analyzed above. However, it is subject to the usual caveats, namely its reliance on a strong assumption about the functional form of the regression equation and the considerable uncertainty about the magnitude of the coefficient of this explanatory variable even if the assumed functional form were correct. Furthermore, it is not reasonable to believe that improved targeting of housing assistance without substantial outreach activities will reach the chronic street homeless who suffer from severe mental illness and substance abuse. These people do not typically apply for housing assistance.

Recent legislation has improved the targeting of housing assistance toward the poor. However, these policies do not go far enough in their targeting the poorest of the poor and a larger reduction in the number of homeless can be obtained at the same cost to taxpayers by switching the type of housing assistance offered low-income households. The Quality Housing and Work Responsibility Act of 1998 required that at least 75 percent of new recipients of tenant-based housing vouchers have incomes less than limits based on 30 percent of the local median.[12] Although a larger fraction of new recipients of vouchers should be from the ranks of the very poor, the larger opportunity for eliminating homelessness involves providing less assistance through publicly and privately owned projects. Due to concerns about concentrating large numbers of the poorest households in projects, the 1998 Act set much lower targeting standards for HUD's project-based assistance, namely 40 rather than 75 percent with incomes less than limits based on 30 percent of the local median. Since the empirical evidence is unanimous that it costs substantially more to provide equally good housing with any form of

[12] Specifically, four-person households must have incomes less than 30 percent of the local median in order to be in the targeted group. The limits are lower for smaller families and higher for larger families and are obtained by multiplying the four-person limit in the locality by nationally uniform constants.

project-based assistance than with tenant-based vouchers and the weaker targeting toward very poor households in housing projects hinders the goal of eliminating homelessness, there is a strong case for phasing out project-based assistance as rapidly as possible.[13]

Increasing the number of the poorest households who receive housing assistance without increasing the total number of assisted households could not occur rapidly without major changes in program rules. Given current turnover rates in housing programs (about ten percent a year), it would require many years to increase the percentage of recipients with annual incomes less than $5,000 from 15 to 76 percent. Furthermore, the greater targeting of assistance on the poorest households would increase government expenditure on housing assistance because poorer households receive larger subsidies. Although HUD currently has the budget authority to spend this larger amount, Congress could respond to a large increase in expenditure on housing assistance in a number of ways, including reducing the number of households assisted.[14] Finally, we would not want to use the percentage of recipients with annual incomes less than $5,000 as our measure of targeting for purposes of focusing assistance on the poorest of the poor because it takes no account of differences in family size or prices across areas.[15] Some fraction of the upper income limit for eligibility for families of each size in each locality would be a superior measure for this purpose.

[13] GAO (2001) produces evidence of the per-unit cost of apartments of the same size in the same metropolitan area under the largest programs currently used to produce new subsidized housing. The Low Income Housing Tax Credit accounts for the bulk of these units. Olsen (2000) provides a description and critical appraisal of the data and methods used in the major older studies of the cost-effectiveness of low-income housing programs.

[14] When Congress authorizes additional units under housing programs, it provides budget authority for the duration of the government's contractual commitment based on the assumption that families with no income will occupy all of the units. This is the maximum expenditure on these units over the life of the contract. The distinction between the budget authority for the newly authorized units in a year and actual expenditure on housing assistance in that year is a major source of confusion in discussions of housing policy. Many confuse a reduction in budget authority with a decrease in housing assistance.

[15] We used this variable in the analysis due to its accessibility.

Table 4 reports the results of regressions identical to those in Table 3 except that shelter availability is treated as exogenous. Since we believe that shelter availability is positively correlated with the error term in these regressions, we expect that the OLS estimator of its coefficient will be biased upwards in each regression. The results in Tables 3 and 4 support this expectation for the regressions explaining total and sheltered homelessness. Otherwise, the qualitative results in the two tables are similar.

Since it is reasonable to believe that the sheltered homeless count is fairly accurate but the street homeless count is much too low, we estimated relationships explaining the overall rate of homelessness based on the assumptions that the shelter count was completely accurate but that only a fifth and a tenth of the street homeless were counted. The first column of Table 5 reports results based on the actual census counts. These are the results reported in the first column of Table 3. The last two columns report results based on the alternative assumptions about the extent of the undercount of the street homeless.

The estimated coefficients in Table 5 can be calculated from the estimated coefficients in Table 3 and the assumptions about the extent of the undercount. For example, if a fifth of the street homeless are counted everywhere, the estimated coefficient of the fraction very poor in Table 5 must be equal to its estimated coefficient in the sheltered homeless regression in Table 3 plus five times its estimated coefficient in the street homeless regression. If the estimated coefficients of a variable in the regressions explaining sheltered and street homelessness in Table 3 have the same sign, the estimated coefficient of the variable in the regression explaining the overall rate of homelessness in Table 5 must increase in absolute value with the assumed extent of the undercount.

The estimated standard deviations of the estimators of the coefficients in Table 5 cannot be determined from the results reported in Table 3. They depend in part on the extent of the sample correlation across metropolitan areas between the residuals in the sheltered and street homeless regressions in Table 3. If these residuals are positively correlated, the estimated standard deviations of the estimators of the coefficients in Table 5 will increase with the assumed extent of the undercount. This is the case in Table 5.

Accounting for an undercount of the street homeless that is the same percentage everywhere does affect the t-scores in a regression explaining the overall rate of homelessness and hence the results of tests of hypotheses. However, even after accounting for undercounts of the assumed magnitudes, we are rarely able to reject at usual levels of significance the hypothesis that a coefficient is zero against the alternative that it has the expected sign. Under the three assumptions regarding the undercount of the street homeless, we find a positive and statistically significant relationship between the rate of homelessness and the price of rental housing. If only a fifth or a tenth of the street homeless are counted, we can confidently conclude that locations with higher average March temperatures have a higher overall rate of homelessness.

Table 6 contains the results of sensitivity analyses involving the functional form of the systematic part of the relationship explaining the overall rate of homelessness and an alternative assumption about the nature of the error term. The results reported in the first column of Table 6 are the same as the results reported in the first column of Table 3. The second column assumes that the standard deviation of the error term in the preceding regression equation is proportional to the number of poor in the metropolitan area and provides the appropriate weighted least

squares estimates.[16] The third column reports the results of a regression in which the dependent variable is the logarithm of the overall rate of homelessness and the error term is assumed to be homoskedastic. The last column reports results of a similar regression in which the explanatory variables (except those which had zero values for some observations) are the logarithms of the explanatory variables in the third regression.

We use each regression to predict the overall rate of homelessness and calculate the sum of squared deviations between the actual and predicted values. The results reported at the bottom of Table 6 reveal that the initial specification involving untransformed variables performed best in predicting the rate of homelessness.[17]

Table 7 summarizes the results of the preceding regressions explaining the overall incidence of homelessness that account for the endogeneity of shelter availability. The estimated coefficients of the percent of the poor who have incomes below 50 percent of the poverty line, the price of rental housing service, real annual spending per shelter bed, percent of subsidized units occupied by households with incomes less than $5000, and average March temperature have the expected sign in all cases. No other coefficient has the expected sign in the majority of cases. For the price of rental housing services, we can reject the null hypothesis of no effect on the incidence of homelessness against the alternative of the expected effect in all cases. We can do it in the majority of cases only for this variable and for average March temperature. The consistency of the signs of the three other estimated coefficients across the different specifications give us some confidence about the direction of the effects of these variables.

[16] This is analogous to QRS's assumption that the standard deviation of the error term is proportional to the population of the metropolitan area. Their dependent variable is the fraction of the entire population rather than the poor who are homeless.

[17] Two metropolitan areas reported no homeless. These two observations were deleted in estimating the third and fourth regressions.

Table 8 reports the results of regressions explaining the fraction of the homeless who were in shelters on the night of the Census count under alternative assumptions about the extent of the undercount of the street homeless. It includes only the most theoretically compelling explanatory variables rather than the complete list of explanatory variables used in the previous regressions. For these regressions, we used the number of shelter beds per homeless person instead of the number of shelter beds per poor person in the metropolitan area. Temperature and precipitation during the week of the S-Night Census count are used to explain the fraction of the homeless in shelters rather than the average March values of these variables. These results confirm the well-known fact that shelters fill up on the coldest nights. For the coefficient on temperature during the S-Night count, we reject the hypothesis that the coefficient is zero in favor of the alternative that the coefficient has the expected sign at the 1 percent level of significance in all three regressions. In the regressions assuming that a fifth or a tenth of the street homeless were counted, shelter quality is positively and statistically significantly related to the fraction of the homeless in shelters. The coefficients of crime in the regressions based on the assumptions that a fifth and a tenth of the street homeless were counted are highly significant but have unexpected signs. The results imply that a higher crime rate leads to a smaller fraction of the homeless being in shelters. These unexpected signs may be due to reverse causation. Staying in shelters may deter crime by the homeless and hence reduce the overall crime rate. The other coefficients are not estimated precisely enough to be at all confident of the effects of the corresponding variables.

Table 9 gives the results explaining the fraction of the homeless in shelters on the night of the Census enumeration using a full set of explanatory variables. The relationship between the fraction of the homeless in shelters and temperature and crime are similar to the results using the

most theoretically compelling variables. However, a statistically significant relationship between shelter quality and the fraction in shelters is no longer evident.

5. Conclusions

Although the 1990 Decennial Census is the best source of aggregate cross-sectional data available to study the determinants of homelessness, it did not produce sufficiently accurate counts, especially of the street homeless, to permit very precise estimates of the effects of many factors which surely affect the rate of homelessness. We are rarely able to reject at the usual significance levels the hypothesis that a coefficient is zero against the alternative that it has the expected sign. The primary exception is the price of housing in many regressions explaining the overall incidence of homelessness and the rates of sheltered and street homelessness. Coefficients on average March temperature usually had the expected sign and were statistically significant in the majority of regressions explaining the overall rate of homelessness and in the regression explaining street homelessness.

With respect to the policy variables that motivated this study, only the coefficient on the fraction of subsidized units going to households with annual incomes less than $5000 had the expected sign in all regressions and was statistically significant in some regressions. A reasonable interpretation of our results is that a sufficiently large increase in the targeting of housing assistance on the poorest eligible households would eliminate homelessness among those who apply for housing assistance. Achieving this goal rapidly with little or no additional expenditure requires phasing out project-based assistance.

Almost all knowledgeable observers agree that the 1990 Decennial Census significantly undercounted the street homeless. Plausible models of the undercount imply that in regressions

with a rate of homelessness as the dependent variable estimators of the coefficients of explanatory variables will be biased towards zero. In regressions with the logarithm of a rate of homelessness as the dependent variable, only the estimator of the constant term will be biased downwards. Since no credible estimate of the extent of the undercount exists for either the street or sheltered homeless, we have no basis for correcting for the bias in the estimated coefficients. Therefore, care should be taken in predicting the effects of policy interventions on the number of homeless based on the results in this paper and previous studies.

The official position of the U.S. Bureau of the Census is that they did not attempt to count the homeless but rather attempted to include some additional homeless people in the census who would not have been counted using standard census procedures. This paper casts doubt on the value of this activity. The additional households counted clearly had little impact on any political jurisdiction's share of the total count or its count of low-income persons and hence no significant effect on Congressional representation or the allocation of federal funds. They added less than one tenth of a percent to the total count. This paper shows that these counts, especially of the street homeless, were not sufficiently accurate to permit good estimates of the effects of many factors that surely affect the rate of homelessness. This raises questions about the justification for the enormous cost per additional household counted. Interestingly, the Bureau does not intend to release data on the homeless people counted in 2000, thereby precluding the use of these data to study the causes of homelessness.

References

Allgood, Sam; Moore, Myra L.; and Warren, Ronald S. "The Duration of Sheltered Homelessness in a Small City." *Journal of Housing Economics* 6 (1997): 60-80.

American Appraisal Associates. *Boeckh Building Cost Index Numbers*. Milwaukee: American Appraisal Associates, Inc., January-February 1990.

Appelbaum, Richard P., Dolny, Michael, Dreier, Peter and Gilderbloom, John I. "Scapegoating Rent Control: Masking the Causes of Homelessness." *Journal of the American Planning Association* 57 (1991): 153-64.

Bohanon, Cecil. "The Economic Correlates of Homelessness in Sixty Cities." *Social Science Quarterly* 72 (1991): 817-25.

Burt, Martha R. *Over the Edge; The Growth of Homelessness in the 1980's.* Washington, D.C.: Urban Institute Press, 1992.

Burt, Martha R., and Taeuber, Cynthia M. "Overview of Seven Studies That Counted or Estimated Homeless Populations," in Cynthia M. Taeuber (ed.) *Conference Proceedings for Enumerating Homeless Persons: Methods and Data Needs*. Washington, D.C.: U.S. Department of Commerce, Economics and Statistics Administration, Bureau of the Census, March 1991.

Committee on Ways and Means. *1994 Green Book: Overview of Entitlement Programs, Background Material and Data on Programs within the Jurisdiction of the Committee on Ways and Means*. Washington, D.C.: U.S. House of Representatives, 1994.

Congressional Quarterly Inc. *Politics in America, 1992, The 102nd Congress*, edited by Phil Duncan. Washington, D.C.: Congressional Quarterly Inc., 1992.

Cragg, Michael and O'Flaherty, Brendan. "Do Homeless Shelter Conditions Determine Shelter Population? The Case of the Dinkins Deluge." *Journal of Urban Economics* 46 (1999): 377- 415.

Culhane, Dennis, and Kuhn, Randall. "Patterns and Determinants of Public Shelter Utilization among Homeless Adults in New York City and Philadelphia." *Journal of Policy Analysis and Management*. 17 (1998): 23-43.

Culhane, Dennis; Lee, Chang-Moo; and Wachter, Susan M. "Where the Homeless Come From: A Study of the Prior Address Distribution of Families Admitted to Public Shelters in New York City and Philadelphia." *Housing Policy Debate*. 17 (1996): 327-65.

Culhane, Dennis, and Metraux, Stephen. "One-Year Rates of Public Shelter Utilization by Race/Ethnicity, Age, Sex and Poverty Status for New York City (1990 and 1995) and Philadelphia (1995)." *Population Research and Policy Review*. 18 (1999): 219-36.

Early, Dirk W. "A Microeconomic Analysis of Homelessness: An Empirical Investigation Using Choice-Based Sampling." *Journal of Housing Economics* 8 (1999): 312-327.

Early, Dirk W. "The Role of Subsidized Housing in Reducing Homelessness: An Empirical Investigation Using Micro-Data." *Journal of Policy Analysis and Management* 17 (1998): 687-696.

Early, Dirk W. and Olsen, Edgar O. “Rent Control and Homelessness.” *Regional Science and Urban Economics* 28 (1998): 797-816.

Elliott, Marta, and Krivo, Lauren. "Structural Determinants of Homelessness in the United States." *Social Problems* 38 (1991): 113-31.

Gaquin, Deirdre A. and Littman, Mark S., eds. *1999 County and City Extra: Annual Metro, City, and County Data Book, Eighth Edition*. Washington, D.C.: Bernan Press, 1999.

Grimes, Paul and Chressanthis, George. “Assessing the Effect of Rent Control on Homelessness.” *Journal of Urban Economics* 41 (1997): 23-37.

Honig, Marjorie and Filer, Randall K. "Causes of Intercity Variation in Homelessness." *American Economic Review* 83 (1993): 248-55.

Interagency Council on the Homeless. *Homelessness: Programs and the People they Serve*. Washington, D.C.: Interagency Council on the Homeless, 1999.

Interagency Council on the Homeless. *The Federal Plan to Break the Cycles of Homelessness*. Washington, D.C.: Interagency Council on the Homeless, 1994.

Interagency Council on the Homeless. *The 1990 Annual Report of the Interagency Council on the Homeless*. Washington, D.C.: Interagency Council on the Homeless, 1991.

Malpezzi, Stephen, Chun, Gregory H. and Green, Richard K. “New Place-to-Place Housing Price Indexes for U.S. Metropolitan Areas, and their Determinants. *Real Estate Economics* 26 (1998): 235-274.

National Climatic Data Center. *Normal Daily Minimum Temperature*. Asheville, NC: National Climatic Data Center. Available World Wide Web page <http://www.ncdc.noaa.gov/ol/climate/online/ccd/mintemp.html> (last modified August, 1997[1]).

National Climatic Data Center. *Normal Monthly Precipitation*. Asheville, NC: National Climatic Data Center. Available World Wide Web <http://www.ncdc.noaa.gov/ol/climate/online/ccd/nrmlprcp.html> (last modified August, 1997[2]).

Olsen, Edgar O. “The Cost-Effectiveness of Alternative Methods of Delivering Housing Subsidies.” Thomas Jefferson Center for Political Economy, Working Paper 351, December 2000.

Quigley, John M., Raphael, Steven and Smolensky, Eugene. “Homeless in America, Homeless in California.” *Review of Economics and Statistics* 83 (2001): 37-51.

Quigley, John M. "Does Rent Control Cause Homelessness? Taking the Claim Seriously." *Journal of Policy Analysis and Management* 9 (1990): 89-93.

Rossi, Peter H. *Down and Out in America: The Origins of Homelessness*. Chicago: University of Chicago Press, 1989.

Taeuber, Cynthia M., ed. *Conference Proceedings for Enumerating Homeless Persons: Methods and Data Needs*. Washington, D.C.: U.S. Department of Commerce, Economics and Statistics Administration, Bureau of the Census, March 1991.

Taeuber, Cynthia M., and Siegel, Paul. "Counting the Nation's Homeless Population in the 1990 Census" in Cynthia M. Taeuber (ed.) *Conference Proceedings for Enumerating Homeless Persons: Methods and Data Needs*. Washington, D.C.: U.S. Department of Commerce, Economics and Statistics Administration, Bureau of the Census, March 1991.

Troutman, William, Jackson, John D. and Ekelund, Robert. "Public Policy, Perverse Incentives, and the Homeless Problem." *Public Choice* 98 (1999): 195-212.

Tucker, William. *The Excluded Americans*. Washington, D.C.: Regnery Gateway, 1990.

Tucker, William. "Where do the Homeless come From?" *National Review* (1987): 32-43.

Uccello, Cori E., McCallum, Heather R. and Gallagher, L. Jerome. "State General Assistance Programs, 1996." Manuscript. Washington, D.C.: Urban Institute. Available World Wide Web <www.newfederalism.urban.org/pdf/sgap96.pdf> 1996.

U.S. Bureau of Labor Statistics. *Consumer Price Indexes for Rent and Rental Equivalence*. Washington, D.C. Available World Wide Web <http://www.bls.gov/cpifact6.htm> (last modified, December 28, 1999).

U.S. Census Bureau. *1990 Census of Population and Housing – CPH-4: Population and Housing Characteristics for Congressional Districts of the 103rd Congress*. Washington, D.C. Available World Wide Web <www.census.gove/prod/1/90dec/cph4/90cph4.htm> (last modified, April 9, 1999).

U.S. Department of Agriculture. *Weekly Weather and Crop Bulletin*. Washington, D.C.: National Agricultural Statistics Service and World Agricultural Outlook Board, 1990.

U.S. Department of Commerce. *Census of Governments, 1992, Volume 1: Government Organization, Number 1: Government Organization*. Washington, D.C.: Economics and Statistics Administration, 1992[1].

U.S. Department of Commerce. *Census of Population and Housing, 1990*. Summary Tape File 3 on CD-ROM [machine-readable data files] / prepared by the Bureau of the Census – Washington, D.C. Available World Wide Web <http://venus.census.gov/cdrom/lookup>, 1992[2].

U.S. Department of Commerce. *Commerce News: Census Bureau Releases 1990 Decennial Counts for Persons Enumerated in Emergency Shelters and Observed on Streets*. Washington, D.C.: Economics and Statistics Administration, 1991.

U.S. Department of Health and Human Services. *State and County Mental Hospitals- United States and Each State, 1986*. Washington, D.C.: Survey and Reports Branch, Division of Biometry and Applied Sciences, 1990.

U.S. Department of Housing and Urban Development. *The FY 1998 Budget of the Department of Housing and Urban Development: Executive Summary*. Available World Wide Web <http://www.hud.gove/budsum1.html#I> Washington, D.C.: U.S. Department of Housing and Urban Development, 1997.

U.S. Department of Housing and Urban Development. *A Picture of Subsidized Housing*. [Computer File] / prepared by the Office of Policy Development and Research – Washington, D.C., 1996.

U.S. Department of Housing and Urban Development. *Report to Congress on Rent Control*. Washington, D.C.: Office of Policy Development and Research, 1991.

U.S. Department of Housing and Urban Development. *A Report on the 1988 National Survey of Shelters for the Homeless*. Washington, D.C.: Office of Policy Development and Research, 1989.

U.S. Department of Housing and Urban Development. *A Report to the Secretary on the Homeless and Emergency Shelters*. Washington, D.C.: Office of Policy Development and Research, 1984.

U.S. Department of Justice. *Uniform Crime Reports for the United States*. Washington, D.C.: Federal Bureau of Investigation, 1991.

U.S. General Accounting Office, *Federal Housing Programs: What They Cost and What They Provide*, GAO-01-901R, July 18, 2001.

U.S. General Accounting Office. *1990 Census: Limitations in Methods and Procedures to Include the Homeless*. Washington, D.C.: General Accounting Office, 1991.

Table 1. Description and source of variables

Variable	Mean (std. dev.)	Description and source
Sheltered homeless	794.95 (2893.77)	S-Night count of the number of homeless staying in shelters March 21, 1990 (U.S. Department of Commerce, 1992[2]).
Street homeless	124.71 (769.91)	S-Night count of the number of homeless staying on the streets March 21, 1990 (U.S. Department of Commerce, 1992[2]).
Poverty population	96630.86 (259408.30)	Persons in the metropolitan area with 1989 incomes below adjusted poverty line controlling for differences in prices across areas. Constructed by the authors (U.S. Department of Commerce, 1992[2], Malpezzi, Chun, and Green, 1998, U.S. Bureau of Labor Statistics, 1999).
Dependent Variables:		
Total homelessness	73.05 (64.09)	Sheltered + street homeless divided by the poverty population controlling for differences in prices across areas X 10,000.
Sheltered homelessness	67.68 (56.57)	Sheltered homeless divided by the poverty population controlling for differences in prices across areas X 10,000.
Street homelessness	5.37 (19.45)	Street homeless divided by the poverty population controlling for differences in prices across areas X 10,000.
Percent of homeless in shelters	95.197 (11.518)	Sheltered homeless divided by the total number of homeless X 100.
Explanatory variables:		
Percent very poor	44.07 (5.14)	Percent of the poor with 1989 incomes less than 50% of the adjusted poverty line controlling for differences in prices across areas (U.S. Department of Commerce, 1992[2]).
Unemployment Rate	6.41 (1.89)	Metropolitan area civilian unemployment rate, 1989 (U.S. Department of Commerce, 1992[2]).
Price of rental housing service	418.09 (94.10)	Renter housing price index (Malpezzi, Chun, and Green, 1998).
Rental vacancy rate	8.53 (3.37)	Percent of all rental units that are vacant and available for rent (U.S. Department of Commerce, 1992[2]).
Shelter beds per poor person	262.65 (342.82)	Number of shelter beds (Burt, 1992) divided by the poverty population controlling for differences in prices across areas X 10,000. Actual value if observed, fitted otherwise.
Shelter beds per poor person (fitted value)	262.65 (312.91)	Fitted value of shelter beds per poor person from the first stage of the estimation procedure.
Shelter beds per homeless person	3.93 (6.22)	Number of shelter beds (Burt, 1992) divided by the number of homeless persons. Actual value if observed, fitted otherwise.
Shelter beds per homeless person (fitted value)	3.93 (5.57)	Fitted value of the number of shelter beds divided by the number of homeless persons from the first stage of the estimation procedure.

Table 1. Description and source of variables *(continued)*

Shelter quality	902.226 (505.820)	Average annual spending on a bed in an emergency homeless shelter in 1988 divided by housing price index. Constructed by the authors using data from the 1988 National Survey of Shelters for the Homeless (HUD, 1989). Varies by region and city size.
McKinney Act funding	55.45 (22.95)	State McKinney Act funding per poor person, fiscal years 1987-1990 (Interagency Council on the Homeless, 1991) divided by price index for all goods.
Availability of subsidized housing	1555.02 (680.38)	Number of subsidized units divided by the poverty population adjusted for price changes X 10,000. Constructed by the authors from A Picture of Subsidized Housing (HUD, 1996).
Availability of subsidized housing missing	0.0313 (0.1743)	1 if availability of subsidized housing is missing, 0 otherwise.
Subsidized housing targeted toward the very poor	14.78 (9.09)	Percent of subsidized units going to households with income less than $5,000 in 1995. Constructed by the authors from A Picture of Subsidized Housing (HUD, 1996).
Subsidized housing targeted toward the very poor missing	0.0446 (0.2070)	1 if subsidized housing targeted toward the poor is missing, 0 otherwise.
Rent control	0.03 (0.16)	1 if the metropolitan area had rent control, 0 otherwise (HUD, 1991).
AFDC	405.90 (165.32)	Maximum monthly AFDC payment for a family of three (Committee on Ways and Means, 1994) divided by price index for all goods.
GA	118.73 (134.84)	Maximum monthly general assistance payments for a single person living along in 1989 (Burt, 1992; Uccello, McCallum, and Gallagher, 1996) divided by price index for all goods.
GA missing	0.1116 (0.3156)	1 if GA is missing, 0 otherwise.
Spending on mental hospitals	26.38 (18.60)	State per capita expenditures by state and county mental hospitals in 1986 (U.S. Department of Health and Human Services, 1990) divided by price index for all goods.
Percent of poor African-American	23.94 (20.17)	Percent of the unadjusted 1989 poverty population African-American (U.S. Department of Commerce, 1992[2]).
Percent of poor female	57.80 (2.52)	Percent of the unadjusted 1989 poverty population female (U.S. Department of Commerce, 1992[2]).
Average March temperature	36.84 (9.90)	Average temperature for March (National Climatic Data Center, 1997[1]).
Temperature during S-Night count	47.75 (11.35)	Average temperature during the week of the S-Night homeless count (U.S. Department of Agriculture, 1990).
Average March precipitation	3.12 (1.36)	Average precipitation for March (National Climatic Data Center, 1997[2]).

Table 1. Description and source of variables *(continued)*

Precipitation during S-Night count	0.27 (0.33)	Precipitation during the week of the S-Night homeless count (U.S. Department of Agriculture, 1990).
Crime rate	669.65 (539.34)	Number of violent crimes per 100,000 persons in 1990 (U.S. Department of Justice, 1991).
Crime rate missing	0.1518 (0.3156)	1 if crime rate missing, 0 otherwise.
Northeast	0.15 (0.36)	1 if metropolitan area is in the Northeast census region, 0 otherwise.
South	0.39 (0.49)	1 if metropolitan area is in the South census region, 0 otherwise.
West	0.20 (0.40)	1 if metropolitan area is in the West census region, 0 otherwise.
Additional variables used in the first stage regressions explaining shelter availability:		
Liberal political climate	45.00 (32.20)	Index of the voting record of representative to congress, 100 = liberal, 0 = conservative (Congressional Quarterly, Inc., 1992)
Price of land	2448.74 (1344.30)	Price of an acre of farmland in 1997 (Gaquin and Littman, 1999) divided by price index for all goods.
Price of land missing	0.1116 (0.3156)	1 if price of land is missing, 0 otherwise.
Building cost index	1456.03 (211.67)	Boeckh building cost index for Apartments, Hotels and Office Buildings (American Appraisal Associates, 1990) divided by price index for all goods.
Median income	31295.48 (3664.00)	Median income of the metropolitan area (U.S. Department of Commerce, 1992[2]) divided by price index for all goods.
Population of the metropolitan area	822.21 (1822.81)	Population of the metropolitan area in 1,000s (U.S. Department of Commerce, 1992[2]).
Number of governments	148.10 (238.03)	Number of governments in the counties in which the metropolitan area lies in 1992 (U.S. Department of Commerce, 1992[1]).
Number of governments per capita	2.49 (1.87)	Number of governments in the counties in which the metropolitan area lies in 1992 (U.S. Department of Commerce, 1992[1]) divided by the population of these counties X 10,000.
Unless otherwise noted, all variables are for 1990.		

Table 2. Expected sign of coefficients explaining rates of homelessness and percent of homeless in shelters

Explanatory variable	Total homeless	Sheltered homeless	Street homeless	Percent of homeless in shelters
Percent very poor	+	+	+	0
Unemployment rate	+	+	+	0
Price of rental housing service	+	+	+	0
Rental vacancy rate	–	–	–	0
Shelter beds per poor person	+	+	–	na
Shelter beds per homeless person	na	na	na	+
Shelter quality	+	+	–	+
McKinney Act funding	+	+	–	+
Availability of subsidized housing	–	–	–	0
Subsidized housing targeted toward the very poor	–	–	–	0
Rent control	?	?	?	0
AFDC	–	–	–	0
GA	–	–	–	0
Spending on mental hospitals	–	–	–	0
Percent of poor African-American	?	?	?	?
Percent of poor female	–	?	–	+
Average March temperature	+	–	+	na
Temperature during S-night count	na	na	na	–
Average March precipitation	–	+	–	na
Precipitation during S-night count	na	na	na	+
Crime rate	–	+	–	+

Table 3. Regressions explaining rates of homelessness accounting for potential endogeneity of shelter availability

Dependent variable	Total homeless	Sheltered homeless	Street homeless
Explanatory variable	Coefficient (Std. Error)	Coefficient (Std. Error)	Coefficient (Std. Error)
Percent very poor	0.8015 (1.2642)	0.8725 (1.1116)	-0.0710 (0.4355)
Unemployment rate	-2.8954 (2.6384)	-2.5554 (2.3200)	-0.3400 (0.9090)
Price of rental housing service	0.2517*** (0.0841)	0.2149*** (0.0740)	0.0368 (0.0290)
Rental vacancy rate	3.7598 (1.6300)	3.3627 (1.4333)	0.3971 (0.5615)
Shelter beds per poor person (fitted value)	-0.0518 (0.0641)	-0.0462 (0.0563)	-0.0056 (0.0221)
Shelter quality	0.0056 (0.0097)	0.0051 (0.0085)	0.0005 (0.0033)
McKinney Act funding	-0.0464 (0.2740)	0.0218 (0.2409)	-0.0682 (0.0944)
Availability of subsidized housing	0.0296 (0.0089)	0.0283 (0.0078)	0.0014 (0.0031)
Subsidized housing targeted toward the very poor	-1.1965* (0.9130)	-1.2013* (0.8028)	0.0047 (0.3145)
Rent control	-20.3784 (30.5926)	-15.3805 (26.9010)	-4.9979 (10.5395)
AFDC	0.0274 (0.0649)	0.0129 (0.0571)	0.0146 (0.0224)
GA	0.0151 (0.0534)	0.0104 (0.0470)	0.0047 (0.0184)
Spending on mental hospitals	0.4233 (0.4403)	0.3800 (0.3872)	0.0433 (0.1517)
Percent of poor African-American	-0.7218 (0.4563)	-0.6254 (0.4012)	-0.0963 (0.1572)
Percent of poor female	2.4464 (3.1047)	2.4531 (2.7301)	-0.0066 (1.0696)
Average March temperature	0.5184 (0.7502)	-0.1362 (0.6597)	0.6546** (0.2584)
Average March precipitation	3.9326 (4.0095)	2.4546 (3.5256)	1.4780 (1.3813)
Crime rate	0.0027 (0.0089)	0.0015 (0.0078)	0.0012 (0.0031)
R-squared	0.3838	0.3885	0.2062
N	224	224	224
F	5.09	5.25	2.03

The dependent variable is (number of homeless / the number of poor persons) X 10,000.
Coefficients on region dummies, dummies for missing observations, and the constant term are omitted.
* = reject hypothesis that coefficient is zero against the alternative that it has the expected sign at the 10% level of significance.
** = reject hypothesis that coefficient is zero against the alternative that it has the expected sign at the 5% level of significance.
*** = reject hypothesis that coefficient is zero against the alternative that it has the expected sign at the 1% level of significance.

Table 4. Regressions explaining rates of homelessness treating shelter availability as exogenous

Dependent variable	Total homeless	Sheltered homeless	Street homelesss
Explanatory variable	Coefficient (Std. Error)	Coefficient (Std. Error)	Coefficient (Std. Error)
Percent very poor	0.7942 (1.2533)	0.8649 (1.0983)	-0.0707 (0.4355)
Unemployment rate	-2.4947 (2.5591)	-2.1404 (2.2427)	-0.3543 (0.8892)
Price of rental housing service	0.2173*** (0.0693)	0.1793*** (0.0607)	0.0380* (0.0241)
Rental vacancy rate	3.7690 (1.6159)	3.3722 (1.4161)	0.3968 (0.5615)
Shelter beds per poor person (actual value)	-0.0081 (0.0235)	-0.0009 (0.0206)	-0.0072 (0.0082)
Shelter quality	0.0031 (0.0090)	0.0025 (0.0079)	0.0006 (0.0031)
McKinney Act funding	0.0400 (0.2453)	0.1113 (0.2150)	-0.0713 (0.0852)
Availability of subsidized housing	0.0271 (0.0082)	0.0257 (0.0072)	0.0015 (0.0028)
Subsidized housing targeted toward the very poor	-0.9641 (0.8490)	-0.9605* (0.7440)	-0.0036 (0.2950)
Rent control	-14.5829 (29.3023)	-9.3775 (25.6792)	-5.2054 (10.1817)
AFDC	0.0523 (0.0549)	0.0386 (0.0481)	0.0137 (0.0191)
GA	0.0011 (0.0495)	-0.0041 (0.0434)	0.0052 (0.0172)
Spending on mental hospitals	0.2251 (0.3448)	0.1746 (0.3022)	0.0504 (0.1198)
Percent of poor African-American	-0.5250 (0.3661)	-0.4216 (0.3208)	-0.1034 (0.1272)
Percent of poor female	0.9746 (2.3504)	0.9285 (2.0597)	0.0461 (0.8167)
Average March temperature	0.5504 (0.7425)	-0.1030 (0.6507)	0.6534*** (0.2580)
Average March precipitation	3.9499 (3.9749)	2.4725 (3.4834)	1.4774 (1.3812)
Crime rate	0.0032 (0.0088)	0.0019 (0.0077)	0.0012 (0.0030)
R-squared	0.3944	0.4031	0.2063
N	224	224	224
F	5.16	5.35	2.06

The dependent variable is the (number of homeless / number of persons) X 10,000.
Coefficients on region dummies, dummies for missing observations, and the constant term are omitted.
* = reject hypothesis that coefficient is zero against the alternative that it has the expected sign at the 10% level of significance.
** = reject hypothesis that coefficient is zero against the alternative that it has the expected sign at the 5% level of significance.
*** = reject hypothesis that coefficient is zero against the alternative that it has the expected sign at the 1% level of significance.

Table 5. Regressions explaining homelessness under alternative assumptions about undercount of street homeless

Explanatory variable	Assumptions about street count: Actual street counts Coefficient (Std. Error)	Street counts x 5 Coefficient (Std. Error)	Street counts x 10 Coefficient (Std. Error)
Percent very poor	0.8015 (1.2642)	0.5175 (2.6156)	0.1626 (4.6831)
Unemployment rate	-2.8954 (2.6384)	-4.2553 (5.4589)	-5.9553 (9.7738)
Price of rental housing service	0.2517*** (0.0841)	0.3988** (0.1741)	0.5826** (0.3117)
Rental vacancy rate	3.7598 (1.6300)	5.3483 (3.3725)	7.3338 (6.0382)
Shelter beds per poor person (fitted value)	-0.0518 (0.0641)	-0.0742 (0.1325)	-0.1022 (0.2373)
Shelter quality	0.0056 (0.0097)	0.0075 (0.0201)	0.0098 (0.0359)
McKinney Act funding	-0.0464 (0.2740)	-0.3193 (0.5669)	-0.6605 (1.0150)
Availability of subsidized housing	0.0296 (0.0089)	0.0351 (0.0184)	0.0419 (0.0330)
Subsidized housing targeted toward the very poor	-1.1965* (0.9130)	-1.1775 (1.8890)	-1.1538 (3.3822)
Rent control	-20.3784 (30.5926)	-40.3699 (63.2970)	-65.3592 (113.3289)
AFDC	0.0274 (0.0649)	0.0857 (0.1343)	0.1585 (0.2405)
GA	0.0151 (0.0534)	0.0339 (0.1106)	0.0573 (0.1980)
Spending on mental hospitals	0.4233 (0.4403)	0.5967 (0.9110)	0.8134 (1.6311)
Percent of poor African-American	-0.7218 (0.4563)	-1.1071 (0.9440)	-1.5888 (1.6902)
Percent of poor female	2.4464 (3.1047)	2.4200 (6.4238)	2.3870 (11.5014)
Average March temperature	0.5184 (0.7502)	3.1368** (1.5522)	6.4097** (2.7790)
Average March precipitation	3.9326 (4.0095)	9.8446 (8.2957)	17.2346 (14.8529)
Crime rate	0.0027 (0.0089)	0.0076 (0.0184)	0.0138 (0.0329)
R-squared	0.3838	0.2921	0.2503
N	224	224	224
F	5.09	3.28	2.63

The dependent variable is (number of homeless / number of poor persons) X 10,000.
Coefficients on region dummies, dummies for missing observations, and the constant term are omitted.
* = reject hypothesis that coefficient is zero against the alternative that it has the expected sign at the 10% level of significance.
** = reject hypothesis that coefficient is zero against the alternative that it has the expected sign at the 5% level of significance.
*** = reject hypothesis that coefficient is zero against the alternative that it has the expected sign at the 1% level of significance.

Table 6. Regressions explaining homelessness under alternative assumptions about error term and functional form

Explanatory variable	(1) Coefficient (Std. Error)	(2) Coefficient (Std. Error)	(3) Coefficient (Std. Error)	(4) Coefficient (Std. Error)
Percent very poor	0.8015	4.7415***	0.0149	0.3291
	(1.2642)	(1.3180)	(0.0158)	(0.6210)
Unemployment rate	-2.8954	0.3758	-0.0551	-0.2023
	(2.6384)	(2.6024)	(0.0333)	(0.2207)
Price of rental housing service	0.2517***	0.2254***	0.0030***	0.5804*
	(0.0841)	(0.0659)	(0.0010)	(0.4239)
Rental vacancy rate	3.7598	0.3009	0.0591	0.3622
	(1.6300)	(1.6794)	(0.0204)	(0.1805)
Shelter beds per poor person	-0.0518	-0.0163	0.0000	0.4222***
(fitted value)	(0.0641)	(0.0770)	(0.0008)	(0.1133)
Shelter quality	0.0056	0.0012	0.0001	0.0181
	(0.0097)	(0.0092)	(0.0001)	(0.1013)
McKinney Act funding	-0.0464	-0.4860	-0.0031	-0.0045
	(0.2740)	(0.2909)	(0.0034)	(0.2078)
Availability of subsidized	0.0296	0.0414	0.0005	0.0004
housing	(0.0089)	(0.0114)	(0.0001)	(0.0001)
Subsidized housing targeted	-1.1965*	-2.3669**	-0.0064	-0.0045
toward the very poor	(0.9130)	(1.1699)	(0.0114)	(0.0105)
Rent control	-20.3784	-77.7885	-0.3298	0.0048
	(30.5926)	(18.2889)	(0.3806)	(0.3195)
AFDC	0.0274	0.0502	0.0010	0.5341
	(0.0649)	(0.0675)	(0.0008)	(0.2394)
GA	0.0151	-0.0525	0.0005	-0.0001
	(0.0534)	(0.0675)	(0.0007)	(0.0006)
Spending on mental hospitals	0.4233	1.5933	0.0038	-0.1263
	(0.4403)	(0.4588)	(0.0055)	(0.1670)
Percent of poor African-	-0.7218	-1.5673	-0.0108	0.0151
American	(0.4563)	(0.5116)	(0.0057)	(0.0573)
Percent of poor female	2.4464	2.3193	0.0388	-2.3143*
	(3.1047)	(3.7195)	(0.0390)	(1.6926)
Average March temperature	0.5184	1.5486**	0.0096	0.5672*
	(0.7502)	(0.6821)	(0.0094)	(0.3565)
Average March precipitation	3.9326	11.2084	-0.0007	-0.1034
	(4.0095)	(4.6613)	(0.0500)	(0.1141)
Crime rate	0.0027	0.0064	0.0001	0.0001
	(0.0089)	(0.0036)	(0.0001)	(0.0001)
N	224	224	222	222
F	5.09	12.37	7.54	8.17
MSE	2519.72	3289.75	2766.80	2771.85

(1) Untransformed dependent and explanatory variables, error term assumed homoskedastic.
(2) Untransformed dependent and explanatory variables, standard deviation of error term assumed proportional to poverty population.
(3) Log dependent variable, error term assumed homoskedastic.
(4) Log dependent variable, appropriate explanatory variables in logs, error term assumed homoskedastic.
MSE = mean squared error when equation is used to predict the rate of homelessness.
Coefficients on region dummies, dummies for missing observations, and the constant term are omitted.
* = reject hypothesis that coefficient is zero against the alternative that it has the expected sign at the 10% level of significance.
** = reject hypothesis that coefficient is zero against the alternative that it has the expected sign at the 5% level of significance.
*** = reject hypothesis that coefficient is zero against the alternative that it has the expected sign at the 1% level of significance.

Table 7. Results based on regressions explaining overall rate of homelessness

Explanatory variable	Table-Regression					
	3-1	5-2	5-3	6-2	6-3	6-4
Fraction very poor	•	•	•	***	•	•
Unemployment rate				•		
Price of rental housing service	***	***	***	***	***	*
Rental vacancy rate						
Shelter beds per poor person						***
Shelter quality	•	•	•	•	•	•
McKinney Act funding						
Availability of subsidized housing						
Subsidized housing targeted toward the very poor	*	•	•	**	•	•
Rent control						
AFDC						
GA				•		•
Spending on mental hospitals						•
Fraction of poor African-American						
Fraction of poor female						*
Average March temperature	•	**	**	**	•	*
Average March precipitation					•	•
Crime rate						

3-1 Untransformed dependent and explanatory variables, homoskedastic error term.
5-2 Untransformed dependent and explanatory variables, homoskedastic error term, street counts x 5.
5-3 Untransformed dependent and explanatory variables, homoskedastic error term, street counts x 10.
6-2 Untransformed dependent and explanatory variables, standard deviation of error term assumed proportional to poverty population.
6-3 Log dependent variable, homoskedastic error term.
6-4 Log dependent variable, appropriate explanatory variables in logs, homoskedastic error term.
Coefficients on region dummies, dummies for missing observations, and the constant term are omitted.
• Expected sign but not significant at the 10% level
* Expected sign and significant at the 10% level
** Expected sign and significant at the 5% level
*** Expected sign and significant at the 1% level

Table 8. Regressions explaining the percent of the homeless in shelters with most theoretically compelling variables

	Assumptions about street count		
	Actual street counts	Street counts x 5	Street counts x 10
Explanatory variable	Coefficient (Std. Error)	Coefficient (Std. Error)	Coefficient (Std. Error)
Shelter beds per homeless person (fitted value)	0.3723 (0.6116)	0.3046 (1.1360)	0.3609 (1.3896)
Shelter quality	0.0013 (0.0017)	0.0047* (0.0032)	0.0070** (0.0039)
McKinney Act funding	0.0277 (0.0426)	0.0700 (0.0792)	0.0843 (0.0968)
Percent poor African-American	0.0510 (0.0665)	0.0672 (0.1235)	0.0641 (0.1510)
Percent poor female	0.0660 (0.5382)	0.3868 (0.9997)	0.4396 (1.2229)
Temperature during S-Night count	-0.3675*** (0.1093)	-0.6718*** (0.2031)	-0.7930*** (0.2484)
Precipitation during S-Night count	-1.2604 (3.4515)	-4.7517 (6.4113)	-6.5104 (7.8426)
Crime rate	-0.0022 (0.0017)	-0.0086 (0.0032)	-0.0121 (0.0039)
R-squared	0.1689	0.2361	0.2508
N	222	222	222
F	4.28	5.49	5.91

The dependent variable is the (number of homeless in shelters / total number of homeless persons) X 100.
Coefficients on region dummies, dummies for missing observations, and the constant term are omitted.
* = reject hypothesis that coefficient is zero against the alternative that it has the expected sign at the 10% level of significance.
** = reject hypothesis that coefficient is zero against the alternative that it has the expected sign at the 5% level of significance.
*** = reject hypothesis that coefficient is zero against the alternative that it has the expected sign at the 1% level of significance.

Table 9. Regressions explaining the percent of the homeless in shelters with exhaustive list of variables

	Actual Street Counts	Street counts x 5	Street counts x 10
Explanatory variable	Coefficient (Std. Error)	Coefficient (Std. Error)	Coefficient (Std. Error)
Percent very poor	0.0659 (0.2538)	-0.1831 (0.4840)	-0.3629 (0.5978)
Unemployment rate	0.1980 (0.5137)	0.6686 (0.9797)	0.8571 (1.2101)
Price of rental housing service	-0.0224 (0.0135)	-0.0493 (0.0258)	-0.0638 (0.0319)
Rental vacancy rate	-0.2405 (0.3135)	-0.1109 (0.5980)	-0.0767 (0.7385)
Shelter beds per homeless person (fitted value)	-0.0336 (0.5998)	-0.8920 (1.1439)	-1.4917 (1.4128)
Shelter quality	-0.0002 (0.0022)	0.0028 (0.0041)	0.0055 (0.0051)
McKinney Act funding	0.0331 (0.0487)	0.0524 (0.0928)	0.0543 (0.1147)
Availability of subsidized housing	0.0010 (0.0017)	0.0031 (0.0032)	0.0041 (0.0040)
Subsidized housing targeted toward the very poor	0.0913 (0.1711)	0.2345 (0.3263)	0.2888 (0.4030)
Rent control	2.3497 (5.8882)	2.3398 (11.2299)	3.5378 (13.8700)
AFDC	0.0053 (0.0135)	-0.0005 (0.0258)	-0.0066 (0.0318)
GA	-0.0048 (0.0143)	0.0059 (0.0272)	0.0121 (0.0336)
Spending on mental hospitals	-0.0037 (0.0658)	0.0690 (0.1254)	0.1039 (0.1549)
Percent poor African-American	0.0365 (0.0675)	0.0567 (0.1288)	0.0561 (0.1591)
Percent poor female	0.1804 (0.6079)	0.6027 (1.1594)	0.8150 (1.4319)
Temperature during S-Night count	-0.2705** (0.1302)	-0.4697** (0.2483)	-0.5378** (0.3067)
Precipitation during S-Night count	-0.8895 (3.4945)	-3.3436 (6.6647)	-4.3517 (8.2316)
Crime rate	-0.0017 (0.0018)	-0.0063 (0.0034)	-0.0088 (0.0042)
R-squared	0.2435	0.2670	0.2671
N	222	222	222
F	2.53	3.21	3.43

The dependent variable is the (number of homeless in shelters / total number of homeless persons) X 100.
Coefficients on region dummies, dummies for missing observations, and the constant term are omitted.
* = reject hypothesis that coefficient is zero against the alternative that it has the expected sign at the 10% level of significance.
** = reject hypothesis that coefficient is zero against the alternative that it has the expected sign at the 5% level of significance.
*** = reject hypothesis that coefficient is zero against the alternative that it has the expected sign at the 1% level of significance.

Appendix A. First stage regressions

	(1)	(2)	(3)
	Coefficient (Std. Error)	Coefficient (Std. Error)	Coefficient (Std. Error)
Percent very poor	0.2959 (3.6745)		-0.0865 (0.0727)
Unemployment rate	0.3147 (9.7773)		-0.1844 (0.1932)
Price of rental housing service	0.3220 (0.2233)		-0.0036 (0.0044)
Rental vacancy rate	-3.3765 (4.8730)		0.0411 (0.0943)
Shelter quality	0.0441* (0.0265)	-0.0003 (0.0010)	0.0019*** (0.0005)
McKinney Act funding	-1.6676** (0.7294)	-0.0181 (0.0249)	0.0011 (0.0146)
Shelter beds per poor person	0.0569** (0.0238)		-0.0007 (0.0005)
Subsidized housing targeted toward the very poor	-5.3959** (2.4398)		-0.0086 (0.0493)
Rent control	-81.5119 (103.4879)		1.0445 (2.0700)
AFDC	-0.5898*** (0.1590)		-0.0164*** (0.0031)
GA	0.1921 (0.1458)		0.0143*** (0.0029)
Spending on mental hospitals	4.2731*** (0.9712)		-0.0066 (0.0193)
Percent poor African-American	-5.4379*** (1.0439)	-0.0688** (0.0305)	-0.0395** (0.0183)
Percent poor female	26.7688*** (6.6218)	0.7282*** (0.2238)	0.5617*** (0.1339)
Average March temperature	1.8205 (2.2931)		
Temperature during the S-Night count		0.0375 (0.0661)	-0.0294 (0.0392)
Average March precipitation	11.7829 (11.9292)		
Precipitation during S-Night count		-0.5444 (1.9485)	-0.2031 (1.0158)
Crime rate	-0.0139 (0.0290)	-0.0004 (0.0011)	0.0000 (0.0006)

Appendix A. First stage regressions *(continued)*

	(1) Coefficient (Std. Error)	(2) Coefficient (Std. Error)	(3) Coefficient (Std. Error)
Liberal political climate	0.3313 (0.3789)	0.0008 (0.0144)	0.0107 (0.0077)
Price of land	-0.0128 (0.0108)	-0.0006* (0.0004)	-0.0003 (0.0002)
Building cost index	-0.0270 (0.1203)	0.0077** (0.0039)	0.0032 (0.0024)
Median income	0.0167*** (0.0057)	0.0002 (0.0001)	0.0001 (0.0001)
Population of the metropolitan area	-0.0028 (0.0144)	0.0005 (0.0005)	0.0001 (0.0003)
Number of governments	0.0412 (0.1019)	-0.0051 (0.0039)	-0.0006 (0.0020)
Number of governments per capita	-30.2479*** (8.7630)	0.1756 (0.3199)	-0.8875*** (0.1774)
R-squared	0.8331	0.1542	0.8016
N	224	222	222
F	29.80	1.94	23.86

(1) First stage regression used in Table 3, Table 5, and Table 6 (regressions 1 and 2). Dependent variable is shelter beds per poor person. First stage regressions used in Table 6 (regressions 3 and 4) are not reported. They involve fewer observations and, in the case of regression 4, logarithmic transformations of many variables.
(2) First stage regression used in Table 7. Dependent variable is shelter beds per homeless person.
(3) First stage regression used in Table 8. Dependent variable is shelter beds per homeless person.
Coefficients on region dummies, dummies for missing observations, and the constant term are omitted.
* = reject hypothesis that coefficient is zero at the 10% level of significance.
** = reject hypothesis that coefficient is zero at the 5% level of significance.
*** = reject hypothesis that coefficient is zero at the 1% level of significance.