Valuing Housing Subsidies:

A Revised Method for Quantifying Benefits in a New Measure of Poverty

Sharon Stern
U.S. Census Bureau
Housing & Household Economics Statistics Division
Federal Office Building 3, Room 1472
Washington, DC 20233

sharon.m.stern@census.gov

Prepared for the Annual Conference of the
American Statistical Association

August 16, 2000

Indianapolis, Indiana

Abstract

According to the recommendations of the National Academy of Science's Panel on Poverty and Family Assistance, a new measure of poverty should include a value of noncash benefits in the estimation of family resources. This paper addresses the issue of valuing one such non-cash benefit, housing subsidies. This analysis devises a revised method for valuing subsidies using data from the American Housing Survey (AHS) and the Current Population Survey (CPS).

The present paper diverges from previous attempts to measure housing subsidies in two important ways. First is the inclusion of geographic location as a key factor for specifying differences in housing prices. Second is the use of a statistical match of families in the AHS to families in the CPS, the official source of poverty statistics, which allow values of subsidies to be allocated across the two surveys in a method more in line with a missing data approach. The analysis then compares estimated subsidies using this new method to subsidies under the current one. Finally, the paper discusses the poverty rates and the distribution of the poverty population when the housing subsidies are taken into account.

Introduction

The National Academy of Science's Panel on Poverty and Family Assistance analyzed the methods, concepts, and definitions currently used to determine the official poverty level from the March Supplement of the Current Population Survey (CPS.) In addition, the Panel investigated the possible effects of implementing their recommendations for changes to the poverty concept and the measurement methods. The Panel then released their recommendations for improvements to the official measurement of poverty in three main areas: the concept of a threshold, the definition of resources, and the adjustments necessary for geographic and family size equivalence.

In the area of resources, the Panel recommended that the value of noncash benefits be added to a family's resources for the determination of the family's poverty status. Although the CPS March Supplement does release an estimate on the monetary value of the housing subsidy, the procedure used to produce those estimates has several shortcomings. For instance, the estimates are based on 1985 American Housing Survey (AHS) data, which are then updated for inflation using the Consumer Price Index (CPI.) This paper will describe and compare several methods for valuing subsidies, including the method used to create the subsidy values currently released with the CPS March Supplement.

While the Panel did not offer a definitive alternative to the current method for estimating the value of housing subsidies, they did suggest several key elements that should be present in any such method which are not presently part of the Census Bureau's procedures. One element is an adjustment for the value of housing subsidies to reflect the local price level. Despite evidence in the real estate market that housing prices vary between housing markets, the present method distinguishes only between the four Census regions, Northeast, Midwest, South and West.

Addressing these geographic concerns in a revised method for estimating housing subsidies inevitably brings up two related matters. First, any use of geographic indicators to represent price differences in the estimation of a family's resources should be equivalent in approach to the geographic adjustments being made in the calculation of poverty thresholds. Second, because CPS is a survey of only 50,000 households, geographic codes for smaller places are not identified on the public use files. Use of a geographic adjustment in estimation of housing subsidies presents complications for issues of confidentiality.

The general approach of this work is to evaluate several different methods for estimating the value of housing subsidies on the CPS. These methods differ from the method currently used by the U.S. Census Bureau and other proposals by Naifeh and Eller in two key ways. First, places are grouped together based on housing markets rather than region or size of population. A stepwise regression estimation identified the 139 metropolitan areas in a hedonic housing cost equation which, as a group, improved the fit of the model. Second, the American Housing Survey (AHS) and the CPS subsidized households were statistically matched.

The background section will briefly review the current method and a few of the alternatives analyzed by Naifeh and Eller. The methods section will separately detail the hedonic housing cost method used to estimate the value of housing subsidies in the AHS and the methods used to transfer these estimates to families in the CPS. The evaluation section will explore the impact of these varied procedures on the distribution of the monetary value of subsidies and on the final poverty distribution. The final section will address proposals for additional analysis and conclusions.

Background

The Current Population Survey March Supplement, the source for official poverty measurement, is a nationally representative sample of households. Because this survey asks detailed income and program participation questions, it is a rich source for determining poverty under the current definition. However, experimental poverty measures which include non-cash benefits as resources cannot be determined from the CPS alone. Specifically, the value of housing subsidies integrally relies on characteristics of the housing unit which are not available on the CPS. However, the CPS does identify people living in two types of subsidized housing. The first type is public housing, in which housing units are owned and operated by local housing authorities. The second type includes privately owned housing units which are rented at a reduced cost with reimbursement of the discount to the owner from a federal, state, or local government program, such as the Section 8 program sponsored by the Department of Housing and Urban Development (HUD.)

As explained in detail below, the Census Bureau has been using the AHS, a nationally representative sample of housing units, as the source of information on the housing subsidies. Because the two surveys have different purposes and different designs, using the information from one survey in the analysis of the other includes inherent difficulties, including different geographic representation and few common variables. Despite these obstacles, the AHS is a natural choice for the process of estimating housing subsidy values due to its extensive detail of the housing characteristics. The AHS identifies subsidized housing units in a more detailed fashion than the CPS. Furthermore, the AHS renters report the amount of rent and utilities they pay. As a result, the reported characteristics of the unit can be used to predict what the market value of the unit would be if the unit were unsubsidized. And finally, the value of the subsidy can be estimated from the difference between the estimated market value and the amount paid by the renter[1].

Current Method

The current Census Bureau method uses the 1985 AHS in the first stage to estimate a value of housing subsidies. For later years, these are updated for inflation using the Consumer Price Index Residential Rent Index. The model estimated considers the following factors as independent influences on the monthly cost of a two bedroom unit: the number of bathrooms in the unit, whether the unit has three specific kitchen appliances (refrigerator, dishwasher, and garbage disposal), whether the unit has any of four specific problems (holes in the walls, holes in the floor, peeling paint, or rats), and an index of satisfaction with community services.

In this system, the monthly housing costs for unsubsidized two bedroom units are regressed on each of the four independent variables separately for each of the four regions, Northeast, Midwest, South, and West. The estimated coefficients are then applied to the characteristics of the subsidized units yielding an estimated market value of the two bedroom units. The average predicted two-bedroom monthly cost less the average two bedroom reported rent paid for each of the four regions is the average subsidy for two-bedroom units in each of the four regions.

The region-specific average subsidies are not applied directly to families in the CPS. Rather, the AHS subsidies are adjusted for number of bedrooms in the unit (more than two, two, or less than two) and family income ($10,000 or more, $6,000-$9,999, or less than $6000). The result is a 36 cell matrix of income by number of bedrooms by region. Each family in the CPS is assigned a subsidy from the table according to its own family income, number of bedrooms, and region.

[1] HUD programs subsidize the “monthly cost” of the unit, rent plus utilities.

One additional element in the current method is that the CPS does not collect information about the number of bedrooms in the unit. As a result, the current method includes a complicated method for imputing the number of bedrooms based on the composition of the primary family and the related subfamilies. The aim of this procedure is to associate a family with the number of bedrooms for which they would be eligible under some standard housing subsidy programs.

Naifeh and Eller Household Methods

Previous research by Naifeh and Eller (1997) examined the present method in detail and also made efforts to revise certain key elements. In their work, they took several new approaches to the modeling of market rents in the AHS and the method for matching AHS subsidies to the CPS including the method by which the number of bedrooms is imputed on the CPS.

The changes to the modeling of market rents included a reformulation of the equation for estimating the market value of a subsidized rental unit. This approach modeled rents, for all units, as a function of characteristics of the unit and the household and geographic identifiers. Naifeh and Eller looked at two constructs for geographic identification based on the size of the Consolidated Metropolitan Statistical Area (CMSA) and the Primary Metropolitan Statistical Area (PMSA). Since PMSAs are areas within CMSAs, Naifeh and Eller define the size of the area in which a household resides by both. In one construct they use only 17 categories of possible sizes of the two areas. In the second construct, metropolitan areas are assigned to groups by size and by region, resulting in 48 metropolitan categories.

They also investigated a third approach to assigning the AHS subsidies to the families in the CPS. They estimated a model in which the AHS housing subsidies are the dependent variable determined by the following independent variables: the number of bedrooms for which a family was eligible, family income, family income squared, and the metropolitan area size categories.

The coefficients from this model were applied to the families in the CPS to predict the relevant subsidy amount.

FMR Approach

Finally, Naifeh and Eller compare the results using these experimental methods to a fair market rent approach. This process benefits from the extensive work already conducted by the Department of Housing and Urban Development (HUD). Each year HUD publishes a list of fair market rents (FMRs), which are essentially estimates of the 40th percentile rent in the relevant local housing market. These rents are set for the purpose of administering the Section 8 Housing Assistance Payments Program.[2] Consequently, it is a natural extension to use these rents to estimate the value of the housing assistance in the CPS by subtracting 30 percent of a family's income from the appropriately chosen FMR. A brief summary of how FMRs are determined demonstrates their usefulness for the purposes described here.

Base market rents are calculated separately for each bedroom size category. To determine these with statistical reliability for every geographically unique housing market, HUD starts with the most recent Decennial (1990) Census data. For the largest metropolitan areas, HUD updates the rents intercensally using the AHS. For other FMR areas, HUD uses random digit dialing telephone surveys in conjuncture with trending factors based on the CPI to update the rents with statistical accuracy. The result of all this work is FMRs for every metropolitan area and every non-metropolitan area county in the United States updated yearly.

The appeal of this approach is three-fold. First, the methods used to establish these rents are consistent with standard housing economic and statistical principles. Second, HUD has

[2] Families in the Section 8 Certificate program pay 30 percent of their adjusted income in rent. The housing authority pays the remainder: rent charged less the amount paid by the tenant. The maximum allowable rent for a unit of a given size is the fair market rent. Details on HUD programs are available on the web site, www.hud.gov.

published these rents for all geographic areas, eliminating the problem of small sample size when using the AHS national sample alone. Third, since the local housing authorities administering Section 8 and other programs use the FMRs to set the amounts for vouchers, then any effort to estimate the value of vouchers would do well to use the same source of market rent information. As a result, this paper uses the FMR technique as a comparison to the methods below.

Method of Hedonic Housing Cost to Estimate Subsidy

The first step in revising the method for determining the value of housing subsidies in the CPS is to examine the model used to generate estimates of housing subsidy value in the AHS. The general approach in the housing literature for predicting rents is to use a hedonic housing price equation for the relevant housing market. In general, a hedonic housing price equation is one in which the cost of a unit is described as a function of the characteristics of the housing unit. The estimated coefficients on the characteristics are interpreted as the marginal cost of that characteristic. For instance, if the equation is in semi-log form, a coefficient of 0.12 on the number of bedrooms in the unit means that if the unit had 1 more bedroom, the monthly cost would be 12% higher.

While the entire USA is by no means a single housing market, it is difficult to decide exactly how small the geographic area must be to constitute a single housing market. The sample size in the AHS is not large enough to treat each metropolitan area with a population of 100,000 or more separately. Furthermore, some large housing markets might have only a small sample and some smaller housing markets are not represented by any of their own housing units.[3] Therefore, using a single hedonic price equation for all units in the United States assumes that all the marginal prices for housing characteristics are constant across geographic units and that the relevant differences between those geographic units can be adequately represented by a constant difference in price. The technique described below, therefore, relies on a single geographic identifier to express the price differences between a given metropolitan area the rest of the country. Since the main concern being addressed is the observed differences in prices across housing markets, this assumption is critical to the process used in this analysis.

The standard hedonic price equation used in the housing literature is a regression of housing expenditures on the housing characteristics which are the explanatory variables. These characteristics generally take several forms including characteristics of the structure and the neighborhood. Many researchers have used this type of equation to predict rents or measure implicit prices for certain characteristics.[4]

The hedonic housing price equation has a set of explanatory variables which includes as many characteristics as are available. Researchers are often in a difficult position between choosing between a smaller geographic area with less information on housing and a broader geographic area with more information on housing. For example, the AHS National Sample includes many measures of structural and location characteristics but only identifies a broad geographic area. Despite the many housing characteristics available, the sample size prohibits a complete housing price model estimated at the MSA level. Alternatively, some researchers use

[3] It is for this reason that many researchers use the Decennial Census for analyzing variation in housing cost. The two main drawbacks to using the Decennial Census are the limited number of questions about housing characteristics and the frequency of data.

the Decennial Census as the data source, since it provides enough data to estimate a housing equation at the MSA level. In exchange for more observations at the MSA level, however, researchers such as Malpezzi, Chun, and Green compute the hedonic housing equation using a reduced set of explanatory variables. Importantly, Malpezzi et al. point out that a simpler specification does not invalidate the method if the goal is using the predicted rents. These will not be severely compromised by a smaller set of variables.

Following the standard methods outlined by Malpezzi et al. and others, the housing price equation in the present model is in semi-log form with the characteristics of the housing unit as the independent variables and the log of rent paid as the dependent variable. The set of structural characteristics included in the model are as follows:

BATHS	Number of bathrooms in the unit.
BEDRMS	Number of bedrooms in the unit.
Set of dummy variables for the number of rooms in the unit:	
RM1	The unit has one room.
RM2	The unit has two rooms.
RM3	The unit has three rooms.
RM4	The unit has five rooms.
RM5	The unit has six or more rooms.
Omitted:	Units with four rooms.
Age of dwelling:	
DWELLAGE	The age the building.
AGESQ	The age of the building squared
Set of dummy variables for type of dwelling:	
ONEUNITD	The dwelling is a single family unit, detached structure.
ONEUNITA	The dwelling is a single family unit, attached structure.
MOBILE	The dwelling is a mobile home.
Omitted:	Dwellings with multiple units.
Amenities of the unit:	
DECK	The unit has a porch or deck.
AIRCOND	The unit has air conditioning.
FIREPL	The unit contains a fireplace.

[4] For an example of creating housing indexes from predicted rents, see "New Place-to-Place Housing Price Indexes for the U.S. metropolitan Areas and their Determinants" by Stephen Malpezzi, Gregory Chun, and Richard Green 1998.

CARPORT	The unit has a garage or carport.
ALLAPPL	The unit has three major appliances: refrigerator, garbage disposal, dish washer.

Since the aim was to identify which locations had a significant impact on housing prices, a set of geographic area dummy variables was created for the analysis. In order to determine which geographic areas were significant in explaining housing market variation, the equation was estimated repeatedly using a stepwise procedure. At the outset, all 236 metropolitan statistical areas in the US were possibilities for inclusion as an indicator of differences in the housing market. One geographic variable at a time was added to the model which included all the structural characteristics until no additional improvements to R-squared could be made. Further, the geographic variables were tested jointly and independently with F-tests to assure that the set of geographic indices was significant.

Once the subset of 148 metropolitan areas was chosen, the model was estimated in its final version. The coefficients on the geographic dummy variables are the incremental difference in the rental price between a unit in a particular MSA and a unit that is either not in an MSA or is in an MSA whose rental price is not significantly different from non-MSA units. Appendix A details the results of the first stage estimation equation, providing a full listing of the independent variables used in the final model.

The market value of subsidized rental units was calculated using the estimated coefficients for non-subsidized rental units. The difference between the predicted monthly cost and the reported amount paid is the value of the housing unit's subsidy. Table 1 shows some summary statistics about the weighted distribution of housing subsidies in the AHS.

Table 1. Statistics on Subsidies in AHS Households	
Monthly Subsidies:	
Mean	$238.5
Maximum	1274.5
75th percentile	405.6
Median	207.6
25th percentile	0
Minimum	0
Aggregate yearly subsidy (in 000s)	16,700,256

Methods of Assigning Subsidies from AHS to CPS

This research explores three different methods for applying the AHS subsidies to the CPS. All the methods that use the AHS as the source for the subsidies assign the subsidy to a household in the CPS. Since poverty is a family measure, in the final analysis each household subsidy needs to be scaled down to the size of the family or families within the household.

The first two methods are variations of the subsidy value table approach currently in use. After calculating an average subsidy for a relevant group of characteristics from households in the AHS, these averages are applied to the households in the CPS with the same set of characteristics. In the simpler approach, Metropolitan Statistical Area (MSA) was the only dimension used in the subsidy value table. Since only 148 of the MSAs in the AHS were significant in the housing equation, the subsidy value table has only 149 cells. In the CPS, a household in one of the 148 identified MSAs received a subsidy equal to the average household subsidy in the same MSA in the AHS. The remainder of the CPS households were treated as non-MSAs and were assigned the average household subsidy of the non-MSA households in the AHS.

In the more complicated variation, households in each MSA were grouped further by the number of people in the household as one person, two people, three people, four people, or five or more people. Since the correlation between the number of bedrooms and the number of people in the household for subsidized households is 0.685 in the AHS, this approach accounts for general geographic differences in prices and a proxy for the size of the unit. Unfortunately, the data is rather sparse at this degree of detail. In cases where the CPS household had no match on number of people and MSA, the MSA overall subsidy average was used instead.

The third method tested in this research is a statistical match of households in the CPS to households in the AHS. Unlike an exact match of a particular unit across two data sources, in a statistical match, each record from one data source is matched with a record from a second data source, where the matched record represents a similar unit. The general procedure for a statistical match of any sort is to first identify a cohort variable and then to define a distance function. The cohort variable (or variables) is used to stratify the two data sources such that matches can only occur within a single strata. The distance function defines the "similarity" between all the possible matches. Similarity can be defined with continuous or categorical variables. With continuous variables, the distance function can be defined as the degree of difference or the relative difference. With a categorical variable, a positive value is assigned as the distance if a match is not made.

As an example, matching each person on data source A with a person from data source B, the most similar record on source B might be defined as the person closest in age to the person on file A. The distance could be defined as the absolute value of the difference in ages: $distance(a,b)=|age_a-age_b|$, where age_a is the age of the person on file A and age_b is the age of a person on file B. The person record on file B with the minimum distance to the person record on file A is the selected match. In this example, file B is the donor data set, meaning it contains information that is needed for analysis of data set A, but not available on A. File A is the recipient data set.

In most cases, the distance function will be defined not by a single variable, but by many variables. In that case, the distance function must include a weighting scheme which indicates which match characteristics are more important. For example, a match between two data sources of the number of bedrooms may include the number of people in the household and the number of children in the household. The distance between a record on the recipient data set and each record on the donor dataset could be calculated using the following weighted distance formula: $distance(a,b)=weight_{people}*|(\text{number of people})_a-(\text{number of people})_b| + weight_{children}*|(\text{number of children})_a-(\text{number of children})_b|$. The researcher must evaluate the appropriate weighting scheme to achieve the best distribution of bedrooms on the recipient data set. In many cases, the weights that generate a bedroom distribution on the recipient data set closest to the distribution on the donor data set is the best.

The statistical match strategy used in this work is unconstrained. An unconstrained match allows a particular record in the donor data to be used as many times as it is chosen as minimizing the distance function. Since this can (and probably will) result in some records on the donor data set being used many times and other records not being used, it requires some finesse to choose a match that best represents the distribution of the variables being moved from the donor data set.

The object of the statistical match used in this research is to match each subsidized renter in the CPS (the recipient file) to a subsidized renter in the AHS (the donor file). Once an acceptable match is completed, the estimated value of the housing subsidy can be used as a resource in a new measure of poverty. The match between these two data sets was done on a household record level, the unit of measurement for which the AHS has the best information.

As mentioned above, the purpose of the cohort variable is to group together records that should be matched only to records in the same group in the other data set. In this analysis, the fair market rents were used to construct a cohort variable. Since the FMRs are chosen as the 40^{th} percentile of the rental price distribution for an apartment with a specific number of bedrooms, they serve in this analysis as a proxy for general rental price level in a specific area. Using cluster analysis on the FMR data, the resulting 14 distinct groups are called CLUSTER in the following analysis. Since the FMR data includes a rental amount for every location either MSA or county in the US, clustering the areas into groups by FMR grouped together locations with similar rental prices. Table 2 reports summary statistics on the clusters created from the FMR data.

In the matching procedure, certain observations are sometimes used in matches more frequently than others. In the CLUSTERs numbered in Table 2 as 1, 2, and 3, it looks as if the analysis would have plenty of sample such that overuse of a single record would not occur. However, since the FMRs are based on county level prices, many of the areas which make up these clusters are not represented in either the AHS or the CPS sample. Therefore, a second formulation of the clusters was tested in the analysis, CLUST2. This alternate cluster variable, referenced in the evaluation below, collapses groups 1, 2, and 3 from CLUSTER, which are of similar price level but have smaller sample sizes in the CPS and AHS. This leaves a larger group with mean of 390.33 and standard deviation of 25.5.

Table 2, Summary Statistics on Clusters

CLUSTER	Number of FMR Areas in Cluster	Mean Two Bedroom FMR ($s)	Standard Deviation
*1	548	357.81	5.90
*2	490	385.90	6.79

Table 2, Summary Statistics on Clusters			
CLUSTER	Number of FMR Areas in Cluster	Mean Two Bedroom FMR ($s)	Standard Deviation
*3	798	415.38	8.79
4	289	463.58	13.47
5	183	500.25	10.82
6	101	547.94	12.42
7	72	592.32	11.70
8	53	636.36	13.88
9	34	680.41	13.69
10	33	722.42	12.62
11	16	763.50	14.93
12	10	820.90	9.92
13	22	903.50	43.13
14	13	1,113.62	71.21

* These three groups were combined into one in variable CLUST2.

The variables tested for the distance function include: the number of people in the household, the number of children in the household, household's MSA, state, marital status of the householder, senior citizen status of householder, race of householder, and the sex of householder. The first two variables were entered into the distance function as the degree of difference. This means that the weight increases by the absolute value of the difference between the value on the donor and recipient files.

The match was run 74 times using different subsets and weights on each of these variables. As an example of the differences, Table 3 reports the parameters for the match with the highest aggregate and mean estimated subsidy as well as the match with the lowest aggregate and mean subsidy.

Table 3, Parameters from Two Statistical Matches		
Distance Variables	Weight - Low Match	Weight - High Match
Number of people in household	10	2
Number of children in household	7	1
MSA[5]	1	5

[5] Only the MSAs which were specifically identified in the first stage estimation equation are used as possible indicators in the statistical match.

Table 3, Parameters from Two Statistical Matches		
Whether or not householder >64	0	3
Note: The Low Match used CLUST2 as the cohort variable and the High Match used CLUSTER.		

Finally, this paper also includes subsidies calculated with the fair market rent method. The use of FMR was implemented in the same general manner as in the Naifeh and Eller paper. Each CPS family was assigned the appropriate FMR based on the family's number of bedrooms[6]. The subsidy is equal to the FMR less thirty percent of the family's income. For reasons described in the background section, the FMR method is a solid alternative to the other methods and should continue to be examined as an option.

Evaluation of Methods

In order to evaluate the methods of transfer between the AHS and the CPS, Table 4 reports the distributions of the household level subsidies under the different household level transfer methods detailed above.

Table 4, Statistics on Household Level Subsidies in the CPS

	Distribution in AHS	Average by MSA	Average by MSA and Household Size	Low Match	High Match
Monthly Subsidies: Mean	*$238.5*	$240.5	$242.1	$241.6	$261.4
Maximum	*1274.5*	1043.7	1043.7	1209.9	1274.5
75th percentile	*405.6*	283.8	289.2	401.4	415.9
Median	*207.6*	208.6	236.2	215.8	228.2
25th percentile	*0*	208.6	178.3	14.0	41.6
Minimum	*0*	30.4	0	0	0
Aggregate yearly subsidy (in 000s)	*16,700,256*	13,876,440	13,967,916	13,939,980	15,083,088

[6] Since the CPS respondents are not asked about the number of bedrooms in the unit, post-CPS processing includes a technique for estimating the number of bedrooms based on family, not household, composition. For more information on this process, see Naifeh and Eller. In their work, they explain the current method and produce and test several alternatives to this technique.

The two methods that assign averages based on specific characteristics do a reasonable job of replicating the AHS mean subsidy on the CPS. The average by MSA and average by MSA and household size methods result in mean subsidies on the CPS of 240.5 and 242.1 respectively. These are quite close to the mean subsidy on the AHS, 238.5. Conversely, they do not bring in any of the depth or richness of the information on the AHS. The standard deviation of the AHS subsidies is 228.2, much larger than in the standard deviation of the subsidies in the CPS using these two methods. Assigning CPS households the average subsidies by MSA only has a standard deviation of 80.0. Assigning by MSA and the household size in the household has standard deviation of 111.7.

Table 4 also reports that both of the statistical matches give reasonable weighted distributions of subsidies on the CPS compared to the original distributions on the AHS. For example, the aggregate amount of annual housing subsidies on the AHS is $16.7 billion compared with $15.1 billion on the CPS with the high match and $13.9 billion on the CPS with the low match. The minimum, maximum, mean, median, 25th percentile, and the 75th percentile estimated subsidies are all similar between the AHS and the two sample statistical matches.

To compare the FMR method to the others, Table 5 reports some characteristics of the distributions at the family level. The household level subsidies from the first methods are scaled down by the number of people in the family relative to the number of people in the household. The term families includes primary families, unrelated subfamilies, and unrelated individuals.

Each person in a household is included in a family unit either alone or in combination with others in the household to whom the person is related.[7]

Table 5, Statistics on Family Level Subsidies in the CPS					
	FMR Method	Average by MSA	Average by MSA and Household Size	Low Match	High Match
Monthly Subsidies:					
Mean	$349.0	$228.1	$229.6	$229.2	$247.8
Maximum	1662.9	1043.7	1043.7	1156.4	1274.5
75th percentile	515.6	274.7	279.9	384.9	401.4
Median	322.9	208.6	229.8	196.7	212.2
25th percentile	166.0	208.6	178.3	12.7	32.7
Minimum	0	0	0	0	0
Aggregate yearly subsidy (in 000s)	21,241,872	13,886,580	13,977,864	13,949,076	15,081,924

This table clearly shows that the fair market rent method estimates much larger subsidy values than any of the other methods. In work not yet completed, these raw distributions of FMRs applied to the CPS will be compared to the raw distribution resulting from applying the FMR method back to the AHS where the number of bedrooms is actually reported rather than imputed as in the CPS. It will be interesting to see if the amount of subsidies estimated that way is equally high.

To isolate the effects of adding subsidies as resources to poverty measurement, poverty rates were calculated using the official definition of poverty with one modification. The monetary equivalent for the housing subsidy under each alternative method was added to the family's resources.

[7] Note that the universe of Table 5 is limited to families in the poverty universe. Given how subsidies are scaled to the family size, this will result in some of the subsidy value being lost. As a example, a household includes a family of 4 plus an unrelated individual under the age of 15. The family of 4 is assigned four-fifths of the value of the housing subsidy. The unrelated individual is assigned one-fifth of the housing subsidy. In the calculation of the distribution of household level subsidies in Table 4, this sample household contributed one subsidy which is the sum of the two family unit level subsidies. However, the poverty universe does not include unrelated individuals under 15 years old. Therefore, the distribution of subsidies in the poverty universe will include only the scaled subsidy from the family of four in the sample household.

One criticism levied against these different methods is that the value of the housing subsidy should not exceed the amount of money in the poverty threshold which is presumed to be housing expenses. As a result, one final change was made to the estimated subsidies before poverty rates were computed: subsidies were capped at 44.3% of the relevant poverty threshold. Table 6 shows the impact of capping the subsidies on the distribution of family level subsidies.

Table 6, Statistics on Subsidies in the CPS Capped at 44.3% of Poverty Threshold					
	FMR Method	Average by MSA	Average by MSA and Household Size	Low Match	High Match
Monthly Subsidies: Mean	$285.5	$222.2	$222.7	$203.0	$215.5
Maximum	1209.9	625.6	754.0	832.2	804.0
75th percentile	396.1	274.7	279.9	313.1	313.1
Median	288.6	208.6	229.8	196.7	212.2
25th percentile	166.0	208.6	178.3	12.7	32.7
Minimum	0	0	0	0	0
Aggregate yearly subsidy (in 000s)	17,376,924	13,523,808	13,555,632	12,355,788	13,117,416

Table 7 demonstrates how these different methods can affect the poverty rate, when the rate is calculated using a single modification to the official definition, the addition of subsidies to resources.

Table 7, Comparing 1998* Poverty Rates		
	Percent in Poverty	Number in Poverty (in 000)
Official Poverty Definition	12.7%	34,476
Add value of subsidy to income:		
Current Method	12.2	33,158
Average by MSA	12.1	32,683
Average by MSA and number of people	12.1	32,683
Low Match	12.1	32,831
High Match	12.1	32,826
FMR Method	11.8	31,935
*Using March 1999 CPS, 1999 AHS, 1999 FMRs		

While all the different methods lower poverty rates, they do not appear to have a differential impact on the various segments of the population. Table 8 gives poverty rates for selected groups of people. These rates show that certain portions of the population, such as seniors and people living in families with a female household with no spouse present, will experience slightly lower poverty rates under the modified poverty definition. These differences are consistent with the overall lower poverty rates under the alternative definition shown in Table 7.

Table 8, 1998 Poverty Rates for Specific Population Groups				
	Seniors	Northeast	West	Female Householder*
Official Poverty Definition	10.5%	12.3%	14.0%	33.1%
Add value of subsidy to income:				
Current Method	9.2	11.3	13.6	32.2
Average by MSA	9.2	11.0	13.5	31.5
Average by MSA and # of people	9.3	11.0	13.5	31.3
Low Match	9.5	11.3	13.5	31.4
High Match	9.4	11.3	13.6	31.3
FMR Method	9.1	10.6	13.1	30.1
*People living in families with a female householder, no spouse present.				

The method used to evaluate subsidies does not seem to have an impact on the characteristics of people with income below the poverty level. For example, under the current definition of poverty, 39.1% of the people in poverty are children. With the value of subsidies added to the resources of the family, the percent varies between 39.2% and 39.7%. Table 9 gives some characteristics of the people in poverty.

Table 9, Characteristics of People in Poverty				
	% Children	% Seniors	%Female	%South
Official Poverty Definition	39.1%	9.8%	57.3%	37.7%
Add value of subsidy to income:				
Current Method	39.7	9.0	57.1	38.1
Average by MSA	39.6	9.1	57.1	38.2
Average by MSA and # of people	39.4	9.2	57.1	38.3
Low Match	39.2	9.4	57.1	38.0
High Match	39.3	9.3	56.9	38.0
FMR Method	39.3	9.3	56.8	38.8

Conclusions

The methods covered in this paper have distinct impact on the distribution of housing subsidies as they appear on the CPS. The statistical matching methods do a better job of recreating the distribution of housing subsidies as they were originally estimated on the AHS. In addition, all of these new methods improve upon the current method by giving a more accurate distribution of housing subsidies. The only caveat is that the FMR method resulted in consistently higher subsidy estimated until capped at 44.3% of the relevant family threshold. With that change in place, the FMR method was more consistent with the other methods.

At the same time, for the purposes of defining the effect of including housing subsidies as a resource in the measurement of poverty, the various methods were similar. All four new methods based on the AHS had poverty rates of 12.1%. As expected, this is lower than the official poverty rate for 1998, 12.7%. The FMR method did result in a slightly lower poverty rate of 11.8%. In addition, the characteristics of the poor were remarkably similar regardless of the method which produced the value of housing subsidies.

Some additional work, currently in the planning phase, could help verify some of the conclusions reached here. The following are examples.

- Apply the FMR method back to the AHS data. Compare the resulting distribution of subsidies to the hedonic price method. Investigate the implications of this comparison on each methodology separately.
- Use a statistical match to assign each household on the CPS with a number of bedrooms from the AHS. Apply the FMR method using that number of bedrooms. How close is this distribution of subsidies to the other methods.
- Consider using an average approach where the characteristics are the cluster, the number of people in the household, and the number of children in the household.

- Look at the statistical match procedure for viability in a production environment. Examine the degree of difference in the results if only 50 MSAs are identified explicitly in the procedure rather than the full list of significant places.

Appendix A - Hedonic Housing Price Estimation in the American Housing Survey

Dependent Variable: Log of monthly cost, rent plus utilities.
Model R-square: 0.3823

Variable	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob>|T|
Intercept	5.8732	0.0267	220.2	0.0001
Number of bathrooms in the unit	0.1335	0.0106	12.65	0.0001
Number of bedrooms in the unit	0.0562	0.0087	6.5	0.0001
Age of Dwelling	-0.0047	0.0007	-6.72	0.0001
Age of Dwelling Squared	0	0	4.1	0.0001
Unit is a single detached unit	0.0342	0.0114	3	0.0027
Unit is a single attached unit	0.0178	0.0124	1.43	0.153
Unit is a mobile home	-0.1715	0.0219	-7.82	0.0001
Unit has a deck or porch	0.0093	0.0086	1.09	0.2779
Unit has an air conditioner	0.0985	0.0101	9.79	0.0001
Unit has a fireplace	0.0836	0.0122	6.86	0.0001
Unit has a carport or garage	0.0773	0.009	8.57	0.0001
Unit has 3 major appliances (refrigerator, dish washer, garbage disposal)	0.1239	0.0105	11.84	0.0001
Unit has 1 room	-0.2215	0.0366	-6.06	0.0001
Unit has 2 rooms	-0.1328	0.0244	-5.45	0.0001
Unit has 3 rooms	-0.0691	0.0119	-5.81	0.0001
Unit has 5 rooms	0.0523	0.0111	4.73	0.0001
Unit has 6 or more rooms	0.115	0.0163	7.07	0.0001
Albany-Schenectady-Troy NY MSA	0.2158	0.0608	3.55	0.0004
Allentown-Bethlehem PA-NJ MSA	0.2776	0.0595	4.67	0.0001
Altoona PA MSA	-0.1791	0.1022	-1.75	0.0798
Anchorage AK MSA	0.4021	0.0679	5.92	0.0001
Atlanta GA MSA	0.2655	0.0368	7.22	0.0001
Atlantic City NJ MSA	0.5868	0.1224	4.79	0.0001
Austin TX MSA	0.3289	0.0418	7.87	0.0001
Baltimore MD MSA	0.2312	0.0378	6.12	0.0001
Bergen-Passaic NJ PMSA	0.7036	0.0519	13.55	0.0001
Binghamton NY MSA	0.2124	0.1186	1.79	0.0733
Boston MA PMSA	0.6416	0.0318	20.21	0.0001
Boulder-Longmont CO PMSA	0.3827	0.1032	3.71	0.0002
Bridgeport-Milford CT PMSA	0.6012	0.1098	5.47	0.0001
Brockton MA PMSA	0.4877	0.1218	4	0.0001
Brownsville-Harlingen TX MSA	-0.1564	0.0955	-1.64	0.1014
Buffalo NY PMSA	0.1709	0.0531	3.22	0.0013
Charleston SC MSA	0.1507	0.0839	1.8	0.0725
Charlotte-Gastonia-Rock Hill NC-SC MSA	0.0715	0.0427	1.68	0.0938
Chattanooga TN-GA MSA	-0.2144	0.0985	-2.18	0.0296
Chicago IL PMSA	0.4546	0.0233	19.54	0.0001
Cincinnati OH-KY-IN PMSA	0.1682	0.0496	3.39	0.0007
Cleveland OH PMSA	0.1725	0.0431	4.01	0.0001

Variable	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob>|T|
Colorado Springs CO MSA	0.3391	0.0735	4.62	0.0001
Columbia SC MSA	0.1658	0.0898	1.85	0.0649
Columbus GA-AL MSA	-0.2564	0.0785	-3.27	0.0011
Columbus OH MSA	0.177	0.0437	4.05	0.0001
Corpus Christi TX MSA	0.2014	0.0884	2.28	0.0227
Dallas TX PMSA	0.1584	0.0286	5.54	0.0001
Daytona Beach FL MSA	0.368	0.0956	3.85	0.0001
Dayton-Springfield OH MSA	0.1179	0.06	1.97	0.0492
Denver CO PMSA	0.3508	0.0444	7.9	0.0001
Des Moines IA MSA	0.1725	0.0999	1.73	0.0843
Detroit MI PMSA	0.2093	0.0315	6.64	0.0001
Dutchess County NY PMSA	0.5235	0.1072	4.89	0.0001
Eugene-Springfield OR MSA	0.3167	0.0825	3.84	0.0001
Fayetteville NC MSA	0.1874	0.0837	2.24	0.0252
Fort Collins-Loveland CO MSA	0.3599	0.0961	3.75	0.0002
Fort Lauderdale-Hollywood-Pompano Beach FL PMSA	0.3339	0.0517	6.46	0.0001
Fort Myers-Cape Coral FL MSA	0.2107	0.1202	1.75	0.0797
Fort Pierce FL MSA	0.2585	0.0995	2.6	0.0094
Fort Worth-Arlington TX PMSA	0.1385	0.0414	3.35	0.0008
Gary-Hammond IN PMSA	0.1603	0.0831	1.93	0.0539
Glens Falls NY MSA	0.2169	0.1105	1.96	0.0497
Grand Rapids MI MSA	0.1809	0.0625	2.9	0.0038
Greensboro--Winston-Salem--High Point NC MSA	0.1487	0.0554	2.69	0.0073
Hagerstown MD MSA	0.2054	0.0791	2.6	0.0095
Harrisburg-Lebanon-Carlisle PA MSA	0.1843	0.0737	2.5	0.0125
Hartford CT PMSA	0.4122	0.0579	7.11	0.0001
Honolulu HI MSA	0.5555	0.0707	7.86	0.0001
Houma-Thibodaux LA MSA	0.3877	0.1717	2.26	0.024
Houston TX PMSA	0.0626	0.0271	2.31	0.0208
Indianapolis IN MSA	0.1707	0.0456	3.74	0.0002
Jacksonville FL MSA	0.1509	0.06	2.52	0.0119
Jersey City NJ PMSA	0.6343	0.0503	12.62	0.0001
Johnson City-Kingsport-Bristol TN-VA MSA	-0.1615	0.1019	-1.58	0.1133
Kalamazoo MI MSA	0.2245	0.0791	2.84	0.0046
Kankakee IL MSA	0.2182	0.0927	2.35	0.0186
Kansas City MO-KS MSA	0.1031	0.043	2.4	0.0165
Knoxville TN MSA	0.14	0.0711	1.97	0.049
Lafayette LA MSA	-0.1926	0.1173	-1.64	0.1007
Lancaster PA MSA	0.5316	0.1071	4.97	0.0001
Lansing-East Lansing MI MSA	0.2765	0.1078	2.57	0.0103
Las Vegas NV MSA	0.1924	0.045	4.28	0.0001
Lawrence KS MSA	0.1152	0.0613	1.88	0.0604
Lawrence-Haverhill MA-NH PMSA	0.5923	0.112	5.29	0.0001
Lexington-Fayette KY MSA	0.1051	0.0584	1.8	0.0721
Little Rock-North Little Rock AR MSA	0.1747	0.0662	2.64	0.0084

Variable	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob>|T|
Los Angeles-Long Beach CA PMSA	0.4016	0.0174	23.14	0.0001
Louisville KY-IN MSA	0.1122	0.0618	1.81	0.0697
Lowell MA-NH PMSA	0.5235	0.1307	4.01	0.0001
Madison WI MSA	0.4054	0.0757	5.36	0.0001
Manchester NH MSA	0.4667	0.1508	3.1	0.002
Mansfield OH MSA	-0.3567	0.1145	-3.12	0.0018
Miami-Hialeah FL PMSA	0.334	0.034	9.83	0.0001
Middlesex-Somerset-Hunterdon NJ PMSA	0.6781	0.0682	9.94	0.0001
Milwaukee WI PMSA	0.2615	0.0462	5.66	0.0001
Minneapolis-St. Paul MN-WI MSA	0.4232	0.0386	10.96	0.0001
Monmouth-Ocean NJ PMSA	0.5412	0.0707	7.66	0.0001
Myrtle Beach SC MSA	-0.1856	0.1005	-1.85	0.0647
Naples FL MSA	0.2773	0.1178	2.36	0.0186
Nashua NH PMSA	0.3579	0.1456	2.46	0.014
Nashville TN MSA	0.1676	0.0501	3.34	0.0008
Nassau-Suffolk NY PMSA	0.7132	0.0478	14.93	0.0001
New Bedford MA MSA	0.5536	0.1391	3.98	0.0001
New Haven-Meriden CT MSA	0.467	0.0986	4.74	0.0001
New London-Norwich CT-RI MSA	0.4764	0.1413	3.37	0.0008
New York NY PMSA	0.6429	0.0174	36.98	0.0001
Newark NJ PMSA	0.5778	0.0406	14.25	0.0001
Newburgh NY-PA PMSA	0.2636	0.071	3.71	0.0002
Norfolk-Virginia Beach-Newport News VA MSA	0.1249	0.0442	2.83	0.0047
Oakland CA PMSA	0.5829	0.0346	16.84	0.0001
Olympia WA MSA	0.2289	0.0808	2.84	0.0046
Omaha NE-IA MSA	0.1423	0.064	2.23	0.0261
Orange County CA PMSA	0.5281	0.032	16.48	0.0001
Orlando FL MSA	0.2369	0.0474	5	0.0001
Philadelphia PA-NJ PMSA	0.3664	0.0284	12.9	0.0001
Phoenix AZ MSA	0.1579	0.0365	4.33	0.0001
Pittsburgh PA PMSA	0.2158	0.0459	4.7	0.0001
Pittsfield MA MSA	0.7701	0.2053	3.75	0.0002
Portland ME MSA	0.4334	0.0921	4.71	0.0001
Portland OR PMSA	0.278	0.0382	7.28	0.0001
Portsmouth-Dover-Rochester NH-ME MSA	0.4409	0.0894	4.93	0.0001
Providence RI PMSA	0.3237	0.0487	6.65	0.0001
Racine WI PMSA	0.2589	0.1065	2.43	0.0151
Raleigh-Durham NC MSA	0.1997	0.0497	4.02	0.0001
Reading PA MSA	0.2055	0.0846	2.43	0.0152
Riverside-San Bernardino CA PMSA	0.1483	0.0357	4.16	0.0001
Roanoke VA MSA	0.178	0.0921	1.93	0.0533
Rochester NY MSA	0.3567	0.0532	6.71	0.0001
Rockford IL MSA	0.4832	0.1357	3.56	0.0004
Sacramento CA MSA	0.1743	0.042	4.15	0.0001
Saginaw-Bay City-Midland MI MSA	0.3155	0.1249	2.53	0.0115

Variable	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob>|T|
Salem OR MSA	0.1544	0.0765	2.02	0.0437
Salinas-Seaside-Monterey CA MSA	0.5004	0.0749	6.68	0.0001
Salt Lake City-Ogden UT MSA	0.1847	0.0647	2.86	0.0043
San Diego CA MSA	0.4845	0.0303	15.99	0.0001
San Francisco CA PMSA	0.7937	0.036	22.04	0.0001
San Jose CA PMSA	0.7718	0.0457	16.89	0.0001
Santa Barbara-Santa Maria-Lompoc CA MSA	0.3763	0.0885	4.25	0.0001
Santa Cruz CA PMSA	0.7318	0.0871	8.41	0.0001
Santa Rosa-Petaluma CA PMSA	0.4954	0.1199	4.13	0.0001
Sarasota FL MSA	0.2771	0.0863	3.21	0.0013
Savannah GA MSA	0.2064	0.079	2.61	0.009
Scranton--Wilkes-Barre PA MSA	0.153	0.0785	1.95	0.0512
Seattle WA PMSA	0.4222	0.0332	12.71	0.0001
Sioux Falls SD MSA	0.1669	0.0896	1.86	0.0625
Springfield MA MSA	0.3118	0.091	3.43	0.0006
Springfield MO MSA	-0.4118	0.0937	-4.39	0.0001
St. Louis MO-IL MSA	0.1524	0.0383	3.98	0.0001
Stamford CT PMSA	0.811	0.1026	7.91	0.0001
Steubenville-Weirton OH-WV MSA	-0.2312	0.1112	-2.08	0.0375
Syracuse NY MSA	0.3677	0.082	4.48	0.0001
Tacoma WA PMSA	0.2278	0.0598	3.81	0.0001
Tampa-St. Petersburg-Clearwater FL MSA	0.2354	0.0429	5.49	0.0001
Trenton NJ PMSA	0.5141	0.1102	4.67	0.0001
Tucson AZ MSA	0.1209	0.0633	1.91	0.0561
Tuscaloosa AL MSA	-0.128	0.0673	-1.9	0.0573
Utica-Rome NY MSA	0.1945	0.1229	1.58	0.1136
Vallejo-Fairfield-Napa CA PMSA	0.2543	0.0746	3.41	0.0007
Ventura CA PMSA	0.5074	0.0658	7.72	0.0001
Vineland-Millville-Bridgeton NJ PMSA	0.3488	0.1547	2.26	0.0241
Washington DC-MD-VA MSA	0.474	0.027	17.55	0.0001
Waterbury CT MSA	0.4082	0.116	3.52	0.0004
WestPalm Beach-Boca Raton-Delray Beach FL MSA	0.2433	0.0693	3.51	0.0004
Wilmington DE-NJ-MD PMSA	0.3951	0.0932	4.24	0.0001
Worcester MA MSA	0.4573	0.0993	4.6	0.0001
Yolo, CA MSA	0.4456	0.1258	3.54	0.0004
Yuma AZ MSA	-0.5922	0.239	-2.48	0.0132

References

Citro, Connie F. and Robert T. Michael (ed.), **Measuring Poverty: A New Approach**, Washington, DC; National Academy Press, 1995.

Malpezzi, Stephen, Gregory H. Chun and Richard K. Green. 1998. “New Place-to-Place Housing Price Indexes for U.S. Metropolitan Areans, and Their Determinants,” *Real Estate Economics*, 26(2):235-274.

Naifeh, Mary and T.J. Eller, "Valuing Housing Subsidies in a New Measure of Poverty Using the Current Population Survey," Poverty Measurement Working Paper, U.S. Census Bureau, 1997.