Where We Live

Geographic Differences in Poverty Thresholds

Kathleen Short
U.S. Census Bureau
January 2001

One of the most troublesome concepts of poverty measurement is that of geographic cost-of-living differences. The Measure of Poverty: A Report to Congress as Mandated by the Education Amendments of 1974. Washington DC, GPO, April 1976.

Background

The current official poverty measure provides individuals and policy makers with information about the types of families and people who are not making ends meet. One important dimension along which this information is provided has to do with where people live. It is quite obvious to all that consider it that the cost of living is different depending upon where you live. But here we find one of the more contentious issues in measuring poverty, how to accurately portray the difficulty of making ends meet when the cost of basic necessities varies widely for people living in different places.

The current official poverty measure makes no adjustments for geographic area. It presents the cost of getting along as uniform across the United States. Poverty rates by state, for example, as shown in the first figure, are calculated using the same poverty thresholds regardless of where you live. In her book on measuring poverty, Patricia Ruggles (1990) wrote that, "Other than differences in family size and type, the major source of possible differences in need that analysts have considered as a basis for adjustment are differences in place of residence." Ruggles provided evidence that a strong case can be made for adjustment of the poverty thresholds for price differences by geographic area. She argued that, even though we cannot make precise adjustments for very small areas, we quite possibly would make fewer errors if reasonable adjustments were made for larger geographic areas. The challenge then is, as Ruggles stated, "the analyst must balance the possibility that a given adjustment will distort needs estimates in some cases against the potential error that may results if no adjustment is done." In her view, "…the evidence for real differences in price levels across regions has become too compelling to ignore."

I. Education Amendments of 1974

Over the last thirty years, since the inception of an official poverty measure, the issue of geographic differences in cost of living has often been raised. Under the Education Amendments of 1974, the Department of Education was required to conduct a thorough study of the use of the measure of poverty in Title I of the Elementary and Secondary Education Act of 1965. More specifically it stated that the measure of poverty should be adjusted for regional, climatic, metropolitan, urban suburban and rural

differences. It was recommended that the study should consider the availability of data more current than the decennial census, the availability and usefulness of cost-of-living data, the usefulness of cost-of-housing data, labor market and job availability data, data on wages, unemployment rates and income distribution, and data with respect of eligibility criteria for welfare programs. It was further stipulated that if data were not found to be available then the study should include an explanation of the reasons therefor.

An interagency group of government researchers undertook this research effort. They examined Bureau of Labor Statistics family budgets series, state standards of need for the AFDC program, the 1960 and 1970 Censuses of Housing, the Current Population Survey, and a series of other federal surveys. Their conclusion at that time was that there was no known way to make satisfactory geographic adjustments to the poverty cutoffs.

II. GAO Report 1995

In 1995, a report from the General Accounting Office was prepared at the request of Congress concerning the statistical data requirements that would be needed to construct a Cost of Living index that could be used to adjust for geographic difference in living costs. This report specifically describes the function of market baskets in determining an index, identified methodologies that might have potential for calculating an index, and finally surveyed 'experts' on how best to make these adjustments. The conclusions from this work were exceedingly pessimistic, suggesting that constructing a market basket is difficult, consensus problematic, and expert opinion on methodologies conflicting and unresolved.

The GAO report identified twelve different methodologies that could be used. Fifteen experts were then asked to provide an assessment of the strengths and weaknesses of each of the methods. The 12 methods were budgets, norms, housing data, family budgets, consumption data, interarea price indexes, economic modeling, consumer price indexes, estimation models, local indexes, polling, and comparable pay data. Of these methods, the expert GAO panel regarded five as having little or no promise: local indexes, polling, family budgets, consumption data, and the consumer price index methodologies. The comparable pay method was found to be not promising at all.

The housing data methodology was among the methods that were considered to be strong in that housing costs were agreed to be a large source of variation in the cost of living. Specifically the report referred to Fair Market Rents (FMR) -- data from the 1990 decennial census, analyzed with a methodology developed by the Department of Housing and Urban Development (HUD) to estimate rents for comparable units in different localities. One important criticism of housing data was that the quality of housing should be controlled in the measure to prevent downward bias in low-income areas. Two other methods got somewhat favorable responses from the reviewers; budgets (basic needs budget, Renwick and Bergman) and norms (Runzheimer plan of living cost standards).

III. National Academy of Sciences Report 1995

In the spring of 1995 the National Academy of Sciences (NAS) released a report recommending that the official poverty measure be revised. Their report listed a series of recommendations, one of which was the inclusion of an adjustment for geographic differences in the poverty threshold. This panel of experts stated that variations in housing costs can be significant across areas of the country and housing expenditures are a large component of the poverty budget. As a first and partial step to account for cost-of-living differences among geographic areas, the NAS panel recommended that the housing component of poverty thresholds be adjusted for geographic differences in the cost of housing. The panel developed a set of indexes that were used to adjust poverty thresholds for metropolitan areas that fall within several population size groups and non-metropolitan areas in each of the nine Census Bureau divisions of the country. The panel stated that, while these indexes contained inaccuracies, they were a marked improvement over the current measure, which makes no adjustment at all for geographic differences in costs of basic needs. (The panel also urged additional research on these issues.)

The NAS panel implemented a modified version of the HUD FMR approach with 1990 census data to determine whether inter-area housing cost index values could be developed that were reasonably similar to findings in the scientific literature. The panel concluded that their approach represented a modest step in a necessary direction but that limitations remain.

The NAS indexes were based on 1990 census data on rents for two-bedroom apartments that had plumbing facilities, kitchen facilities, and electricity and in which the occupant had moved into within the last 5

years. These data are the same as those used by HUD to develop FMRs. Metropolitan areas were then grouped by 6 population size categories and non-metropolitan areas by the 9 census defined divisions and the indexes were computed using the cost of housing at the 45th percentile of the value of the distribution for each area. This resulted in a set of indexes for 41 geographic areas. Table 1 shows indexes calculated for each of the nine census divisions. Note that this method assumes that housing prices are comparable within each of the divisions.

Their procedure takes account of geographic differences in housing costs, but not differences in other costs, and even for housing costs it assigns index values to people in some areas that are in error. Because of limitations of the available data, the panel recommended additional research to determine a method for updating measures of housing costs more frequently than every ten years using decennial census data. In order to account for the fact that the index values represented only the cost of housing, on average 44 percent of the cost of the basic bundle, they produced a fixed weight index for application to the thresholds which included other goods; food, clothing, and a little bit more.

Housing costs vary widely and often within small distances of one another. In order to most accurately capture housing cost differences, therefore, one should have data in very fine detail. The NAS panel started the calculation of indexes with information from the census on housing costs for the entire U.S. by county. They were forced to a smaller group of 341 metro areas and non-metropolitan areas for each state. They finally were reduced again to the 41 areas described above.

Table 1. NAS Adjustments for Housing Costs for Proposed Poverty Thresholds*	
Area and Population Size	**Index Value**
Northeast	
New England (Connecticut, Maine, Massachusetts,	
New Hampshire, Rhode Island, Vermont)	
Non-metropolitan areas and Metropolitan areas under 250,000	1.128
Metropolitan areas 250,000-500,000	1.128
Metropolitan areas 500,000-1,000,000	1.148
Metropolitan areas 1,000,000-2,500,000	1.141
Metropolitan areas 2,500,000 or more	1.209
Middle Atlantic (New Jersey, New York, Pennsylvania)	
Non-metropolitan areas and Metropolitan areas under 250,000	0.908
Metropolitan areas 250,000-500,000	0.997
Metropolitan areas 500,000-1,000,000	1.020
Metropolitan areas 1,000,000-2,500,000	0.975

Metropolitan areas 2,500,000 or more	1.187
Midwest	
East North Central (Illinois, Indiana, Michigan,	
Ohio, Wisconsin)	
Non-metropolitan areas and Metropolitan areas under 250,000	0.896
Metropolitan areas 250,000-500,000	0.959
Metropolitan areas 500,000-1,000,000	0.987
Metropolitan areas 1,000,000-2,500,000	0.995
Metropolitan areas 2,500,000 or more	1.059
West North Central (Iowa Kansas, Minnesota, Missouri, Nebraska,	
North Dakota, South Dakota)	
Non-metropolitan areas and Metropolitan areas under 250,000	0.861
Metropolitan areas 250,000-500,000	0.962
Metropolitan areas 500,000-1,000,000	0.981
Metropolitan areas 1,000,000-2,500,000	1.028
Metropolitan areas 2,500,000 or more (use areas 1-2.5 million)	N.A.
South	
South Atlantic (Delaware, District of Columbia, Florida, Georgia,	
Maryland, North Carolina, South Carolina, Virginia, West Virginia)	
Non-metropolitan areas and Metropolitan areas under 250,000	0.899
Metropolitan areas 250,000-500,000	0.961
Metropolitan areas 500,000-1,000,000	1.007
Metropolitan areas 1,000,000-2,500,000	1.043
Metropolitan areas 2,500,000 or more	1.119
East South Central (Alabama, Kentucky, Mississippi, Tennessee)	
Non-metropolitan areas and Metropolitan areas under 250,000	0.827
Metropolitan areas 250,000-500,000	0.935
Metropolitan areas 500,000-1,000,000	0.947
Metropolitan areas 1,000,000-2,500,000	N.A.
Metropolitan areas 2,500,000 or more	N.A.
West South Central (Arkansas, Louisiana, Oklahoma, Texas)	
Non-metropolitan areas and Metropolitan areas under 250,000	0.858
Metropolitan Areas 250,000-500,000	0.911
Metropolitan areas 500,000-1,000,000	0.942
Metropolitan areas 1,000,000-2,500,000	0.962
Metropolitan areas 2,500,000 or more	1.005
West	
Mountain (Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, Utah, Wyoming)	
Non-metropolitan areas and Metropolitan areas under 250,000	0.888
Metropolitan areas 250,000-500,000	0.976
Metropolitan areas 500,000-1,000,000	1.039
Metropolitan areas 1,000,000-2,500,000	1.003
Metropolitan areas 2,500,000 or more	N.A.
Pacific (Alaska, California, Hawaii, Oregon, Washington)	
Non-metropolitan areas and Metropolitan areas under 250,000	0.969
Metropolitan areas 250,000-500,000	1.018
Metropolitan areas 500,000-1,000,000	1.028
Metropolitan areas 1,000,000-2,500,000	1.104
Metropolitan areas 2,500,000 or more	1.217

*Table copied from Table 1-1, page 41: Connie F. Citro and Robert T. Michael (eds.)
Measuring Poverty: A New Approach, Washington, D. C.: National Academy Press, 1995.

One important reason to collapse housing cost information into fewer groups is to simplify the poverty thresholds. More important, though, publicly available microdata files typically do not contain the necessary information to use housing values in every small area because of confidentiality concerns, limiting disclosure of geography to areas of at least 100,000 population. This precludes the application of cost indexes at very detailed levels when using public microdata files such as the Current Population Survey (CPS) if we wish to make poverty thresholds available to the public.

Our task, then, is to group housing costs of small geographic areas into a relatively small number of groups based solely on housing costs, as represented by some meaningful and consistent variable or group of variables, without distorting the representation of housing cost for any given area of analytical interest. This leads us to the question of determining the area of analytical interest.

The Census Bureau releases official poverty statistics every year for states, as shown in table 2 for 1992. These averages are calculated using poverty rates for 1990, 1991, and 1992. This is the lowest level of geography that are currently published as official poverty statistics, and even so, they are calculated using a 3-year average. This is done to reduce the variance of the estimates and increase reliability. While state-level poverty rates are the smallest geographic area statistics reported, they are of great interest. State-level poverty rates are reported and analyzed widely and, thus, the focus of this paper is on the effect of geographic indexes on state-level poverty estimates. Our goal, then, is to look for geographic adjustments that most accurately estimate poverty rates by state and protect confidentiality concerns for release of public data.

IV. Census Bureau report on experimental measures
The Census Bureau's report on experimental poverty measures (Short et al. 1999) presented results from implementing the panel's procedure for adjusting thresholds for geographic differences in the costs of housing using the current official thresholds as well as all other recommendations of the panel. That report presented the effects that resulted from each recommendation separately as well as incorporating all recommendations taken together. To demonstrate the effect of adjusting thresholds for cost of living

differences by area, the NAS indexes were used with the official poverty thresholds. Money income was used to compare to these thresholds to determine poverty status. This approach yielded a poverty rate of 13.0 percent in 1997, slightly lower than the official rate of 13.3 percent.

Table 2: Three-Year Average Poverty Rates for 1992: Official versus NAS Poverty Measures			
	Official	**NAS**	**Difference**
Alabama	18.40	15.82	-2.58
Arizona	14.50	14.73	0.23
Arkansas	18.10	15.14	-2.96
California	15.10	18.94	3.84
Colorado	11.60	11.10	-0.50
Connecticut	8.00	7.86	-0.14
Delaware	7.30	9.28	1.98
District of Columbia	20.00	22.96	2.96
Florida	15.00	15.50	0.50
Georgia	16.90	15.53	-1.37
Idaho	14.60	12.15	-2.45
Illinois	14.20	13.98	-0.22
Indiana	13.50	13.08	-0.42
Iowa	10.40	9.28	-1.12
Kansas	11.20	11.02	-0.18
Kentucky	18.60	15.83	-2.77
Louisiana	22.30	20.07	-2.23
Maine	13.50	16.58	3.08
Maryland	10.20	11.08	0.88
Massachusetts	10.60	12.56	1.96
Michigan	14.00	13.01	-0.99
Minnesota	12.60	11.13	-1.47
Mississippi	24.60	18.98	-5.62
Missouri	14.60	13.34	-1.26
Montana	15.10	12.91	-2.19
Nebraska	10.00	9.60	-0.40
Nevada	11.90	12.82	0.92
New Hampshire	7.40	9.07	1.67
New Jersey	9.60	12.08	2.48
New Mexico	21.40	17.86	-3.54
New York	15.00	16.32	1.32
North Carolina	14.40	13.27	-1.13
North Dakota	13.40	11.79	-1.61
Ohio	12.40	11.90	-0.50
Oklahoma	17.00	14.97	-2.03
Oregon	11.30	12.28	0.98
Pennsylvania	11.20	11.79	0.59
Rhode Island	10.00	11.48	1.48
South Carolina	17.20	16.16	-1.04
South Dakota	14.00	12.11	-1.89
Tennessee	16.50	14.18	-2.32
Texas	17.10	15.09	-2.01
Utah	10.10	9.74	-0.36
Vermont	11.30	13.11	1.81
Virginia	10.10	10.77	0.67

Washington	9.80	10.24	0.44
West Virginia	19.40	16.91	-2.49
Wisconsin	10.00	9.23	-0.77
Wyoming	10.40	8.84	-1.56

While overall rates are quite close, however, the important effect of the geographic adjustments is that there are fewer poor in areas of relatively low living costs and more in areas of relatively high living cost. As one might expect, poverty rates estimated with these geographic adjustments are lower in areas with lower housing costs, such as the South and the Midwest, and higher in the Northeast and the West.

While the NAS report and the Census Bureau's work in this area suggest the importance of adjusting poverty thresholds for geographic differences in cost of living, there are some particular problems with their approach. Over and above the fact that their indexes were developed from a relatively old set of data, the procedure that they used to calculate the indexes had some undesirable results. This is particularly apparent when one calculates poverty rates by state.

First, 1992 poverty rates by state were calculated using the NAS indexes. The resulting differences in state poverty rates from the official estimates are shown in table 2. These estimates are based on 3-year averages of the poverty rates, both official and with geographic indexes. The table shows the percentage point change in poverty rates when we apply the geographic adjustment.

As can be seen in the table, which shows the difference between the official and the experimental poverty rates, some resulting differences in poverty rates are reasonable and expected, such as higher poverty rates in California where housing costs are higher than generally in the nation. Increased poverty rates in most of the New England states and the western states is expected. Decreases in the south and states such as North Dakota and South Dakota are also not surprising.

But there are also some surprising results. The large increase in poverty rates for Maine where housing costs are somewhat lower than average, with a small decline in Connecticut, where costs are quite high, are not expected results. The explanation for these unexpected results stems from the method the NAS panel used to calculate their indexes. The main problem being that indexes were calculated based on geographic

location rather than on housing costs. So for example, all metro areas in the new England division were given the same index value, even though there is quite a lot of variation in housing costs in this area. So given standards of living in the New England states as represented in reported incomes and earnings, the poverty thresholds in Maine were increased too much while perhaps they were not increased enough in Connecticut to adequately reflect differences in cost of living.

A final important issue that arises is that of controlling for differences in housing quality in the development of housing cost indexes. For example, a housing unit that rents for $200 per month in Mississippi may not be similar in amenities to a housing unit that rents for $200 per month in New York City. Therefore, it is important to attempt to control for differences in housing quality when constructing a housing price index. Malpezzi points out that there is a well-developed literature on housing price index construction and that hedonic price indexes are a generally accepted, easily understood method of constructing housing indexes adjusted for quality differences.

In sum, there are four main criticisms of the NAS indexes as they are presented in the 1995 NAS report and subsequently used by the Census Bureau. First, the indexes are based on outdated information. Second, the indexes are grouped in a way that yields unexplained results for some geographic areas. Third, there is no control for housing quality across geographic areas. Finally, geographic indexes, including the NAS indexes, present confidentiality problems for release of microdata files. We address these criticisms partially in the work described below.

V. Addressing the problems

1. Data sources and other indexes

Before presenting their indexes, the panel provided a comprehensive discussion of the available data for this purpose. Their discussion is similar to that in the GAO report. One of the most promising methods that was discussed in both publications included hedonic models, particularly Bureau of Labor Statistics (BLS) interarea indexes for the 30 largest metropolitan areas and some smaller ones (Kokoski, Cardiff, and Moulton). These indexes are most promising but are currently not available for other areas of the U.S. In their research they showed that the component of spending with the largest variation was shelter and utilities. Together these two components account for 33 percent of the Consumer Price Index (CPI) market

basket. These results supported the notion that adjusting for housing costs alone in poverty thresholds would represent a significant improvement for measuring poverty.

Since the interarea indexes from BLS are not available for all areas they cannot be used to adjust thresholds. However, HUD fair market rents are estimated annually for 2416 counties that are outside metro areas and for all 341 metro areas. They are defined to be gross rent with utilities at the 45th percentile for the rent distribution of a standard quality of rental housing. There are some problems with FMRs. One of those being that they do not fully control for housing quality as the definition of 'standard' is not a strict one. And only recent movers are included in the calculations, while long-term renters may pay lower rents on average. So they are biased upward. There are also some legislative adjustments that present some inconsistencies across areas.

Analysts at HUD have described reasons why FMRS should not be used to adjust a poverty threshold[1]. They list the following reasons;

    1. FMRs are only developed for use in section 8 certificate and voucher program
    2. They measure rents not total costs
    3. They use gross rents of recent movers
    4. Only major metro areas are checked using Random Digit Dialing surveys
    5. Rental markets are volatile
    6. For 99 large areas, rents are adjusted using CPI rent and utility factors. While only available for 32 Consolidated Metropolitan Statistical Areas (CMSAs), they are applied to all Primary Metropolitan Statistical Areas (PMSAs) within the CMSAs
    7. There are updates of rents for small areas with Random Digit Dialing procedures that may result in generalizations of rent changes not applicable to individual areas
    8. The percentile standard is not consistent over time (the 50th percentile from 1975 to 1983, the 45th percentile from 1985 to 1994, and the 40th percentile starting in 1995).
    9. The percentile measure is administratively determined and not based on measurement criteria
    10. The treatment of non-metropolitan counties has changed over time
    11. In 1996 a state minimum FMR was instituted.

FMRs do have problems, as noted by both HUD analysts and the NAS panel. However, the FMR methodology is straightforward and applicable to all areas of the U.S. The data are also updated on a regular basis from the base year of 1990. The information is very detailed and provides us with a good look at the variability of housing costs in different areas of the country.

So, for the purpose of this paper we begin with FMRs for 1999 as the 'best' data available at this time. Mean FMRs indexes by state are shown in table 3. The FMR for a given PMSA or non-metropolitan county

is attached to the CPS file and used to construct an index by which a portion of the poverty threshold is adjusted to account for shelter costs. In the first case FMRs are used directly for each metropolitan area or county in nonmetro areas and applied to each family in the CPS sample. The indexes are then normalized so that the average is equal to one across people on the CPS file. The figure shows mean indexes by state calculated in this way.

Table 3: Mean normalized indexes by state						
	NAS	FMR	FMR clustered	Malpezzi clustered	Malpezzi collapsed	FMR state/msa
Alabama	0.86	0.84	0.84	0.85	0.85	0.84
Alaska	0.95	1.14	1.13	1.09	1.07	1.13
Arizona	0.96	0.99	0.98	0.96	0.96	0.99
Arkansas	0.85	0.84	0.84	0.85	0.83	0.84
California	1.14	1.12	1.13	1.17	1.18	1.13
Colorado	0.95	1.02	1.03	0.97	0.95	1.02
Connecticut	1.15	1.14	1.13	1.19	1.19	1.13
Delaware	1.02	1.02	1.03	1.02	1.02	1.02
District of Columbia	1.09	1.17	1.17	1.35	1.23	1.17
Florida	0.98	1.01	1.01	1.01	1.00	1.01
Georgia	1.00	0.96	0.96	0.97	0.98	0.95
Hawaii	0.99	1.23	1.25	1.17	1.19	1.23
Idaho	0.89	0.88	0.88	0.86	0.93	0.88
Illinois	1.00	1.04	1.03	1.03	1.00	1.03
Indiana	0.93	0.90	0.90	0.89	0.89	0.90
Iowa	0.87	0.87	0.87	0.88	0.90	0.87
Kansas	0.90	0.88	0.88	0.90	0.93	0.88
Kentucky	0.86	0.85	0.84	0.87	0.86	0.85
Louisiana	0.89	0.86	0.86	0.90	0.91	0.87
Maine	1.10	0.94	0.94	0.95	0.95	0.94
Maryland	1.07	1.05	1.06	1.09	1.09	1.05
Massachusetts	1.16	1.15	1.15	1.14	1.17	1.16
Michigan	0.99	0.98	0.98	0.96	0.96	0.97
Minnesota	0.92	0.97	0.98	0.96	0.98	0.97
Mississippi	0.84	0.83	0.83	0.89	0.84	0.83
Missouri	0.94	0.87	0.87	0.90	0.90	0.87
Montana	0.87	0.88	0.89	0.84	0.84	0.88
Nebraska	0.89	0.89	0.89	0.88	0.88	0.89
Nevada	0.98	1.06	1.05	0.95	0.93	1.06
New Hampshire	1.15	1.07	1.07	1.11	1.11	1.07
New Jersey	1.16	1.19	1.18	1.21	1.20	1.18
New Mexico	0.92	0.91	0.91	0.88	0.88	0.92
New York	1.09	1.17	1.17	1.15	1.14	1.16
North Carolina	0.95	0.92	0.92	0.93	0.92	0.92
North Dakota	0.84	0.85	0.85	0.82	0.81	0.85
Ohio	0.98	0.93	0.93	0.92	0.89	0.92
Oklahoma	0.88	0.85	0.84	0.87	0.87	0.85
Oregon	1.02	0.98	0.97	0.94	0.94	0.98
Pennsylvania	1.02	0.98	0.98	0.98	1.02	0.97

1 Personal communication with Alan Fox at HUD.

Rhode Island	1.11	1.05	1.06	1.04	1.07	1.05
South Carolina	0.94	0.89	0.88	0.90	0.93	0.89
South Dakota	0.84	0.89	0.89	0.86	0.86	0.89
Tennessee	0.89	0.88	0.89	0.87	0.88	0.89
Texas	0.92	0.95	0.95	0.91	0.92	0.95
Utah	0.95	0.98	0.98	0.89	0.90	0.98
Vermont	1.10	0.99	0.99	1.00	0.98	0.99
Virginia	0.99	1.01	1.02	1.03	1.03	1.00
Washington	1.10	1.01	1.01	0.96	0.99	1.01
West Virginia	0.91	0.85	0.85	0.83	0.81	0.86
Wisconsin	0.92	0.92	0.92	0.91	0.90	0.92
Wyoming	0.87	0.88	0.87	0.85	0.84	0.88

These indexes can be compared to the NAS indexes. They show some slight differences from the NAS indexes shown above. Most notable, however, are the higher values in the New England states under the NAS measure. There are also higher index values for Oregon, Kansas, and West Virginia, South Carolina, Indiana, and Pennsylvania. One further note, the NAS indexes tend to follow the boundaries of the Census division with little variation within those boundaries. The FMR indexes show more variation within census divisions.

2. Grouping the data

In order to address the problems with confidentiality described above it is necessary to follow the NAS panel in attempting to reduce the detail of indexes while not disturbing the estimate of poverty rates by state. The main purpose of this method is to investigate the effect on poverty rates of using indexes that are grouped by housing costs only, rather than by geographic division.

Cluster analysis is used to group small geographic areas by housing costs. Rather than basing the groups on geographic location, an index should be based on the exact items that it is intended to represent. This method uses various mathematical formulae to put together a large number of items into clusters that are alike. Cluster analysis puts data into groups, called clusters, in order to meet designated objectives of a particular application, such as to minimize the variation between members of each cluster while maximizing the difference between groups, all within certain constraints (such as a fixed number of clusters).

Using FMRs for two-bedroom units and performing a cluster analysis specifying, arbitrarily, the outcome of 20 separate groups or clusters, we divide the metro and county areas based on the two-bedroom unit FMR. A further step combines several groups that have only a few members with the nearest group in value of housing costs. This results in 15 clusters. Using the means of each of the final set of clusters, divided by the grand mean across all geographic areas, an index is computed. The index values are constained to apply to only 44 percent of the poverty threshold. The mean indexes by state are shown below. One noticeable result is that they are not different from the nonclustered FMRs. The appropriate index is then applied to the poverty threshold of each household depending on their location and recomputed poverty rates.

3. Housing Quality: Malpezzi, Chun, and Green Indexes

While the FMRs offer us more updated and more detailed housing price indexes than the NAS measures, one thing that is left unaddressed is the problem that in the construction of these indexes there is little control for difference in housing quality.

To examine this issue we turn to work by Malpezzi, Chun and Green (1998). In this work the authors present new place-to-place indexes of the flow price of rental housing and the stock or asset price of owner-occupied housing in 272 United States metropolitan areas. They also develop, though not presented in the article cited, comparable indexes for non-metropolitan portions of each state.

Their indexes are computed by applying hedonic methods to decennial census data for 1990, which has the advantage of adjusting for constant housing quality. Using census data the authors have estimated a set of hedonic regression equations that include a limited set of information available from the decennial census including, structural, neighborhood characteristics, location within the market and contract conditions or characteristics such as utilities included in rent. These indexes are shown in table 3.

In order to see what differences we perceive when we use indexes that control for housing quality, a set of housing price indexes is created from rents prepared from this work. Using the rent indexes, as they are most comparable to the NAS and FMR indexes above, I calculate an index for each metropolitan area and remainder of state for nonmetro areas, by dividing the rents by the grand mean. As in the calculations above the housing price index is applied to the housing portion of the threshold assuming shelter represents

44 percent of the basic bundle. After applying the fixed weight indexes to the thresholds, the indexes are normalized to the CPS sample by dividing by the mean of the indexes across people in the poverty universe.

Mean indexes by state are calculated and presented in the table for comparison to earlier indexes. When this is done we see that the Malpezzi indexes are very similar to the FMR indexes, except that they are lower for a few states: Washington, Nevada, Utah, Colorado, and New Mexico. Comparing to NAS indexes, both being based in 1990 census data, suggests that controlling for housing quality and/or clustering in a more appropriate manner, yields lower housing costs for Washington state, Oregon, Utah, New Mexico, South Carolina, Indiana, West Virginia, and Kansas, Pennsylvania, Vermont, and Maine. The Malpezzi indexes yield higher costs for Illinois, Virginia, and Florida. Generally we see more within-division differences between the two indexes and more differences between this index and the NAS index than we saw between this index and the FMR index. The differences observed suggest that the NAS indexes differ due to the inappropriate grouping rather than difference in housing quality controls.

4. Confidentiality Issues

One important step that is taken next in this exercise is to examine more carefully any breach of confidentiality that we might inadvertently commit by linking these indexes with the microdata. To do this, all of the indexes are examined against geographic information already available on the Census Bureau's public use files. As a quick check, I cross-tabulate state by metropolitan status with each of the indexes to see if additional information is provided to this cross-classification that might allow identification. In fact, all of the indexes, when cross tabulated in this way yield very small cells with weighted counts below 100,000 persons. This represents a rough guide to judge the possibility of identifying respondents in a microdata file.

As an illustrative exercise, indexes are further collapsed until this particular cross-tabulation yields no additional geographic detail below certain levels. Beginning with the Malpezzi indexes, already collapsed into 15 cluster groups, we can examine which particular geographic areas are revealing too much information. By doing this on a case by case basis, further aggregation across index values can be done.

Examining what steps are taken, and the resulting set of indexes, gives us information by which to assess whether or not the degree of collapsing required for confidentiality purposes results in indexes that actually improve our measure of poverty over no adjustment at all. In other words, does this exercise so obscure the detail of housing cost differences that the adjustments result in poverty rates that are more or less uninformed on the issues of geographic differences in housing costs? And if so, is there really a benefit to making these adjustments?

Here are the steps taken to allow possible disclosure of poverty thresholds adjusted for geographic differences based on the Malpezzi indexes.

- collapse 15 index categories to 5 for all states
- for Colorado, Illinois, Kentucky, Minnesota, Missouri, and Ohio further collapse the lowest two categories together, yielding 4 categories for these states
- for Massachusetts and New Jersey collapse the three lowest categories into one, yielding 3 categories
- For Delaware, Idaho, Maine, Maryland, Mississippi, Montana, Nebraska, New Hampshire, New Mexico, Oklahoma, and South Dakota, set the index for the entire state equal to the mean index as computed with the clustered indexes. These states have only one category.

Note that, while the number of categories have been reduced for each state in this exercise severely, overall, including the addition of mean values by state, there remains a set of 10 index values that may be assigned to adjust poverty thresholds.

As before the indexes are recomputed by normalizing to the sample mean and calculated mean indexes by state to compare with previous indexes. Mean indexes by state calculated in this way are shown. Comparing first with the non-collapsed Malpezzi indexes we see an increase for Idaho, Kansas, and Louisiana, and lower for Wisconsin and South Carolina. We also see lower values for Hawaii and Alaska, not shown here in the figures but listed in the table. Overall, however, the indexes are not that different as a result of our collapsing.

As a final exercise housing indexes are collapsed in a slightly different way. This time starting with FMRs, the average FMR is computed for each state, and within each state, by metropolitan or nonmetropolitan

area. This computation yields 100 separate indexes and state means are presented in the final index map shown here. Not surprisingly the indexes, also shown in table 3, are very similar to those calculated using FMRs directly.

VI. Application of indexes to CPS poverty estimates by state

Our final exercise is to compare the resulting poverty rates by state from applying these indexes to poverty thresholds, and comparing the results. I depart from usual practice here by only comparing single estimates for poverty rates by state, rather than 3-year averages as the Census Bureau normally publishes.

We begin comparing official poverty rates by state for 1997with the poverty rates estimated by using the NAS indexes. Comparing poverty rates based on unadjusted versus NAS adjusted thresholds suggests some important problems with unadjusted thresholds that are true regardless of which index you use. See table 4 for all poverty rates.

Table 4: Poverty rates by state 1997						
			--------------------------------------Experimental Rates* -----------------------			
	Official	NAS	FMR	FMR clustered	Malpezzi clustered	Malpezzi collapsed	FMR state/msa
U.S.	**13.3**	**13.0**	**12.9**	**12.9**	**12.9**	**13.0**	**13.0**
Alabama	15.7	12.5	11.6	11.3	12.1	12.2	11.3
Alaska	8.8	8.4	10.3	10.3	9.4	9.2	10.5
Arizona	17.2	16.3	17.1	17.1	16.4	16.3	17.0
Arkansas	19.7	14.9	14.2	14.2	14.9	14.5	14.3
California	16.6	19.1	18.4	18.3	19.3	19.6	19.0
Colorado	8.2	7.9	8.3	8.6	7.9	7.9	8.3
Connecticut	8.6	9.9	9.6	9.5	11.4	11.7	10.0
Delaware	9.6	9.0	9.7	9.7	9.6	9.6	9.7
District of Columbia	21.8	22.7	23.7	23.7	26.6	25.4	23.7
Florida	14.3	13.9	14.1	14.1	14.1	14.0	14.3
Georgia	14.5	14.8	13.4	13.6	14.6	14.6	13.4
Hawaii	13.9	13.5	16.6	16.5	15.3	15.3	16.3
Idaho	14.7	11.5	10.8	10.8	10.3	13.1	11.0
Illinois	11.2	11.4	12.2	12.1	12.1	11.6	12.0
Indiana	8.8	7.1	6.9	6.9	6.3	6.5	6.5
Iowa	9.6	7.4	7.4	7.3	7.4	7.9	7.3
Kansas	9.7	8.4	7.5	7.5	8.0	8.6	7.4
Kentucky	15.9	11.9	11.8	11.9	12.4	13.1	11.8
Louisiana	16.3	14.6	14.1	14.1	15.2	15.2	14.3
Maine	10.1	12.5	9.9	9.9	9.6	9.5	9.7
Maryland	8.4	8.5	8.5	8.6	8.9	8.5	8.4
Massachusetts	12.2	14.3	13.7	13.7	13.6	14.1	14.3
Michigan	10.3	10.5	10.2	10.1	10.1	9.7	9.9

Minnesota	9.6	8.2	9.0	9.0	9.0	9.0	8.8
Mississippi	16.7	12.1	11.6	14.7	13.5	11.3	11.6
Missouri	11.8	10.7	10.1	10.0	10.1	10.9	10.3
Montana	15.6	13.5	13.9	14.0	12.9	12.8	14.1
Nebraska	9.8	9.1	9.0	9.0	8.9	9.1	9.1
Nevada	11.0	10.4	12.2	12.4	10.0	10.1	12.2
New Hampshire	9.1	11.1	9.8	9.8	11.0	10.9	9.9
New Jersey	9.3	10.9	11.0	11.0	11.3	11.3	11.2
New Mexico	21.2	19.3	18.7	18.5	18.2	18.8	19.2
New York	16.5	18.0	19.2	19.1	18.7	18.7	19.2
North Carolina	11.4	10.6	10.1	10.1	10.4	10.3	10.3
North Dakota	13.6	10.3	10.8	10.7	10.1	9.5	10.7
Ohio	11.0	10.1	9.8	9.8	9.6	9.4	9.7
Oklahoma	13.7	11.1	9.6	9.6	10.9	10.3	9.9
Oregon	11.6	12.2	11.4	11.1	10.7	10.9	11.4
Pennsylvania	11.2	11.1	10.6	10.5	10.8	11.0	10.5
Rhode Island	12.7	15.3	14.3	14.3	14.1	14.1	14.3
South Carolina	13.1	11.8	10.9	10.9	11.0	11.6	10.3
South Dakota	16.5	14.2	14.7	14.6	14.3	14.2	14.9
Tennessee	14.3	11.6	10.3	11.1	11.0	11.5	10.7
Texas	16.7	15.0	15.6	15.6	14.7	14.8	15.8
Utah	8.9	7.7	8.0	8.0	7.4	7.6	8.4
Vermont	9.3	10.4	9.0	9.0	9.1	9.1	9.1
Virginia	12.7	12.3	12.3	12.1	12.4	12.5	12.8
Washington	9.2	9.7	8.6	8.6	8.4	8.6	8.9
West Virginia	16.4	13.9	12.6	12.6	12.7	12.1	12.2
Wisconsin	8.2	6.7	7.0	7.0	7.0	6.8	6.6
Wyoming	13.5	11.3	11.8	11.8	11.3	11.4	11.9

* Experimental rates vary only by geographic adjustments to poverty thresholds.

Namely, that the unadjusted official rates are too high for states in the south, such as Alabama, Arkansas, Mississippi, and in the Midwest, such as Idaho, Indiana, and Iowa. On the other hand it appears that the official rates are too low for states in the northeast, such as Massachusetts and Rhode Island. Unlike the other indexes shown later, however, the NAS indexes result in similarly low rates for Alaska and Hawaii as the official rates and uniquely higher rates for Maine, New Hampshire and Rhode Island. We attribute this to the inappropriate groupings used in the construction of the NAS indexes.

Poverty rates based on FMRs, either clustered or not, suggest further that the official unadjusted rates are misreporting the poverty rates of states relative to one another. Again we see the pattern that, compared to official rates, poverty is lower in the south and higher in some New England states, but also higher in such states as New York, California, Alaska and Hawaii, then under the official measure.

The next calculations show poverty rates by state using the Malpezzi measures, both collapsed and uncollapsed. Here we see that compared again with official poverty rates we show lower rates for the south, considerably higher rates for some New England states, New York, California, Alaska and Hawaii.

Compared with FMR indexes, those Malpezzi rates are quite a bit higher for the District of Columbia and for such states as California, Connecticut, and Massachusetts, but also Alabama, Kentucky and Louisiana. On the other hand, they are a bit lower for Alaska, Mississippi, Nevada, and Texas.

Finally we show the last set of indexes calculated – mean FMRs by state and metro/nonmetro status. Once again, we see patterns in state poverty rates that look like the FMRs applied directly to small geographic areas. Overall, however, it is clear that there is a general pattern between the unadjusted official rates and any of the index-adjusted rates by state.

To summarize the results presented, this paper has presented an examination of geographic adjustments to poverty thresholds. I have discussed available data, previous application of such work, and a reapplication of various indexes to the computation of poverty rates using these adjustments. The focus in this comparison has been the estimate of poverty rates by state.

Using a set of housing price indexes at various levels of aggregation suggests to us that the official unadjusted numbers generally misreport poverty rates by state. Some consistent patterns of misreporting are clear regardless of which adjustments are used and these are that the official poverty rates are too high in the south and west, and too low for some New England states, Alaska and Hawaii. Even with gross aggregation to alleviate problems with confidentiality, the index-adjusted numbers appear to report more reasonable poverty rates for the specific areas.

VII. Future research

While the final set of indexes examined were prepared to be compatible with publicly released data, further investigation is required to insure that confidentiality is not breached. Further, the operations performed here only roughly approximate those that may actually be required to allow placement of these indexes, and poverty thresholds based on them, on public use data files.

Nevertheless, our purpose has been to show the problems encountered in applying these kinds of indexes to a measure that must subsequently be made available for public use. Balancing precision with disclosure issues is a difficult challenge. Given the exercises described in this paper, a reasonable conclusion is one that agrees with Ruggles that "…a strong case can be made for adjustment of the poverty thresholds for price differences by geographic area and that, even though we cannot make precise adjustments for very small areas, we quite possibly would make fewer errors if reasonable adjustments were made for larger geographic areas." This being so, we hope to use Fair Market Rents in the short run to further investigate the calculation of poverty estimates by state in experimental measures

In the long run, considerable research can be done. Data from the 2000 decennial census will soon be available at which time it would be interesting to re-estimate the indexes following Malpezzi et al. with more recent data. Even more useful, perhaps future development of housing price indexes would utilize the American Community Survey data, which provides information similar to the census long form on housing costs but on a more frequent basis.

Bibliography

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

Kokoski, Mary F., Patrick Cardiff and Brent Moulton, "Interarea Price Indices for Consumer Goods and Services: An Hedonic Approach Using CPI Data," BLS Working Paper 256, Office of Prices and Living Conditions, U.S. Department of Labor, Washington, DC, 1994.

Malpezzi, Stephen, Gregory Chun, and Richard Green, "New Place-to-Place Housing Price Indexes for U. S. Metropolitan Areas, and Their Determinants; An Application of Housing Indicators," *Real Estate Economics, 26(2), Summer 1998, pp. 235-275.*

Malpezzi, Stephen, "Remarks on Geographic Variation in Housing Prices and the Measurement of Poverty", Poverty Measurement Research Working Paper, August, 1998.

Renwick, Trudi J. and Barbara R. Bergmann, "A Budget-Based Definition of Poverty with an Application to single-Parent Families," *The Journal of Human Resources.* Winter 1993.

Ruggles, Patricia, *Drawing the Line – Alternative Poverty Measures and Their Implications for Public Policy,* Washington DC, Urban Institute Press, 1990.

SAS Institute Inc., *SAS/STAT User's Guide: Statistics, Version 6, Fourth Edition,* Cary, NC:SAS Institute, Inc., 1989.

Short, Kathleen, Thesia Garner, David Johnson, and Patricia Doyle, *Experimental Poverty Measures: 1990 to 1997,* U.S. Census Bureau, Current Population Reports, Consumer Income, P60-205, U.S. Government Printing Office, Washington, D.C. 1999.

The Runzheimer Plan of Living Cost Standards, Runzheimer International (Rochester, WI: June 1994).

U.S. Department of Health, Education, and Welfare, *The Measure of Poverty: A Report to Congress as Mandated by the Education Amendments of 1974,* Washington, and DC. US Government Printing Office, April 1976.

U.S. General Accounting Office, *Poverty Measurement: Adjusting for Geographic Cost-of-Living Differences,* GAO/GGD-95-64, March 1995.