Fully Utilizing Housing Cost Data in the American Community Survey PUMS Data: Identifying Issues and Proposing Solutions

Keith E. Wardrip
Danilo Pelletiere
National Low Income Housing Coalition

Abstract

The American Community Survey (ACS) is emerging as a valuable tool for analyzing annual trends and patterns in housing in the United States. Researchers often use the housing cost-to-income ratios (HCIRs) provided in the ACS Public Use Microdata Sample housing file to evaluate the level of housing cost burden for renters and owners and to estimate the proportion of households spending more than a specified level of income, often 30 percent or 50 percent, on shelter. In this article, we show that these variables should be used with caution, identifying 3.2 million households in the 2006 ACS for which the Census Bureau does not calculate an HCIR, even though useful housing cost and income data are available for these households. We also identify 2.8 million owner households for which the HCIR is underestimated because monthly costs do not

Abstract (continued)

include mobile home fees. This article explores these issues, explains how researchers can develop an alternative HCIR, and describes the resulting distribution of households by housing cost burden.

Introduction

Administered to roughly 3 million households annually by the Census Bureau, the American Community Survey (ACS) has estimated the demographic and socioeconomic characteristics of the U.S. population since 2000 and is slated to replace the long form in the 2010 Census. ACS data are available in a variety of formats, including two Public Use Microdata Sample (PUMS) files. The file analyzed in this article contains more than 1.3 million housing unit records describing the housing stock in 2006, and a second, larger file contains information on the individuals living in those units; the records in both files are weighted, allowing users to estimate national or state-level characteristics from the available survey data.[1]

The ACS questionnaire gathers data on a variety of housing cost components, including rent and mortgage payments, utilities, taxes, insurance, and mobile home and condominium fees.[2] To produce estimates of aggregate renter housing costs in the ACS, the Census Bureau adds contract rent to the cost of utilities and fuels into a variable called gross rent. (GRNTP is the variable name for gross rent used in the PUMS housing file.) For owners, the comparable housing cost variable is referred to as selected monthly owner costs (SMOCP).[3]

The Census Bureau also calculates two variables to quantify the extent to which a household is burdened by its housing costs: gross rent as a percentage of household income (GRPIP) and selected monthly owner costs as a percentage of household income (OCPIP). This article refers to these types of housing cost burden variables as housing cost-to-income ratios, or HCIRs.

As researchers increasingly tap the ACS for the housing cost data that it provides, it is significant that the 2006 PUMS housing file does not include an HCIR for roughly 3.2 million households, 2.8 million of which are renters.[4] The Census-calculated HCIR underestimates the housing cost burden for an additional 2.8 million owner households because it is predicated on an aggregation

[1] Data can also be tabulated for substate areas called Public Use Microdata Areas. For more information on the ACS, visit www.census.gov/acs/www/.

[2] The 2005-06 ACS questionnaire is available at www.census.gov/acs/www/Downloads/SQuest05.pdf.

[3] U.S. Census Bureau, American FactFinder glossary (www.factfinder.census.gov/home/en/epss/glossary_a.html).

[4] Household counts provided in this article are from tabulations of the 2006 ACS PUMS housing file and use the housing weight variable (WGTP) provided therein. Excluding the unweighted records representing those living in group quarters, each record in the file represents roughly 100 households.

of monthly owner costs that does not include mobile home fees. The remainder of this article describes how to calculate alternative HCIR variables with data provided in the ACS PUMS file and explores why, in some cases, user-calculated gross housing costs are not consistent with the summary values provided in the ACS. We conclude with a look at the level of housing cost burden indicated by the alternative HCIR for the roughly 6.0 million affected households.

Calculating Alternative HCIRs for Renters and Owners in the ACS

Because housing costs are rarely as simple or as straightforward as a monthly rent or mortgage payment, the Census Bureau aggregates several secondary components of housing costs with these primary housing expenditures in an effort to develop estimates of shelter costs that are more universally comparable. For renters, this calculation is important because monthly rent payments can include all, some, or no utilities, depending on the terms of the lease or the arrangement with the property owner. Similarly, for owners, this practice recognizes that a mortgage payment does not represent a consistent proportion of total housing costs, because some homeowners have no mortgage and because insurance, taxes, and fees can vary considerably from home to home.

To calculate HCIRs from scratch, we need to accurately calculate both total household income and total housing costs, as defined by the Census Bureau, from the variables provided in the PUMS file. We find that the aggregate household income variable (HINCP) is provided for all households and does not appear to pose any difficulties. To be consistent with the majority of housing costs, we suggest dividing household income by 12 to reflect a monthly income figure.

Developing a gross housing cost variable from its individual components requires a little more manipulation of the PUMS data and varies somewhat by tenure. For renter households, the value provided for contract rent (RNTP)—the basis of the calculation—can be used as is.[5] To calculate an alternative gross rent value, however, the following components must be manipulated as described in the following paragraphs and added to the contract rent.[6]

Electricity (monthly cost) (ELEP). Values of 1 and 2, which indicate that the cost of electricity was included in other fees, there was no charge, or electricity was not used, must be recoded to 0.

Gas (monthly cost) (GASP). Values of 1 through 3, which indicate that the cost of gas was included in other fees, there was no charge, or gas was not used, must be recoded to 0.

Water (yearly cost) (WATP). Values of 1 and 2, which indicate that water/sewer costs were included in other fees or there was no charge, must be recoded to 0. Values over 2 must be divided by 12 to arrive at a monthly cost.

---

[5] Because contract rent—but not gross rent—values are provided for vacant rental units in the PUMS housing file, calculating an alternative gross housing cost variable predicated on contract rent allows researchers the flexibility to consider the affordability of vacant units in their analyses.

[6] The authors will provide the Statistical Package for Social Sciences—SPSS—syntax used to perform these tabulations upon request.

House heating fuel (yearly cost) (FULP). Values of 1 and 2, which indicate that the costs for fuels, such as oil, coal, and kerosene, were included in other fees, there was no charge, or these fuels were not used, must be recoded to 0. Values over 2 must be divided by 12 to arrive at a monthly cost.

For owners, first and second mortgage payments (MRGP and SMP, respectively) replace contract rent in the calculation of housing costs. In addition to mortgage payments and the utilities listed above, owner costs as defined by the Census Bureau should include condo fees (CONP),[7] fire/hazard/ flood insurance (INSP), property taxes (TAXP), and mobile home costs (MHP). Insurance, property taxes, and mobile home costs are annual and must be divided by 12 to reflect monthly costs.

A user-calculated HCIR is simply the ratio of aggregate monthly renter/owner housing costs to one-twelfth of a household's annual income, and, consistent with ACS protocol, HCIRs are topcoded at 101 in this exercise. Although the Census Bureau ignores households with zero or negative income when calculating HCIRs, we assign these households an HCIR of 101, providing housing costs are greater than zero;[8] regardless of income, households for which housing cost components sum to zero are assigned an HCIR of 0. Because Census-calculated HCIRs are integers ranging from 1 to 101, we round the user-calculated HCIRs to the nearest whole number.

Owner Cost Complications

One issue that complicates the development of an alternative HCIR for owner households is the treatment of TAXP in the PUMS file. The ACS questionnaire includes an item to record a household's actual annual property taxes, but in the PUMS file, the variable TAXP is categorical, with each coded response representing a range of actual responses (for example, a value of 30 represents the range $1,800–$1,899). This treatment complicates calculating owner costs from scratch because we would have to approximate the actual taxes paid by using the midpoint of the indicated tax range (for example, $1,850 in the example above).

Given this complication and our analyses indicating that SMOCP accurately summarizes the components of monthly costs for most owner households, we recommend using it in the calculation of an alternative HCIR in most cases.

The PUMS file does not accurately aggregate housing costs for owner households of all types, however. Through the course of our analyses, we discovered that it does not include mobile home costs for the roughly 3.5 million owner households living in mobile homes and paying annual

---

[7] The 2005-06 ACS questionnaire indicates that renters who pay condominium fees in addition to their contract rent should list these fees separately; however, this value is missing for all 36.5 million renter households.

[8] Income can be negative for only two components of total earnings: self-employment (SEMP) and interest, dividends, and net rental income (INTP). In its *Affordable Housing Needs 2005: Report to Congress*, HUD does not calculate a level of cost burden for renter households with zero or negative income and actually imputes a higher income for those paying more than the Fair Market Rent because their lack of income is assumed to be temporary. Because ACS data are intended to capture income earned in a 12-month period and not "typical" income levels, this analysis differs from HUD's methodology and assumes that households reporting housing costs and zero or negative income should be considered burdened by these costs. The HUD study is available at www.huduser.org/publications/affhsg/affhsgneeds.html.

site rent and related fees.[9] For these households, an alternative owner cost variable that includes these fees and the additional components described in the previous section should be calculated. Of necessity, this alternative housing cost value must incorporate the midpoint of the property tax range rather than the actual property taxes reported by the household; however, our analyses show that, although not perfect, in the vast majority of cases this methodology is very effective at approximating monthly owner costs.[10]

HCIR Discrepancies

As exhibit 1 shows, a comparison of user- and Census-calculated HCIRs reveals a significant number of discrepancies worth investigating.

The user-calculated HCIR differs from the housing cost burden value provided in the PUMS file for roughly 2.8 million owner households. With few exceptions, these higher alternative HCIRs can be traced to the exclusion of mobile home costs from aggregate owner costs and the resulting underestimation of housing costs and cost burden.[11, 12] These households reported $1,400 in annual mobile home costs, on average, and the user-calculated HCIR indicates that the median household in this group spent a significantly higher proportion of income on housing (21 percent) than is implied by the Census-calculated HCIR (13 percent).

Exhibit 1

Comparison of ACS-Provided and User-Calculated HCIRs From the 2006 PUMS Housing File

	Renter-Occupied	Owner-Occupied	Total
Total households	36,542,589	75,074,799	111,617,388
HCIRs equal	33,747,467	71,818,807	105,566,274
HCIRs unequal	2,438	2,840,260	2,842,698
HCIR missing in PUMS	2,792,684	415,732	3,208,416

ACS = American Community Survey. HCIR = housing cost-to-income ratio. PUMS = Public Use Microdata Sample.

[9] Census Bureau staff confirmed our conclusion that mobile home costs are erroneously excluded from the SMOCP in the PUMS housing file and, during their investigation, found this exclusion to be the case dating back to 2000.

[10] Most TAXP categories represent a range of $100 or less, so the midpoint is never more than $50 above or below the actual property taxes paid by the homeowner. Because the figure reported is annual, using this midpoint produces an estimate of taxes that is never more than $4.17 from the actual taxes on a monthly basis ($50/12). This methodology is not as effective for households paying annual property taxes in excess of $5,000 because the ranges are too broad for the midpoint to closely approximate the actual taxes paid, but only 0.3 percent of the 3.5 million households with mobile home costs fall into one of these higher property tax brackets.

[11] Although approximately 3.5 million owner households reported costs associated with mobile homes, these costs on a monthly basis were significant enough to affect the calculated HCIR for only 2.8 million households.

[12] Exceptions include 42,883 households with a Census-calculated HCIR of 1, but for which the alternative HCIR rounds to 0, and 1,557 households for which the alternative HCIR is lower than the Census value as a result of approximating property taxes using the midpoint of the appropriate tax range.

For 2,438 renter households, gross rent accurately reflects the sum of housing cost components and totals as much as $150 for some households, although more than half reported costs of $4 per month (13 records representing 1,370 households). In all cases, the HCIR in the PUMS file equals 1, but the user-calculated HCIR rounds to 0.

Missing HCIRs

Census-calculated HCIR values are missing for 3.2 million households (exhibit 1). The omission of these values is significant because tabulations of these cost burden variables would therefore suppress the housing affordability experiences of nearly 3 percent of all U.S. households—and almost 8 percent of all renters. As stated previously, the problem can be traced largely to the aggregation of gross housing cost measures in the ACS.

First, focusing specifically on renters, exhibit 2 separates the 2.8 million households without a Census-calculated HCIR value into those classified as paying cash rent, which have valid values for the contract and gross rent variables, and those paying no cash rent, which are missing these values. The vast majority of the latter—1.8 million out of 2.2 million—reported monthly expenditures for at least one of the four component utility variables. The Census Bureau makes clear in published tables and in the PUMS Data Dictionary that gross rent and HCIR values are calculated only for households that pay cash rent, but these results illustrate that "no cash rent" is not synonymous with "no housing costs." Using these PUMS variables "out-of-the-box" excludes not only 0.4 million households with no obvious housing costs but also 1.8 million that did, in fact, pay at least one of the components of gross rent as defined by the Census Bureau.

HCIR values are also missing in the PUMS file for nearly 0.6 million renter households classified as paying cash rent, presumably because, with the exception of 158 households (2 records),[13] they

Exhibit 2

Renter Households With No HCIR Value in the ACS PUMS Housing File

	Tenure Classification in ACS	
	No Cash Rent	**Cash Rent**
Total	2,212,501	580,183
User-calculated gross rent		
Zero	401,963	0
Greater than $0	1,810,538	580,183
Household income		
Zero or negative	82,264	580,025
Greater than $0	2,130,237	158

ACS = American Community Survey. HCIR = housing cost-to-income ratio. PUMS = Public Use Microdata Sample.

---

[13] These households reported an annual income of $1 in 2006. According to the PUMS Data Dictionary, $1 can represent monetary income or it can be a code for "break-even" status for two components of income (interest, dividends, or rental income [INTP] and self-employment [SEMP]). Therefore, it is unclear whether these households reported $1 of income or reported breaking even on one of these components.

reported zero or negative household income. Each household incurred a monthly housing expense and, given the lack of household income, should therefore be assigned an HCIR of 101 if they are to be included in any housing affordability analysis.

The issue of owner households lacking HCIR values in the PUMS file can also be traced to household income rather than housing costs. As shown in exhibit 3, for the vast majority of these households, the components of selected monthly owner costs sum to a value greater than 0, but none reported an annual household income greater than $1 (see footnote 13). Researchers who wish to include these records in their analyses should assign them an HCIR value of 101. The remaining 1,533 households with no housing costs can reasonably be assigned an HCIR value of 0.

The next section briefly looks at the effect of the user-calculated HCIRs on the level of cost burden associated with these 3.2 million households and the 2.8 million owner households discussed in the preceding section.

Exhibit 3

Owner Households With No HCIR Value in the ACS PUMS Housing File

	Owner Households
Total	415,732
User-calculated owner costs	
Zero	1,533
Greater than $0	414,199
Household income	
Zero or negative	414,909
$1	823

ACS = American Community Survey. HCIR = housing cost-to-income ratio. PUMS = Public Use Microdata Sample.

Impacts on Estimates of Housing Cost Burden

The user-calculated HCIR values described in this article can clearly affect the housing cost burden statistics derived from PUMS housing data.

For example, Census-calculated HCIR values are lower than user-calculated values for nearly all of the 2.8 million owner households for which the Census Bureau excludes mobile home fees in its estimate of monthly owner costs. As a result and as exhibit 4 shows, 29 percent of these owner households were either moderately or severely burdened by total housing expenses, whereas the values provided in the PUMS file indicate that only 15 percent fall into these categories.

For the 3.2 million households for which the PUMS file does not calculate HCIR values, exhibit 5 summarizes their cost burden distribution according to the user-calculated values. Because most such renter households did not pay cash rent, the vast majority (68 percent) were not unduly burdened by their housing costs, even when other costs are considered.

Most of the remaining renter households and nearly all of the owner households in exhibit 5 are considered severely burdened by their housing costs because they reported zero or negative income but some level of housing expenditures over the previous 12 months.

Exhibit 4

Varying Levels of Cost Burden for Owner Households for Which Census- and User-Calculated HCIRs Are Unequal

Percent of Income Spent on Housing	Based on Census-Calculated HCIR		Based on User-Calculated HCIR	
	(number)	(%)	(number)	(%)
Total households	2,840,260	100	2,840,260	100
0-30%—not burdened	2,404,140	85	2,014,972	71
31-50%—moderately burdened	287,497	10	509,064	18
51%+ —severely burdened	148,623	5	316,224	11

HCIR = housing cost-to-income ratio.

Exhibit 5

Actual Level of Housing Cost Burden for Households Lacking HCIR Value in the PUMS Housing File

Percent of Income Spent on Housing	Renter		Owner	
	(number)	(%)	(number)	(%)
Total households	2,792,684	100	415,732	100
0-30%—not burdened	1,907,698	68	1,533	0
31-50%—moderately burdened	113,347	4	0	0
51%+ —severely burdened	771,639	28	414,199	100

HCIR = housing cost-to-income ratio. PUMS = Public Use Microdata Sample.

Viewed in the context of all U.S. households, the effect of calculating alternative HCIRs for only 6 million households does not significantly affect national housing cost burden statistics. Whether using Census- or user-calculated values, the median HCIR is 21 for the 75 million owner households; for the 36 million households who rent, the median user-calculated HCIR is one point lower (29) than the median Census value (30) because the former includes a large number of households paying no cash rent. Even though their impact on national statistics may be negligible, these user-calculated HCIRs may influence housing research that focuses on specific, largely lower-income segments of the population.

Conclusion

The primary intent of this article is to focus the attention of researchers on the peculiarities of the housing cost fields in the ACS PUMS housing file and to sound a cautionary note: Researchers should be aware that using gross housing cost and housing cost burden variables may lead them to unnecessarily exclude roughly 3.2 million households from their analysis, either because they did not pay cash rent (2.2 million) or because they reported housing costs but zero or negative income (1.0 million). Adding to our concerns, fields in the PUMS file exclude certain costs associated with mobile homes, thus materially affecting the HCIR for an additional 2.8 million owner households.

Accepting the Census aggregate housing cost and housing cost burden variables at face value may result in the loss or misinterpretation of housing cost information for a significant number of U.S. households. Because this information is systematically missing or incorrect for specific subsets of the population—households paying no cash rent, reporting zero or negative income, or living in mobile homes—the use of these variables may unintentionally exclude these households from tabulations of the PUMS housing file and bias research findings. Armed with this knowledge, researchers can make informed decisions regarding the appropriateness of the out-of-the-box PUMS file and weigh the advantages and disadvantages of calculating their own housing cost variables.

Acknowledgments

Authors

Keith E. Wardrip is a research analyst at the National Low Income Housing Coalition.

Danilo Pelletiere is a research director at the National Low Income Housing Coalition.