Housing Affordability Data System
David A. Vandenbroucke, HUD
January 28, 2011

I. Introduction

The Housing Affordability Data System (HADS) is a set of housing-unit level datasets that measures the affordability of housing *units* and the housing cost burdens of *households*, relative to area median incomes, poverty level incomes, and Fair Market Rents. The purpose of these datasets is to provide housing analysts with consistent measures of affordability and burdens over a long period. The datasets are based on the American Housing Survey (AHS) national files from 1985 through 2009 and the metropolitan files from 2002 through 2009. Users can link records in HADS files to AHS records, allowing access to all of the AHS variables.

The HADS grew out of a project to provide similar tabulations to the Millennial Housing Commission (MHC) for the years 1985, 1995, and 1999.[1] The strength and value of the HADS is that it incorporates more than twenty years of housing data using assumptions and computations consistent with the practice of the housing analysts that contributed to the MHC. This allows policy deliberations to focus on real policy choices, without the distraction of debating the meaning of the underlying data. Note that HADS development continued after the commission published its report. Thus, tabulations using the HADS datasets will not be exactly the same as those published by the MHC.

This document is a summary of how we constructed the HADS. It is a description of data selection, affordability calculation, and tabulation categories. The document is aimed at housing analysts who are familiar with current issues and techniques of measuring housing affordability. Technical points concerning computer programming are relegated to footnotes. Those who are proficient in reading SAS code are encouraged to consult the program listings that accompany the datasets.

A. Data Sources

The main data sources are the American Housing Survey (AHS) national sample microdata, for the odd-numbered years in 1985-2009 and the AHS metropolitan sample microdata for 2002-2009.[2] Poverty income is based on the Census Bureau's official thresholds[3]. Area median income (AMI) and Fair Market Rent (FMR) data come from HUD calculations, as discussed below. We selected only records representing completed interviews for occupied and vacant units, excluding usual residence elsewhere (URE) and noninterview records.

---

[1] The Millennial Housing Commission was established by the U.S. Congress in 2000 with the mission "to identify, analyze, and develop recommendations that highlight the importance of housing, improve the housing delivery system, and provide affordable housing for the American people, including recommending possible legislative and regulatory initiatives." MHC Research Director Eric Belskey asked HUD for assistance in tabulating American Housing Survey data to meet the Commission's data needs. For more information on the MHC, see http://govinfo.library.unt.edu/mhc/home.html .

[2] We plan to add additional files as new AHS survey results become available. Thus, the HADS may include additional files, for surveys conducted after this document was written. Note that some metropolitan surveys were conducted as supplemental samples of the national survey. The data for these metropolitan areas are included in the national datasets.

[3] http://www.census.gov/hhes/poverty/threshld.html .

B. Disclaimers

The work here is based on consensus agreements made among a group of about six HUD and MHC analysts. The techniques used do not necessarily represent HUD standards. The HADS datasets contain no proprietary or confidential data. The primary data source is the Public Use Files (PUFs) of the American Housing Survey. As the name implies, these files are available to the public.[4] The secondary sources of data are the HUD Fair Market Rents and Income Limits, which are public data.[5] HADS also makes use of the published poverty income thresholds. Thus, there is nothing in the HADS that anyone with the patience and programming skill could not reproduce.

HADS tabulations have no official sanction from the U.S. Department of Housing and Urban Development. We are making the datasets available as a tool for researchers to use as they wish. Any conclusions drawn from the analysis of HADS are those of the analysts, and not of HUD.

II. Custom Variable Definitions for HADS Data

Each record in the HADS dataset represents one sampled housing unit.[6] The variables in the record include a subset of standard AHS variables plus custom, derived variables. This section discusses the latter. See appendix A for a complete list of the variables in each HADS file.

A. Tenure

Tenure for occupied and vacant housing units is collapsed into the OWNRENT variable:

- Rental: Occupied units rented for cash and without payment of cash rent. Vacant for rent, vacant for rent or sale, and rented but not occupied.

- Owner: Owner occupied, vacant for sale, and sold but not occupied.

B. Monthly Housing Costs

1. Housing Cost Concept

HADS measures monthly housing cost using a cash flow concept. It is the monthly expenditure for shelter, including utilities, insurance, and some other costs, as described below. For renters, our measurement is straightforward. For owners, we make four specific assumptions about interest rates, as explained below. The monthly housing cost variables are COST06, COST08, COST12, and COSTMED.

Note that our measure of housing cost does not account for the income tax deductibility of mortgage interest. Neither does it include an estimate of the opportunity cost of owners' capital investment. Finally, it

---

[4] See http://www.huduser.org/datasets/ahs.html for information on how to download the datasets or order them on CD-ROM.

[5] Admittedly, some of the older files are difficult to find, particularly in electronic form.

[6] Since the AHS is a probability sample each sampled household represents many other households when making estimates using the sample weight.

also excludes expenditures for major repairs and replacements. It is simply a month-to-month cash flow measure.

The HADS includes housing cost estimates for both occupied and vacant units.

2. *Definition of Renter Costs*

For renters, housing cost is contract rent plus utility costs[7].

As is explained in the next section, there are four housing cost variables. Their values are *identical* for renters.

3. *Definition of Owner Costs*

Actual housing cost expenditures for owner-occupied units vary whether or not the household has a mortgage and, among mortgaged units, because of the historical contingency of the market interest rate when each household obtained or refinanced its mortgage. In order to measure affordability free of this contingency, we assume that *owners* hold fully amortizing 30-year fixed rate mortgages with 10 percent down payments.[8] The four basic cost measures are estimates of mortgage payments at 6, 8, and 12 per cent interest rates, plus the median interest rate in the AHS dataset for the survey year.[9] The amount financed is assumed to be the value of the unit, minus 10 per cent for the down payment. To each basic cost are added utility cost, other cost (see below), and an allowance for property insurance and property taxes. The combined amount of taxes and insurance is set at 1.5 percent of value per year.

Table 1 compares the median interest rate used in each HADS dataset with the mean rate reported by Freddie Mac's Primary Mortgage Market Survey (PMMS)[10]. Note that the HADS median is based on the stock of owner-occupied units in the given year, regardless of when the mortgage was made, while the PMMS mean is based on mortgage originations during the year.

---

[7] For most renter households, this is the variable ZSMHC. However, in order to be consistent with current practice, cost for no-cash renters (TENURE=3) in 1985-1995 is calculated as UTILITY + OTHERCOST. (These variables are described later in the document.)

[8] Note that when we calculate housing cost burden for occupied units, we use the reported monthly housing cost, not this computed cost.

[9] The mortgage payment calculation uses the SAS function MORT.

[10] http://www.freddiemac.com/pmms/pmms30.htm .

Table 1
AHS and PMMS Interest Rates

Year	AHS Median	AHS Mean	PMMS Mean
1985	9.75	9.90	12.43
1987	9.50	9.41	10.21
1989	9.75	9.66	10.32
1991	9.50	9.58	9.25
1993	8.25	8.52	7.31
1995	8.00	8.37	7.93
1997	8.00	8.12	7.60
1999	7.38	7.68	7.44
2001	7.25	7.60	6.97
2002	7.00	7.32	6.54
2003	6.25	6.60	5.83
2004	6.00	6.11	5.84
2005	6.00	6.10	5.87
2007	6.00	6.27	6.34
2009	5.88	6.00	5.04

4. Imputing Utility Costs for Vacant Units

Housing cost includes utility costs. The AHS does not collect utility costs for vacant units. However, the survey does ask if the costs would be paid by the tenant or included in rent. We use this information to impute utility costs to vacant units using a hot-deck procedure.

For each utility (gas, oil, electricity, other fuel, trash collection, and water), the allocation matrix has four dimensions:

- Monthly rent or mortgage payment (16 categories, $100 intervals). We estimate the mortgage payment the same way as in the housing cost section above, assuming an 8 per cent interest rate.

- Structure type (2 categories). Single-unit or multi-unit.[11]

- Region (4 categories). The four Census regions.

- Tenure (2 categories). Owner or renter.

As the data is processed, we assign the utility costs paid by occupied units to the allocation matrix in the appropriate cells. When a vacant unit is encountered and the appropriate variables indicate that the occupant would pay the costs, then we read the utility cost from the appropriate cell and assign it to the unit.

---

[11] A few cases (n=8) in the 2009 national dataset have missing values for units in structure (NUNITS). These are treated as single family for imputation purposes.

A separate utility cost variable (UTILITY) is included in the dataset, the sum of all applicable utility costs.

5. *Other Costs*

The components of monthly housing cost (ZSMHC) in the AHS have varied over the life of the survey. In order to measure cost consistently, certain non-utility costs are included in the HADS cost measure:

- Home owners' or renters' insurance. For vacant units, this is allocated by the same hot-deck procedure as utility cost.

- Land rent, where distinct from unit rent.

- Condominium fees, where applicable.

- Other mobile home fees, where applicable.

A separate variable (OTHERCOST) is the sum of these items.

C.   Housing Cost Burden

Housing cost burden is simply a household's monthly housing cost divided by its monthly income[12]. In particular, note that we *do not* use mortgage payment assumptions discussed in the "Housing Costs" section above when calculating burden.[13] Households with zero or negative income are given the special code of BURDEN = -1. Vacant units, not being households, have missing values for BURDEN.

D.   Assisted Housing

In the 1985-1995 datasets, a unit is assisted if:

- Proj = 1   [Unit is public housing]
- OR Sub = 1   [Household receiving a subsidy]
- OR SubInc = 1   [Household reports income for setting rent]
- OR SubLoc = 1   [Household receiving state or local subsidy]

---

[12] In most cases, BURDEN = ZSMHC / ZINC2. However, for no-cash rent households, BURDEN = (UTILITY + OTHERCOST)/ZINC2.

[13] Affordability measures where a housing unit fits in the cost distribution, regardless of the circumstances of the current occupants. Housing cost burden measures the actual expenditures of the household that occupies the unit.

In the 1997-2005 datasets, a unit is assisted if:

- Renew = '1' AND (RepHA = '1' OR RepHA = '2'))   [Income reporting needed for lease renewal and this is reported to the landlord, building manager, or housing authority]

- OR SubRnt = '1'   [Household receiving a subsidy]
- OR Proj = '1'   [Unit is public housing]
- OR Vcher='1'   [Household has a voucher]
- OR Apply='1'   [Applied to management to get public housing]

In the 2007 and later datasets, a unit is assisted if:

- VchrMv = '1'   [Can use voucher to move to another location]
- OR SubRnt = '1'   [Household receiving a subsidy]
- OR Proj = '1'   [Unit is public housing]
- OR Vcher='1'   [Household has a voucher]
- OR Apply='2'   [Assigned to specific address]

Note that the AHS has known problems with over reporting of subsidy status[14]. Thus, this classification should be used with caution.

E.   Total Salary Income

Total salary income (TotSal) is useful for identifying the "working poor" and measuring the labor force attachment of a household. This variable is simply the sum of wage and salary income (Sal) over all members of the household.[15]

---

[14] For example, see Gordon, Erika L.; Chipungu, Sandra; Bagley, Lisa Marie; and Zanakos, Sophia I. 2005. *Improving Housing Subsidy Surveys: Data Collection Techniques for Identifying the Housing Subsidy Status of Survey Respondents*. Prepared for Office of Policy Development and Research, U.S. Department of Housing and Urban Development by ORC Macro Calverton, Maryland, March. http://www.huduser.org/publications/pdf/datacollect.pdf

[15] In AHS surveys before 2005, the SAL variable for non-family household members (lodgers, roomates, etc.) included those persons' non-labor as well as labor income. Thus, TotSal may be an overestimate of household labor income.

III. Affordability Measurement Standards

The three commonly used affordability measures express cost of a housing unit relative to 1) the cost of safe, sound and adequate housing, 2) what a family earning the median income could be expected to pay for housing, or 3) what a family with income at the poverty level could be expected to pay for housing. Specifically, we provide three standards for assessing the affordability of housing units: fair market rent (FMR), area median income (AMI), and poverty-level income. For FMR, we express the monthly cost of each unit as a percentage of FMR. For the other two, we rate units as affordable at a certain percentage of the income measure, assuming that 30 percent of income is spent on housing. These three standards were chosen because each is commonly used in affordability discussions. Housing cost relative to AMI is perhaps the most common affordability standard. The FMR is the payment standard for housing assistance programs and is often used as a proxy for the cost of an "affordable" unit in housing literature. Poverty income is a widely recognized threshold, and is often used in the general press.

The affordability standards are used for two purposes:

- To classify *households* according to their ability to pay for housing.

- To classify *housing units* according to how expensive they are to occupy.

HADS includes variables that accomplish both of these purposes.[16]

A. Assigning Area Median Income and Fair Market Rent

HUD estimates AMI and FMR annually for all metropolitan areas and non-metropolitan counties.[17] However, the AHS public use file does not provide county identification, nor does it identify metropolitan areas of less than 100,000 population. Thus, HADS relies on certain procedures for linking AHS records to the appropriate AMIs and FMRs. The data system uses a somewhat crude system for this linking for files covering 1985-2001, but a more straightforward and accurate system is used for national files beginning with 2003 and all metropolitan files.

For the years 1985-2001, we estimate *AMI* and *FMR* for each unit using a computer program that was developed over time by a number of HUD analysts. This program uses the official HUD AMIs and FMRs in a base year and adjusts it for inflation[18] to estimate the corresponding values in the year desired, for each location. Locations include the metropolitan areas that are identified in the AHS (SMSA). Units for which we do not

---

[16] There is a set of continuous and categorical variables for each of the three standards. For example, the continuous variable called IncRelFMRPct (Income Relative to FMR—Percent), and the corresponding categorical variable is IncRelFMRCat (Income Relative to FMR—Category). Similarly, the files contain variables that classify the housing units according to their monthly costs. Note that since there are four definitions of housing costs (based on interest rates) and three affordability standards, there are twelve measures of affordability for each unit, each with a continuous and a categorical variable. Thus, the continuous variable that measures cost relative to FMR, assuming a 6 percent interest rate is Cost06RelFMRPct, and the corresponding categorical variable is Cost06RelFMRCat.

[17] Analysts of the Economic Market Analysis Division, within the Office of Policy Development and Research, perform this task. Data for recent years is available at http://www.huduser.org/datasets/pdrdatas.html . Follow the "income limits" link for AMI.

[18] Inflation adjustment is based on the national Consumer Price Index, all items, all urban consumers.

have this information are classified by heating degree-days, metropolitan status, and Census region[19]. Within each such area, the AMI and FMR are estimated as a weighted average, using 1990 population as weights. Estimates for 1985-1991 use a base year of 1993, adjusted backwards for inflation. Estimates for 1993-1995 use data from the year in question, with no need for an inflation adjustment. Estimates for 1997-2001 use a base year of 1995, adjusted forwards for inflation.

Beginning with the 2003 survey, the national AHS public use file (PUF) includes AMI and FMR estimates based on HUD published data for the current year. Thus, there is no need for inflation adjustment. HUD data were matched to AHS records by metropolitan area or county, using the (confidential) Census internal file. In order to preserve confidentiality, the values for records representing units located outside of the metropolitan areas that the AHS identifies were replaced with weighted averages as described above, except that the standard AHS weight variable is used.

The HADS files based on the AHS metropolitan surveys for 2002 and 2004 simply use the published AMI and FMR data for the appropriate metropolitan areas and years, without further adjustment. (Since the metro area is, necessarily, identified for records in metropolitan surveys, there was no need to use averaged values.)

B.   Measuring Household Income and Unit Affordability Relative to Area Median Income

AMI is used for two purposes: 1) To classify *households* on the basis of income received, and 2) to classify *housing units* on the basis of the income needed to afford them ("affordability"). These two purposes require different kinds of adjustments to the basic income measures.

1. AMI Variables

What we are calling AMI is actually based on four related income measures:

1. *Median Income* (LMed), which is not adjusted for the number of persons in the household.
2. *Low Income* (L80), which *is* adjusted for the number of persons in the household.
3. *Very Low Income* (L50), which *is* adjusted for the number of persons in the household.
4. *Extremely Low Income* (L30), which *is* adjusted for the number of persons in the household. L30 is available in the AHS PUF only for 2003 and later national surveys.

L80 and L50 are HUD programmatic income limits. As their variable names imply, they are based on 80 percent and 50 percent of area median income, respectively. However, these income limits are subject to a number of administrative adjustments, most of which serve to put floors under their values. Thus, these income limits are often *higher* than simply 80 or 50 percent of the corresponding median income.[20] These adjustments affect L50 more than L80. We use the official programmatic income limits in order to classify households and housing units in a way consistent with the rules for program eligibility.

---

[19] That is to say, using DEGREE, METRO or METRO3, and REGION variables. METRO and METRO3 indicate whether a unit is in a central city, suburb, or outside a metropolitan area. Further subdivision is available for some survey years.

[20] Explaining the adjustments that go into Low Income and Very Low Income is beyond the scope of this document. Interested parties might want to read the official summary at http://www.huduser.org/datasets/il/fmr02/briefing02.pdf .

2. *Real Income Growth Adjustment*

While the AMI estimation program for 1985-2001 adjusts for inflation, median incomes rose faster than inflation in that period. In order to correct for this tendancy to underestimate median income, HADS applies an additional real income growth adjustment:[21]

Table 2
Real Income Growth Adjustment Factors

Year	(1) Median Income[a]	(2) Ratio to Base Year[b]	(3) CPI	(4) Ratio to Base Year[b]	(5) Adjustment Factor[c]
1985	28,261	0.7119	107.6	0.7446	0.9560
1987	30,400	0.7657	113.6	0.7862	0.9740
1989	34,000	0.8564	124.0	0.8581	0.9980
1991	38,000	0.9572	136.2	0.9426	1.0155
1993	**39,700**	**1.0000**	**144.5**	**1.0000**	**1.0000**
1995	**40,200**	**1.0000**	**152.4**	**1.0000**	**1.0000**
1997	43,500	1.0821	160.5	1.0531	1.0275
1999	47,800	1.1891	166.6	1.0932	1.0877
2001	52,500	1.3060	177.1	1.1621	1.1238

[a]Source: HUD Economic Market Analysis Division
[b]Base year for 1985-1991 is 1993. Base year for 1997-2001 is 1995.
[c]Column (2) divided by column (4).

Note that the data after 2001 do not need this adjustment.

---

[21] In the dataset, the adjusted quantities are the variables GLMED, GL80, and GL50, "**G**rowth-adjusted income." To simplify programming, these variables are present in all datasets, even when no adjustment is required.

3. Classifying Household Incomes by AMI

HUD programmatic income eligibility rules allow larger households to have higher incomes while still qualifying for program participation. As noted above, the published L80, L50, and L30 income measures include these adjustments for the number of persons in the household, but LMed does not. In order to classify households by income in a consistent manner, we apply the adjustments to LMed as well. We multiply LMed by these standard HUD adjustment factors:[22]

Table 3
Household Size Adjustments to Income

Persons	Adjustment
1	0.70
2	0.80
3	0.90
4	1.00
5	1.08
6	1.16
7+	$(1.16 + 0.08(\text{Persons}-6))$

After making these adjustments, we divide households into income classes, as shown in the table below.[23] Note that the lower income classes are based on L30, L50, and L80, and thus include the effects of the HUD administrative adjustments.

We make a final check of households with zero or negative income. The AHS underestimates household income, compared to other surveys. In particular, an unusually high proportion of households reports zero or negative incomes, compared to other surveys. To partially correct for this bias, households that report zero or negative incomes but are living in adequate, uncrowded[24] units and paying at least the Fair Market Rent in housing costs are assigned to the income category 5 ($\leq$ LMED).

Table 4
Income Categories

Category	Income is…	Label	Comments
1	$\leq 0.60 * L50$	LTE 30% AMI	$0.60*0.50 = 0.30$ (1985-2001, metros)
1	$\leq L30$	LTE 30% AMI	2003 and later nationals
2	$\leq L50$	30.1 - 50% AMI	
3	$\leq 1.20 * L50$	50.1 - 60% AMI	$1.20*0.50 = 0.60$
4	$\leq L80$	60.1 - 80% AMI	
5	$\leq LMED$	80.1 - 100% AMI	
6	$\leq 1.20 * LMED$	100.1 - 120% AMI	
7	$> 1.20 * LMED$	120% AMI +	

---

[22] In the dataset, this is the variable APLMed, with the "AP" standing for "adjusted for persons."

[23] In the dataset, these are the "IncRelAMI" variables.

[24] "Adequate" means ZAdeq=1. "Uncrowded" means no more than 1 person per room (Per/Rooms $\leq$ 1).

4. *Classifying Unit Affordability by AMI*

Unit affordability relative to AMI is the percentage of area median income needed to afford the monthly housing cost associated with that unit. We use the standard assumption that a unit is affordable if the household spends no more than 30 percent of its income on housing. Thus, the basic relationship is:

$$\text{Affordability} = \frac{(\text{Monthly Housing Cost}) \times 12}{\text{AMI} \times 0.30} \times 100$$

Because the number of persons in the current household is irrelevant to measuring the affordability of the unit, we do not apply the household size adjustment to LMED. In addition, we de-adjust L30, L50, and L80 (which are already adjusted for number of persons), by dividing them by the adjustment factors listed above.[25] Because the number of bedrooms in the unit reflects the likely household size the unit would attract, we adjust the AMI for number of bedrooms, as shown in Table 5.[26]

Table 5
Number of Bedrooms Adjustments to Income[27]

Bedrooms	Adjustment
0	0.70
1	0.75
2	0.90
3	1.04
4	1.16
5	1.28
6	1.40
7+	1.40 + 0.12(bedrooms - 6)

We then express the percentage of AMI needed to afford the unit, using the categories described above in Table 4.[28]

C. Measuring Household Income and Unit Affordability Relative to Fair Market Rent

Unit affordability relative to FMR is simply a matter of expressing the monthly housing cost as a percentage of the FMR.[29]

---

[25] The income limits estimation program treats vacant units as 1-person households. Thus, these units are deadjusted as if they were 1-person households.

[26] The income levels adjusted for number of bedrooms are called ABL30, ABL50, ABL80, and ABLMed, where the "AB" stands for "adjusted for bedrooms."

[27] The adjustment factors are the ones required by statute for the Low Income Housing Tax Credit.

[28] In the dataset, these are the "CostnnRelAMIPct" and CostnnRelAMICat variables, where "nn" is a number indicating the interest rate (06, 08, 12, or "med" for median). The continuous variables, CostnnRelAMIPct, uses ABLMed in its divisor. The values of the categorical variables, CostnnRelAMICat, are assigned by beginning at the top of the list in Table 4 and choosing the first condition that is true.

[29] In the dataset, these are the "CostnnRelFMR" variables.

We calculate household income relative to FMR by calculating the maximum amount the household could afford to spend on housing (30 percent of monthly income[30]) and expressing this as a percentage of the FMR.[31] We use these categories:

Table 8
Fair Market Rent Categories

Category	Label
1	LTE 50% FMR
2	50.1-100% FMR
3	GT FMR

D. Measuring Affordability Relative to Poverty Income

We base poverty income on official poverty thresholds, given year, number of persons, number of children[32], and whether the householder is above age 65. For vacant units, we allocate the number of persons and number of adults according to number of bedrooms, as shown in Table 6.

Table 6
Poverty Income Household Size Assumptions
for Vacant Units

Bedrooms	Persons	Children	Elderly
0-1	1	0	No
2	2	1	No
3+	4	2	No

We calculate household income relative to poverty income as a multiple of the poverty threshold.[33]

Unit affordability relative to poverty income is the income a household would need in order to afford the monthly housing cost, expressed as a multiple of the poverty income threshold.[34]

We use these classes for the categorical variables:

---

[30] Thus, $0.30 \ast ZInc2/12$ as a percentage of FMR.

[31] In the dataset, these are the "IncRelFMR" variables.

[32] The number of children is calculated as $Per - ZAdult$. However, if $Zadult = 0$, then the number of children is reduced by 1. In other words, the householder is always considered to be an adult.

[33] In the dataset, these are the "IncRelPov" variables.

[34] In the dataset, these are the "CostnnRelPov" variables.

<u>Table 7</u>
<u>Poverty Income Categories</u>

<u>Category</u>	<u>Label</u>
1	LTE Poverty
2	100-150% Poverty
3	150-200% Poverty
4	200%+ Poverty

IV. Dataset Notes

A. Summary of Classification Categories

This section summarizes the classification intervals used for all variables in the dataset.

- Adequacy: adequate, moderately inadequate, severely inadequate, vacant–no information.

- Affordability and household income with respect to AMI: less than 30%, 30-50%, 50-60%, 60-80%, 80-100%, 100-120%, over 120%.

- Affordability and household income with respect to FMR: less than 50%, 50-100%, greater than FMR.

- Affordability and household income with respect to poverty income: below poverty, 100-150%, 150-200%, over 200%.

- Bedrooms: Studio (0 bedrooms), 1 bedroom, 2 bedrooms, 3 bedrooms, 4 or more bedrooms.

- Burden: Less than 30%, 30% to 50%, 50% or More, No Income.

- Location: central city, suburb, nonmetropolitan.

- Structure type: Single unit, 2-4 units, 5-19 units, 20-49 units, 50+ units, mobile homes.

- Tenure: Owner, renter.

- Year Built (age): After 2010[35], 2000-2009, 1990-1999, 1980-1989, 1960-1979, 1940-1959, Pre-1940.

B. Formatted Variables

Analysts often use the HADS datasets to perform tabulations using SAS, which are then output to spreadsheets for presentation or further analysis. However, SAS PROC EXPORT simply exports the numerical codes of categorical variables instead of the formatted value labels. In order to make spreadsheets easier to understand, the HADS datasets include a set of formatted variables. These are character variables containing value labels rather than numeric codes. Formatted variable names are of the form FMTvarname, where "varname" is the name of the corresponding numeric variable in the dataset. For example, FMTIncRelAMICat is the formatted version of IncRelAMICat (income relative to area median income, category).

C. Variable Names

While we tried to keep variable names consistent across datasets, in some cases variables names differ. The most common reason for this is that variable names changed in the AHS. We decided to keep the HADS

---

[35] For manufactured housing, the BUILT variable is set by model year. Thus, some units in the 2009 datasets will have BUILT=2010.

variable names consistent with the AHS, even if that meant they would differ across years. The table below lists the variable names that are not constant.

Table 8
Variable Names That Differ by Year

Description	Name, 1985-1995	Name, 1997-2009
Age of householder	AGE	AGE1
Formatted metro status	FMTMetro	FMTMetro3
Interview status	ISTATUS	STATUS
Metro status	METRO	METRO3[36]

Any variable name that is common between HADS and AHS is the same as AHS. Whenever we modified an AHS variable for HADS purposes, we did so by creating a new variable. For example, OWNRENT is a collapsed version of TENURE.

D.   Linking Files

The HADS files are based on the AHS public use files, but they include only those AHS variables that we believe to be most useful in analyzing housing affordability. However, users can make use of all the AHS variables by linking the HADS files to the standard AHS PUFs. In addition, since the AHS is a longitudinal survey, successive HADS files can be linked in order examine changes in housing affordability over time, at the housing unit level.

1. Linking to AHS Files

National HADS files may be linked to the corresponding AHS PUFs by matching records on the CONTROL variable. As released, the national HADS files are already sorted in CONTROL order. The metropolitan HADS files may be linked to the corresponding PUFs by matching on SMSA and CONTROL. As released, the metropolitan SAS files are already sorted in SMSA and CONTROL order. Note that some components of the PUF, such as the PERSON file, have a one-to-many relationship with the housing units. For details on linking to these files, see the notes in the main AHS codebook.

2. Longitudinal Linking

HADS files may be linked to one another for longitudinal analysis. The technique for linking records is the same as for linking to the AHS: match on CONTROL (CONTROL and SMSA for metro surveys). Note that the AHS (and hence HADS) follows housing units, not households. Thus, analysts interested in households

---

[36] The metropolitan surveys files use METRO, regardless of year. The meaning of the codes in the metropolitan METRO variable is different from the

must be careful to use the SAMEHH variable[37] to detect whether the household has changed in the interval between surveys. Another consideration is how to weight the linked records. The AHS WEIGHT variable is adjusted to current-year control totals and will not provide consistent longitudinal estimates. For a discussion on longitudinally weighting AHS files, see HUD's Components of Inventory Change reports at http://www.huduser.org/datasets/cinch.html.

E.   Optional Cost Patch for 1985-1993

About 8-9 per cent of owner-occupied cases in the 1985-1993 AHS have missing values for monthly housing cost (ZSMHC). This can occur because mortgage payment information is missing. In order to avoid arbitrary data changes, we left these missing values in HADS. However, we provide an optional program, *HADS85-93 Cost Patch.SAS*, for users who would like to patch these cases with reasonable values. The program estimates ZSMHC values using this procedure:

1. If the mortgage amount is missing, assume 90% of Value (10% down payment).
2. If the interest rate is missing, assume the median interest rate over units in that year's AHS.
3. If term is missing, assume 30 years.
4. If property tax & insurance are missing, assume 1.5% of value per year.

Given the above information, the program estimates the monthly mortgage payment using the SAS MORT function. The allocated value of ZSMHC is the sum of this mortgage payment, utility cost, and the property tax & insurance amount. The optional program adds an additional variable to the dataset, EstCostFlag, which is equal to 1 when this allocation has been performed and zero otherwise. It also calculates BURDEN for the patched records.

See the comments in the program itself for instructions on how to use it.

---

[37] SAMEHH is not on the HADS files. You must link to the AHS to use it.

V.   Contact Information

Dav Vandenbroucke
Senior Economist
U.S. Department of Housing and Urban Development
451 7th Street SW, Room 8222
Washington, DC   20410

david.a.vandenbroucke@hud.gov
Phone 202-402-5890
Fax 202-708-3316

Appendix A. Variable Listing

This is a list of all variables in the HADS datasets. A few variables (noted below) are available only in files for certain years.

Variable	Description	Notes
ABL30	Extremely Low Income Adjusted for # of Bedrooms	Only national, 2003 & later
ABL50	Very Low Income Adjusted for # of Bedrooms	
ABL80	Low Income Adjusted for # of Bedrooms	
ABLMED	Median Income Adjusted for # of Bedrooms	
AGE	Age of head of household	1985-1995
age1	Age of head of household	1997& later
APLMED	Median Income Adjusted for # of Persons	
ASSISTED	Assisted Housing	
BEDRMS	# of bedrooms in unit	
BUILT	Year unit was built	
BURDEN	Housing cost as a fraction of income	
CONTROL	AHS control number	
COST06	Housing cost at 6 percent interest	
COST06RELAMICAT	Cost06 Relative to Median Income (Category)	
COST06RELAMIPCT	Cost06 Relative to Median Income (Percent)	
COST06RELFMRCAT	Cost06 Relative to FMR (Category)	
COST06RELFMRPCT	Cost06 Relative to FMR (Percent)	
COST06RELPOVCAT	Cost06 Relative to Poverty Income (Category)	
COST06RELPOVPCT	Cost06 Relative to Poverty Income (Percent)	
COST08	Housing cost at 8 percent interest	
COST08RELAMICAT	Cost08 Relative to Median Income (Category)	
COST08RELAMIPCT	Cost08 Relative to Median Income (Percent)	
COST08RELFMRCAT	Cost08 Relative to FMR (Category)	
COST08RELFMRPCT	Cost08 Relative to FMR (Percent)	
COST08RELPOVCAT	Cost08 Relative to Poverty Income (Category)	
COST08RELPOVPCT	Cost08 Relative to Poverty Income (Percent)	
COST12	Housing cost at 12 percent interest	
COST12RELAMICAT	Cost12 Relative to Median Income (Category)	
COST12RELAMIPCT	Cost12 Relative to Median Income (Percent)	
COST12RELFMRCAT	Cost12 Relative to FMR (Category)	
COST12RELFMRPCT	Cost12 Relative to FMR (Percent)	
COST12RELPOVCAT	Cost12 Relative to Poverty Income (Category)	
COST12RELPOVPCT	Cost12 Relative to Poverty Income (Percent)	
COSTMED	Housing cost at Median interest	
COSTMedRELAMICAT	CostMed Relative to Median Income (Category)	
COSTMedRELAMIPCT	CostMed Relative to Median Income (Percent)	
COSTMedRELFMRCAT	CostMed Relative to FMR (Category)	
COSTMedRELFMRPCT	CostMed Relative to FMR (Percent)	
COSTMedRELPOVCAT	CostMed Relative to Poverty Income (Category)	
COSTMedRELPOVPCT	CostMed Relative to Poverty Income (Percent)	
FMR	Fair market rent (average)	
FMTASSISTED	Assisted Housing	
FMTBEDRMS	# of bedrooms in unit	
FMTBUILT	Year unit was built	

Variable	Description	Notes
FMTBURDEN	Cost Burden	
FMTCOST06RELAMICAT	Cost06 Relative to Median Income (Category)	
FMTCOST06RELFMRCAT	Cost06 Relative to FMR (Category)	
FMTCOST06RELPOVCAT	Cost06 Relative to Poverty Income (Category)	
FMTCOST08RELAMICAT	Cost08 Relative to Median Income (Category)	
FMTCOST08RELFMRCAT	Cost08 Relative to FMR (Category)	
FMTCOST08RELPOVCAT	Cost08 Relative to Poverty Income (Category)	
FMTCOST12RELAMICAT	Cost12 Relative to Median Income (Category)	
FMTCOST12RELFMRCAT	Cost12 Relative to FMR (Category)	
FMTCOST12RELPOVCAT	Cost12 Relative to Poverty Income (Category)	
FMTCOSTMEDRELAMICAT	CostMed Relative to Median Income (Category)	
FMTCOSTMEDRELFMRCAT	CostMed Relative to FMR (Category)	
FMTCOSTMEDRELPOVCAT	CostMed Relative to Poverty Income (Category)	
FMTINCRELAMICAT	HH Income Relative to Median Income (Category)	
FMTINCRELFMRCAT	HH Income Relative to FMR (Category)	
FMTINCRELPOVCAT	HH Income Relative to Poverty Income (Category)	
FMTMETRO	CENTRAL CITY / SUBURBAN STATUS	National 1985-1995, all metro
FMTMETRO3	CENTRAL CITY / SUBURBAN STATUS	National 1997 & later
FMTOWNRENT	Owner/Renter Status (adjusted)	
FMTREGION	Census Region	National only
FMTSTATUS	Occupancy Status	
FMTSTRUCTURETYPE	Structure Type	
FMTZADEQ	ADEQUACY OF UNIT	
GL30	Growth-adjusted extremely low income	National 2003 & later
GL50	Growth-adjusted very low income	
GL80	Growth-adjusted low income	
GLMED	Growth-adjusted median income	
INCRELAMICAT	HH Income relative to AMI (category)	
INCRELAMIPCT	HH Income relative to AMI (percent)	
INCRELFMRCAT	HH Income Relative to FMR (Category)	
INCRELFMRPCT	HH Income Relative to FMR (Percent)	
INCRELPOVCAT	HH Income Relative to Poverty Income (Category)	
INCRELPOVPCT	HH Income Relative to Poverty Income (Percent)	
IPOV	Poverty Income	
ISTATUS	Interview status	1985-1995
L30	Extremely low income limit (average)	National 2003 & later
L50	Very low income limit (average)	
L80	Low income limit (average)	
LMED	Area median income (average)	
METRO	CENTRAL CITY / SUBURBAN STATUS	National 1985-1995, all metro
METRO3	CENTRAL CITY / SUBURBAN STATUS	National 1997 & later
NUNITS	# of units in building	
OTHERCOST	Insurance, condo, land rent, other mobile home fees	
OWNRENT	Tenure (adjusted)	
PER	# of persons in household	
REGION	Census Region	
ROOMS	# of rooms in unit	
SMSA	1980 design PMSA code	Metro only
STATUS	Interview status	1997 & later

Variable	Description	Notes
STRUCTURETYPE	Recoded structure type	
TENURE	Owner/renter status of unit	
TOTSAL	Total Wage Income	
TYPE	Structure Type	
UTILITY	Monthly utility cost	
VACANCY	Vacancy status	
VALUE	Current market value of unit	
WEIGHT	Final weight	
ZADEQ	Recoded adequacy of housing	
ZINC2	Household Income	
ZSMHC	Monthly housing costs	