

Increase Research Opportunities By Accessing Internal Files of the Census Bureau

Warren A. Brown
University of Georgia

HACE Seminar
March 21, 2011

1

Household Survey Data

- [ACS](#) - American Community Survey
- [CPS](#) - Current Population Survey
- [AHS](#) - American Housing Survey
- [NHANES](#) – National Health and Nutrition Examination Survey
- [NSFG](#) – National Survey of Family Growth

2

Secondary Analysis of Survey Data

- Public Use Files
- Internal Use Files
- Differences

3

Network of Census RDCs



- Boston NBER
- Cornell U
- Baruch C – NYC
- Census Bureau
- Duke/RTI
- U of Michigan
- Chicago Fed
- U of Minnesota
- Berkeley
- Stanford
- UCLA

4

Atlanta Census RDC

5

ACrDC Consortium Members



6

NCHS Data Centers

7

Bureau of Labor Statistics

8

CES Discussion Papers

Survey	Publication	Reason for Restricted Data
NIHANS	Bird, C.E., Semman, T., et al. (2001) Neighbourhood socioeconomic status and biological "wear and tear" in a nationally representative sample of US adults. <i>Journal of Epidemiology and Community Health</i> 54: 865-865.	Census tract identifiers needed to merge contextual level variables
NIHANS	Diao, C.H., P. Patel, et al. (2011) Low Vitamin D among HIV-infected adults: Prevalence of and Risk Factors for Low Vitamin D Levels in a Cohort of HIV-infected Adults and Comparison to the Prevalence among Adults in the US General Population. <i>Clinical Infectious Diseases</i> 52 (2): 396-402	Latitude, state, and date of collection requested to merge IV data from the same source
NIHANS	Pletcher, J.M., Frisvold, D.E. & Todd, N. (2010) The effect of exit drunk taxes on child and adolescent consumption and weight outcomes. <i>Journal of Public Economics</i> 94: 967-974.	State of residence and date of interview (month and year)
NIHANS	Pletcher, J.M., Frisvold, D., & Todd, N. (2010). Taring Soft Drinks and Restricting Access To Vending Machines To Curb Child Obesity. <i>Health Affairs</i> , 29, 1-8.	State of residence and date of interview (month and year)
NIHANS	Lung, S.K., & Shadler, K. (2010) The Impacts of State Health Reform Initiatives on Adults in New York and Massachusetts. <i>Health Services Research</i> .	State of residence needed

9

Research Data Center

10

American Community Survey

PUF

- 1% Sample
- PUMA
- Top Coding
- Collapsed Categories
- Perturbation
- No Administrative

IUF

- 2.5% Sample
- Block/Tract/Place
- Full Distribution
- Full Distribution
- Actual Ages
- Details of Collection

11

12

ACS IUF Research

- Author: Andrew Houtenville
- Title: Complex Survey Questions and the Impact of Enumeration Procedures: Census/American Community Survey Disability Questions
- Need for IUF: Link ACS records to SSI disability data

13

CPS: ASEC (March) Items

- Demographic Characteristics
- Usual Monthly Labor Force Status
- Work Experience
- Income
- Noncash Benefits
- Migration
- Health Insurance

14

Current Population Survey

PUF

- All Months and Topics
- All Records
- State/MSA
- Top Coding
- Collapsed Categories
- No Administrative

IUF

- ASEC (March) Only
- All Records
- Tract
- Full Distribution
- Full Distribution
- Details of Collection

15

CPS IUF Research

- Author: Richard Burkhauser
- Title: Using the P90/P10 Index to Measure U.S. Inequality Trends with Current Population Survey Data: A View From Inside the Census Bureau Vaults
- Need for IUF: Full distribution of income without top-coding

16

CPS IUF Research

- Author: Katharine Abraham
- Title: Exploring Differences in Employment Between Household and Establishment Data
- Need for IUF: Link individual records in CPS to UI Wage Records

17

American Housing Survey

PUF

- All Records
- State/MSA
- Top Coding
- Collapsed Categories
- No Administrative

IUF

- All Records
- Tract
- Full Distribution
- Full Distribution
- Details of Collection

18

Date Last Revised: November 30, 2010
DRAFT

Page 1 of 7

American Housing Survey (AHS): Using the Internal User File (IUF)

The American Housing Survey (AHS) is sponsored by the U.S. Department of Housing and Urban Development (HUD) and the data are collected by the Census Bureau. The first national survey was conducted in 1973 and the first metro area was surveyed in 1974. To date surveys have been conducted in 48 metro areas.

Accessing the AHS

The AHS data are available to the general public in SAS and ASCII format. The Public Use Files (PUFs) are available at <http://www.huduser.org/pocrtal/datasets/ahs.html>. At this site users can download:

- 1) Datasets of all AHS surveys dating back to 1997;
- 2) Order datasets of AHS surveys conducted prior to 1997 on CD-ROM (if 1989 – years previous are available at the National Archives);
- 3) The codebook that provides the name and detailed information on each variable for each survey year (to access each year's codebook, click on that year's survey link on the webpage mentioned above and then click on the "Codebook to the American Housing Survey" link);
- 4) The survey questionnaire/booklet in both English and Spanish; and
- 5) Information on topcoded and bottomcoded variables and the methodology used in the coding.

Why Use the Internal User File (IUF)?

The AHS IUFs include most of the data collected on the AHS, but in order to maintain the privacy of the respondents, many items are not released on the PUF. are recoded and/or topcoded or bottomcoded (i.e., really high and really low values are coded to a single value) and the level of geography available is limited and some of it is masked. The following table shows the differences between the PUF and the IUF.

Public Use File (PUF)	Internal Use File (IUF)
Topcoded/Bottomcoded and Perturbed (Age) data to protect the privacy and confidentiality of respondents	Non-topcoded/bottomcoded data and age is not perturbed
Some variables, such as amenities available in the community, may be recoded to a single Yes/No variable	All individual variables collected in the survey are available
Other variables, such as some variables tracking details on why an unit was not interviewed, are not available	All variables in the interview are available
Merged/Masked geography	Detailed geography (including 1980 census tract)
Recoded/Collapsed information on home improvement	Detailed information on home improvement at the job-level
No information on sampling frames	Detailed information on sampling frames
In some metro areas some cases may be dropped for disclosure purposes, as the population count in certain areas may no longer meet Census' threshold for disclosure (post-Hurricane Katrina New Orleans, for	All cases are available
Not recoded/computed variables that are used to create the publication tables	A file including the recoded/computed variables used to create tables are included
No information from MASTER file	MASTER file which includes detailed information on sampling frame, geography, etc. is included

AHS IUF Research

- Author: Ingrid Ellen
- Title: How Low Income Neighborhoods Change: Entry, Exit, and Enhancement
- Need for IUF: Census tract identifiers needed to merge contextual level variables

21

NHANES IUF Research

- Author: C. Bird
- Title: Neighbourhood Socioeconomic Status and Biological 'Wear and Tear' in a Nationally Representative Sample of US Adults
- Need for IUF: Census tract identifiers needed to merge contextual level variables

22

NSFG IUF Research

- Author: M. Bitler
- Title: Health Disparities and Infertility: Impacts of State-Level Insurance Mandates
- Need for IUF: State IDs to merge in policy variables

23

Useful Links

- *Census Bureau's Center for Economic Studies*
– <http://www.census.gov/ces/>
- *Atlanta Census Research Data Center*
– <http://avspns.gsu.edu/acrdc/>
- *NCHS Data Center*
– <http://www.cdc.gov/rdcl/>
- *BLS Data Center*
– <http://www.bls.gov/bls/blsresda.htm>
- *Handbooks for ACS Data Users*
– http://www.census.gov/acs/www/guidance_for_data_users/handbooks/

24