INFO 7470/ECON 7400/ILRLE 7400
Measuring People and Households

Warren A. Brown
February 11, 2013

Outline

- Surveys and microdata files
- Research questions and access to restricted use data
- Census of Population and Housing
- American Community Survey
- Current Population Survey
- American Housing Survey
- Public Use vs Restricted Use

© Warren Brown 2013, all rights reserved

Outline

- **Surveys and microdata files**
- Research questions and access to restricted use data
- Census of Population and Housing
- American Community Survey
- Current Population Survey
- American Housing Survey
- Public Use vs Restricted Use

© Warren Brown 2013, all rights reserved

What are Microdata Files?

- Computerized versions of the questionnaire
- Individual records about the persons, or households

© Warren Brown 2013, all rights reserved

Microdata File

ID#	Sex	Age	Marital Status
34701	1	34	2
34702	2	33	2
34703	2	8	-1
82301	2	82	3
60201	1	47	4
11801	2	24	6
11802	1	24	6

© Warren Brown 2013, all rights reserved

Marital Status by Sex

	Total	Male	Female
Married	117	59	58
Previously Married	38	12	26
Never Married	122	64	58

© Warren Brown 2013, all rights reserved

Public Use vs Internal Use

- Anonymized
- Sub-Sample
- High level geography
- Top-coding
- Population thresholds
- Statistical disclosure limitation (SDL)

- Anonymized
- Full Sample
- Small area geography
- Full range
- No thresholds
- No SDL or only some of the SDL procedures applied

© Warren Brown 2013, all rights reserved

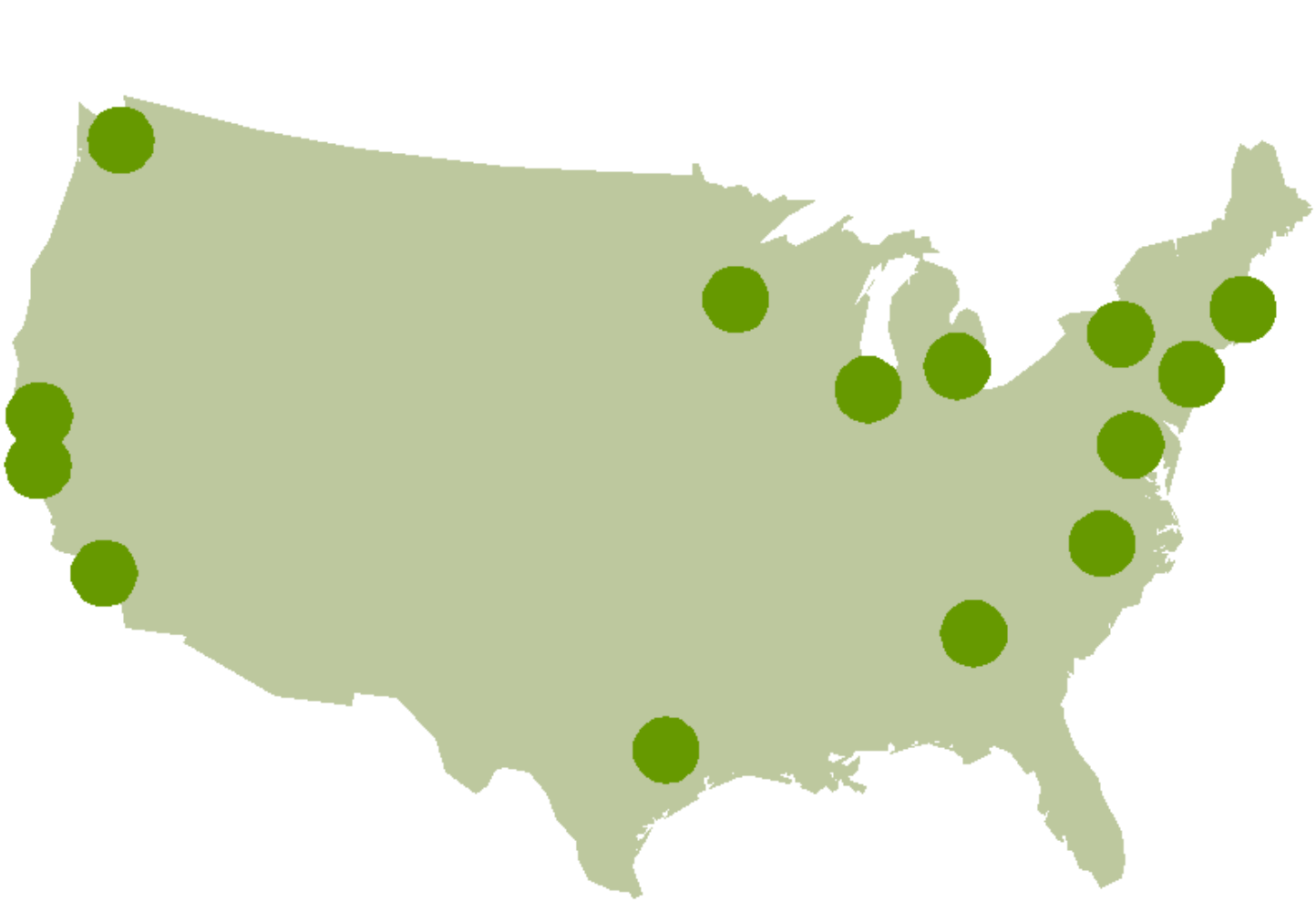
Access to Internal Use Files

- Census Research Data Center
  - Rigorous approval process
  - Remote access to files
  - Rigorous screening for disclosure avoidance (statistical disclosure limitation)

© Warren Brown 2013, all rights reserved

Census Research Data Centers

www.census.gov/ces/rdcresearch/



© Warren Brown 2013, all rights reserved

Documentation for Microdata Files

- Sponsor/Collector
- Purpose
- Universe/Target Population
- Sampling Frame
- Sample Design and Sample Size
- Mode of Administration
- Measurement Instrument
- Editing and Coding (including statistical disclosure limitations, if any)
- Processing and Missing Data
- Weighting
- Estimates and Control Totals

© Warren Brown 2013, all rights reserved

Codebook

- Structure of the file
  - Rectangular or hierarchical
  - Column locations and widths for variables
- Definitions of different record types
- Response codes for each variable
- Codes used to indicate nonresponse and missing data
- Tabulations of key variables for verification

© Warren Brown 2013, all rights reserved

Rectangular Data Structure

ID#	Sex	Age	Marital Status
34701	1	34	2
34702	2	33	2
34703	2	8	-1
82301	2	82	3
60201	1	47	4
11801	2	24	6
11802	1	24	6

© Warren Brown 2013, all rights reserved

Rectangular Data Structure

Housing Unit ID	Geog	HU Type	Poverty
347	36123	1	8
823	42310	2	3
602	13109	8	1
118	47003	9	4

© Warren Brown 2013, all rights reserved

Hierarchical Data Structure

HHld 1	Geog	HU Type	Poverty
Per 1	Age	Sex	Education
Per 2	Age	Sex	Education
Per 3	Age	Sex	Education
HHld 2	Geog	HU Type	Poverty
Per 1	Age	Sex	Education
Per 2	Age	Sex	Education

© Warren Brown 2013, all rights reserved

Sample Sizes

Survey	IUF	PUF
Census 2000S	16%	5%, 1%
ACS	2.5%	1%
CPS	0.1%	0.1%
AHS	0.1%	0.1%

© Warren Brown 2013, all rights reserved

Outline

- Surveys and microdata files
- **Research questions and access to restricted use data**
- Census of Population and Housing
- American Community Survey
- Current Population Survey
- American Housing Survey
- Public Use vs Restricted Use

© Warren Brown 2013, all rights reserved

CES Publications

www.census.gov/ces/publications/index.html

- Annual Research Report
- Discussion Paper Series
- Publications

© Warren Brown 2013, all rights reserved

Outline

- Surveys and microdata files
- Research questions and access to restricted use data
- **Census of Population and Housing**
- American Community Survey
- Current Population Survey
- American Housing Survey
- Public Use vs Restricted Use

© Warren Brown 2013, all rights reserved

Census of Population and Housing

Sponsor: Census Bureau

Collector: Census Bureau

Purpose: "The U.S. Census counts every resident in the United States. It is mandated by Article I, Section 2 of the Constitution and takes place every 10 years. The data collected by the decennial census determine the number of seats each state has in the U.S. House of Representatives and is also used to distribute billions in federal funds to local communities."

Census: Target Population

- All residents of the United States without regard to the person's legal status or citizenship
- Usual Place of Residence – "the place where a person lives and sleeps most of the time." – 6+ months of the year
- People who live at more than one residence during the week, month, or year should be counted at the place where they live most of the time.
- People without a usual residence, however, should be counted where they are staying on Census Day

© Warren Brown 2013, all rights reserved

Census: Sampling Frame
A Survey Wrapped Inside A Census

- Master Address File (MAF) for Census 2000
  - Official inventory of known living quarters
  - Linked to TIGER
- Based on:
  - 1990 Address Control File
  - Updates from the USPS' Delivery Sequence File
  - Field listing operations
  - Addresses supplied by local governments through partnership operations (LUCA)
- DMAF – operational version of the MAF used to control procedures

© Warren Brown 2013, all rights reserved

Census: Design of the Sample

- Every person and housing unit asked basic demographic and housing questions
- Sampling unit for Census 2000 was the housing unit, including all occupants
- Four different housing unit sampling rates: 1-in-8, 1-in-6, 1-in-4, and 1-in-2 (designed for an overall average of about 1-in-6)
- For people living in group quarters the sampling unit was the person and the sampling rate was 1-in-6

© Warren Brown 2013, all rights reserved

Census: Questionnaire(s)

- **Short Form** - Sent to approximately 5 of 6 households, that contains population questions related to household relationship, age, sex, race, Hispanic origin, and tenure (*i.e.*, whether home is owned or rented).
- **Long Form** - Contains all questions on the short form, as well as additional detailed questions relating to the social, economic, and housing characteristics of each individual and household. Information derived from the long form is referred to as sample data and is tabulated for geographic entities as small as the block group level.

© Warren Brown 2013, all rights reserved

Census: Data Collection

Residents of Traditional Housing Units

- **Mailout/mailback.** U.S. Postal Service delivered questionnaires to every ‘‘city style’’ housing unit with a street name and house number. Approximately 95% of population.
- **Update/leave.** Census enumerators delivered questionnaires to housing units without street names and house numbers to be mailed back, mainly in rural areas, and corrected and updated the address list and maps for any additions or errors.
- **List/enumerate.** In remote and sparsely populated areas, enumerators visited every housing unit and completed the enumeration as delivered.

© Warren Brown 2013, all rights reserved

Census: Data Collection

Residents **Not** in Traditional Housing Units

- **Group quarters enumeration -** advance visits to each group quarter followed up by distributing questionnaire packets.
- **Transient night operation -** at racetracks, campgrounds, fairs and carnivals, and marinas.
- **Remote Alaska enumeration**
- **Domestic military/maritime enumeration -** identified living quarters and housing units on military installations and ships assigned to a U.S. home port.
- **Overseas enumeration -** counted federal employees assigned overseas (including members of the armed forces) and their dependents, for apportionment purposes.
- **Homeless** - people with no usual residence or address. This operation was conducted at locations such as shelters and soup kitchens and non- sheltered outdoor locations.

© Warren Brown 2013, all rights reserved

Data Collection Timetable

- Once a decade, tied to April 1, 2000
- Questionnaires mailed out March 15
- Reminder postcards and replacement Q's mailed out
- Nonresponse Follow-Up (NRFU) with 100% on housing units not responding conducted April 27 through June 26.
- Closeout completed September 7.

© Warren Brown 2013, all rights reserved

Census: Response Rates

- Participation is mandatory
- Final Response Rate – 67.4% (9/2/2000)
- Mail Response Rate – 64.3% (4/18/2000)
  - Short-form mail response – 66.4%
  - Long-form mail response – 53.9%

© Warren Brown 2013, all rights reserved

Census: Editing

- Large Households – Q allows list up to 12 but only characteristics of 6. Census Edit Follow-Up.
- Household Relationship – consistency for age and sex
  - Only one spouse per householder
  - Parent of householder 15+ years older
- Same sex married couples
  - 1990 – kept relationship changed sex
  - 2000 – kept sex changed relationship

© Warren Brown 2013, all rights reserved

Census: Missing Data

Imputation

- Substitution
  - Household exists but no “data-defined” person records
  - Characteristics of a nearby household are assigned using “hot-deck” procedure
- Allocation
  - Assignment used other data on respondent
  - Allocation – hot deck procedure
  - Flags variables

© Warren Brown 2013, all rights reserved

Census: Allocation Flags

- Due to imputation, no missing data
- Allocation flags for individual variables on each records identify the type of allocation:
  - Pre-edit
  - Consistency
  - Hot deck
  - Cold deck

© Warren Brown 2013, all rights reserved

Edit/Allocation Flag for Age

Value	Label
0	Consistent as reported
1	Age only
2	Date of birth only
3	Inconsistent age and date of birth
4	Allocated from hot deck
7	Substituted

© Warren Brown 2013, all rights reserved

Census: Weighting

- Records assigned to Weighting Areas
  - Comprise contiguous geographic units w/i counties
  - Minimum of 400 sample units
  - Complete count data for area used as control
  - Controlled using ratio-adjustment (IPF)
- Persons:
  - Probability of selection (initial sampling rate)
  - Stage I: Type of Household
  - Stage II: Sampling Type
  - Stage III: Householder Status
  - Stage IV: Age/Sex/Race/Hispanic Origin
- Sum the weights
  - Single weight for person for all characteristics
  - Weights are integers – proportionally rounded
- Housing: Similar process

© Warren Brown 2013, all rights reserved

Census: Sampling Error

- Rough approximation/less computation
  - Use documentation tables for unadjusted standard errors
  - Adjust using design factors for variable and state
- More precise/more computation
  - Random groups method
  - Draw 100 sub-samples using 2-digit sub-sample number (00 to 99)
  - Compute estimate of variance

© Warren Brown 2013, all rights reserved

Census: PUF / IUF

- 1%, 5% Sample
- PUMA
- Top Coding
- Collapsed Categories
- Data Swapping
- Perturbation
- No Administrative

- 16 % Sample
- Block/Tract/Place
- Full Distribution
- Full Distribution
- Data Swapping (usually)
- Actual Ages
- Details of Collection

© Warren Brown 2013, all rights reserved

Census 2010H Microdata

- Short-form only, also referred to as 100% or H
  - [Long-form see ACS] ;-)
- Person: age, sex, race, Hispanic origin, hhld relationship
- Household: occupancy, tenure, household type
- Availability:
  - PUMS on a State by State flow Mar 2013 – Jun 2013

© Warren Brown 2013, all rights reserved

Outline

- Surveys and microdata files
- Research questions and access to restricted use data
- Census of Population and Housing
- **American Community Survey**
- Current Population Survey
- American Housing Survey
- Public Use vs Restricted Use

© Warren Brown 2013, all rights reserved

American Community Survey

Sponsor: Census Bureau

Collector: Census Bureau

Purpose: "The American Community Survey (ACS) is an ongoing survey that provides data every year -- giving communities the current information they need to plan investments and services. Information from the survey generates data that help determine how more than $400 billion in federal and state funds are distributed each year."

ACS: Purpose

- Key Component for Re-Engineered Census
- Alternative to the Decennial Census "Long-Form"
- Continuous measurement rather than snapshot
- Meet federal legislative and program needs
- Other stakeholders in state and local government and private sectors
- Provide annual data on demographic, social, economic and housing characteristics.
- Improve the Federal Statistical System

© Warren Brown 2013, all rights reserved

ACS: Target Population

- Current residence not “Usual Residence”
  - “two-month rule”
- Household population in 2005
- Total (HH + GQ) in 2006 and ….
- GQ residency is “de facto”

© Warren Brown 2013, all rights reserved

ACS: Sampling Frame

- Master Address File (MAF)
  - Official inventory of known living quarters
  - Linked to TIGER
- Housing Units
  - Based on Census 2000 MAF and updates from the USPS' Delivery Sequence File
- Group Quarters
  - … and updates from the administrative records and the FSCPE
  - Excluded from ACS are domestic violence shelters, soup kitchens, commercial maritime vessels,…

© Warren Brown 2013, all rights reserved

ACS: Design of the Sample

- Annual Sample Size of 3 million addresses
- Series of Monthly Samples of 250,000 addresses
- HU sample in each of the 3,141 Counties
- Areas with smaller populations sampled at higher rates than those with larger populations
- HU Address sampling rate set by Block based on entity (municipality, school district, tract)
- Final sampling rate varies between 1.6% and 10%
- No HU address can be sampled more than once in 5 years

© Warren Brown 2013, all rights reserved

ACS: Questionnaire

- Content designed to meet the needs of federal government agencies
- 21 housing and 48 population questions
- Household Respondent provides responses for all other residents of the household
- Householder or Reference Person is “Person 1”

© Warren Brown 2013, all rights reserved

ACS: Data Collection

- HU addresses by three modes
  - Mailout of paper questionnaire in 1st month
  - Telephone (CATI) non-response follow-up in 2nd month
  - Personal visit (CAPI) non-response follow-up in 3rd month to a sub-sample
- GQ
  - Personal visit within 6 weeks of sample selection

© Warren Brown 2013, all rights reserved

Data Collection Timetable

Month	2005		2006			
	November	December	January	February	March	April
November 2005	Mail	Mail Telephone	Mail Personal Visit			
December 2005		Mail	Mail Telephone	Mail Personal Visit		
January 2006			Mail	Mail Telephone	Mail Personal Visit	
February 2006				Mail	Mail Telephone	Mail Personal Visit
March 2006					Mail	Mail

© Warren Brown 2013, all rights reserved

Non-Response Follow Up

- 1st Month – Mailed Questionnaire
- 2nd Month – Attempt at Telephone Interview
- 3rd Month – Sub-Sample of Non-Respondents
  - 67% unmailable addresses
  - 50% low response tracts
  - 33% high response tracts

© Warren Brown 2013, all rights reserved

ACS: Response Rates

- Participation is mandatory
- Overall response rate is ~98%
- Sample addresses eligible for interviewing:
  - 55% Mail response
  - 59% CATI non-response follow-up
  - 96% CAPI non-response follow-up sub-sample
  - 2% Non-interview

© Warren Brown 2013, all rights reserved

ACS: Missing Data

Item non-response or failed edit check

- FEFU – failed-edit follow-up
  - More than 5 household members
  - Critical questions not answered
  - 33% of mail return questionnaires in 2005
- Imputation
  - Assignment used other data on respondent
  - Allocation – hot deck procedure
  - Flags variables

© Warren Brown 2013, all rights reserved

ACS: Weighting

- Person and Housing Unit weights
- Three stages
  - Probability of selection (initial sampling rate)
  - Adjust for non-response
  - Control to Population Estimates
- Sum the weights
  - Person weights for person characteristics
  - HU weights for family, household or housing unit characteristics

© Warren Brown 2013, all rights reserved

ACS: Sampling Error

- Publishing margins of error in tables
- More sample less error
- 1, 3, and 5 year estimates and associated sampling error
- PUF – records have replicate weights to more precisely estimate sampling error
  (
  link to University of Michigan Population Studies Center methods and programs)
- IUF – no replicate weights currently

© Warren Brown 2013, all rights reserved

PUF / IUF

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

- 2.5% Sample
- Block/Tract/Place
- Full Distribution
- Full Distribution
- Actual Ages
- Details of Collection
- Swapped (usually)

© Warren Brown 2013, all rights reserved

Outline

- Surveys and microdata files
- Research questions and access to restricted use data
- Census of Population and Housing
- American Community Survey
- **Current Population Survey**
- American Housing Survey
- Public Use vs Restricted Use

© Warren Brown 2013, all rights reserved

Current Population Survey

Sponsor: Bureau of Labor Statistics

Collector: Census Bureau

Purpose: Monthly Data for Analysis of Labor Market Conditions. CPS is **Household** survey measuring unemployment rate. It "serves as part of model-based estimates for individual states and other geographic areas"

Secondary Purpose

- Supplemental studies on other topics
- Census Bureau partners with other federal agencies to gather data on:
  - Health Insurance Coverage
  - Food Security
  - Computer Usage
  - Fertility and Childbearing

© Warren Brown 2013, all rights reserved

Advantages of the CPS

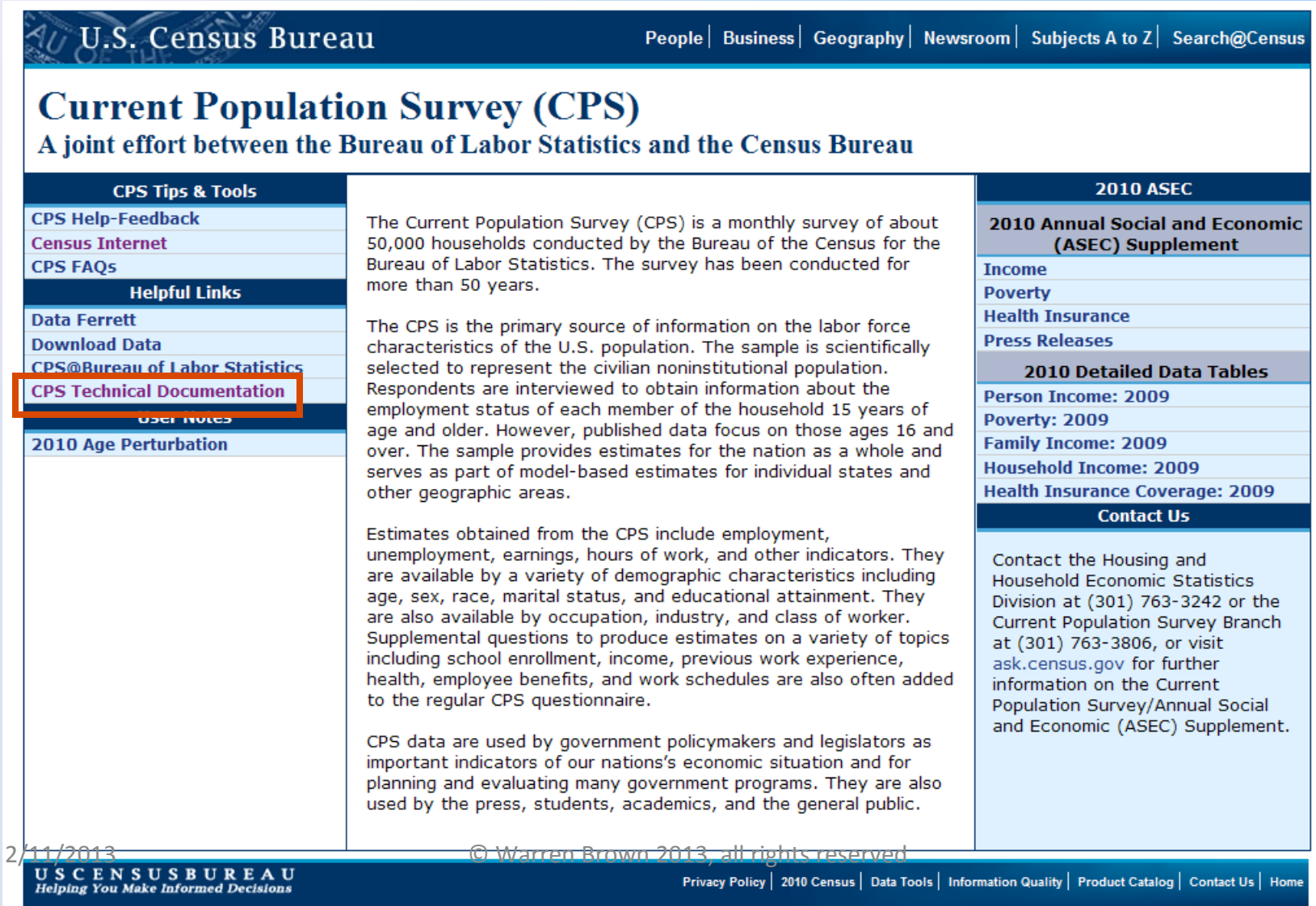
- Timeliness – Quick Turnaround
- Long History
- Expert Interviewers
- Topical Modules

© Warren Brown 2013, all rights reserved

www.bls.gov/cps/

© Warren Brown 2013, all rights reserved

www.census.gov/cps/



U.S. Census Bureau

People | Business | Geography | Newsroom | Subjects A to Z | Search@Census

Current Population Survey (CPS)

A joint effort between the Bureau of Labor Statistics and the Census Bureau

CPS Tips & Tools

- CPS Help-Feedback
- Census Internet
- CPS FAQs

Helpful Links

- Data Ferrett
- Download Data
- CPS@Bureau of Labor Statistics
- CPS Technical Documentation

User Notes

- 2010 Age Perturbation

The Current Population Survey (CPS) is a monthly survey of about 50,000 households conducted by the Bureau of the Census for the Bureau of Labor Statistics. The survey has been conducted for more than 50 years.

The CPS is the primary source of information on the labor force characteristics of the U.S. population. The sample is scientifically selected to represent the civilian noninstitutional population. Respondents are interviewed to obtain information about the employment status of each member of the household 15 years of age and older. However, published data focus on those ages 16 and over. The sample provides estimates for the nation as a whole and serves as part of model-based estimates for individual states and other geographic areas.

Estimates obtained from the CPS include employment, unemployment, earnings, hours of work, and other indicators. They are available by a variety of demographic characteristics including age, sex, race, marital status, and educational attainment. They are also available by occupation, industry, and class of worker. Supplemental questions to produce estimates on a variety of topics including school enrollment, income, previous work experience, health, employee benefits, and work schedules are also often added to the regular CPS questionnaire.

CPS data are used by government policymakers and legislators as important indicators of our nations's economic situation and for planning and evaluating many government programs. They are also used by the press, students, academics, and the general public.

2010 ASEC

2010 Annual Social and Economic (ASEC) Supplement

- Income
- Poverty
- Health Insurance
- Press Releases

2010 Detailed Data Tables

- Person Income: 2009
- Poverty: 2009
- Family Income: 2009
- Household Income: 2009
- Health Insurance Coverage: 2009

Contact Us

Contact the Housing and Household Economic Statistics Division at (301) 763-3242 or the Current Population Survey Branch at (301) 763-3806, or visit ask.census.gov for further information on the Current Population Survey/Annual Social and Economic (ASEC) Supplement.

USCENSUSBUREAU
Helping You Make Informed Decisions

Privacy Policy | 2010 Census | Data Tools | Information Quality | Product Catalog | Contact Us | Home

© Warren Brown 2013, all rights reserved

CPS: Target Population

- US Residents
- Civilian
- Non-Institutional

© Warren Brown 2013, all rights reserved

CPS: Sampling Frame

- National household frame based on the most recent Census of Population and Housing
- Coverage during the decade augmented by lists of building permits

© Warren Brown 2013, all rights reserved

CPS: Design of the Sample

- Monthly Sample Size of 72,000 addresses
- Sample expanded in 2001 to meet needs of SCHIP
- State based sample design
- 1st Stage - 754 Primary Sampling Units (PSUs) chosen from US total of 2,007
- 2nd Stage - Clusters of 4 HUs selected
- 60,000 HUs eligible (occupied, usual residence)
- Yields 110,000 individuals aged 16+
- Household Rotation 4 months in – 4 out – 4 in

© Warren Brown 2013, all rights reserved

CPS: Sample Size

- Sample size is “mandated” by Congress
- At 6 % National Unemployment Rate
- +/- 0.2 % must be significant (at 10% level)

© Warren Brown 2013, all rights reserved

CPS: ASEC (March) Items

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

© Warren Brown 2013, all rights reserved

CPS: Data Collection

- Personal Interview (CAPI) in Months 1 and 5
- Telephone Interview (CATI) in Months 2-4 and 6-8
- Questions are asked of the first in-scope adult encountered in the household
- The respondent provides proxy responses for all other in-scope residents of the household

© Warren Brown 2013, all rights reserved

Non-Response

- Non-Interview Households
  - 7- 8 % of Sample
- Item Non-Response
  - 0.5 % - household relationships
  - 2.0 % - demographic characteristics
  - 2.4 % - labor force status
  - 10.0 % - industry and occupation
  - 18.5 % - earnings

© Warren Brown 2013, all rights reserved

CPS: Editing and Coding

- Industry & Occupation codes—4 digit—are assigned by expert coders at the Census Bureau's National Processing Center
- Missing items are identified in the edit process and handled with various imputation processes.

© Warren Brown 2013, all rights reserved

CPS: Missing Data

Census Bureau Imputation Procedures:

- Relational Imputation
- Longitudinal Edit
- Hot Deck Allocation Procedure

© Warren Brown 2013, all rights reserved

“Hot Deck” Allocation

Labor Force Status

- Employed
- Unemployed
- Not in the Labor Force

© Warren Brown 2013, all rights reserved

Weighting Procedures

- First-stage ratio adjustment
- Differential selection probabilities
- Adjust for nonresponse
- Reduce variance

© Warren Brown 2013, all rights reserved

First-Stage Ratio Adjustment

$$\frac{\text{Stratum Population Total from Frame}}{\text{(Pop of Selected PSU)}}$$

$$\frac{\text{(Pop of Selected PSU)}}{\text{Probability of Selecting PSU}}$$

First-Stage Ratio Adjustment

Pop Total	2,000,000	2,000,000	2,000,000
Pop PSU	50,000	65,000	40,000
Prob Select	0.025	0.025	0.025
Weight	1	1.25	0.77

Selection Weight

Pop Total	199,500,000	24,937,500	174,562,500
Pop PSU	125,000	62,500	62,500
Prob Select	1/1,596	1/399	1/2,793
Weight	1,596	399 or 1	2,793 or 7

Nonresponse Weight

Age	Sample	Response Rate	Nonresponse Adjustment Weight
12-44	31,250	80%	1.25
45+	31,250	88%	1.14

Tallying Weights

Latino 12-44	1.25	1	1.25
Latino 45+	1.14	1	1.14
Non-Latino 12-44	1.25	7	8.75
Non-Latino 45+	1.14	7	7.98

Poststratification Weight

Category	Population	Sample	Post-Strat Weight
Female	52%	50%	0.96
Male	48%	50%	1.04

The Final Weight

First-Stage Ratio Adjustment Weight

X

Selection Weight

X

Nonresponse Weight

X

Poststratification Weight

© Warren Brown 2013, all rights reserved

Poststratification Controls

For the CPS they are:

- Census Bureau estimates of state population by:
  - Age
  - Sex
  - Race
  - Hispanic Origin

© Warren Brown 2013, all rights reserved

PUF / IUF

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

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

© Warren Brown 2013, all rights reserved

Outline

- Surveys and microdata files
- Research questions and access to restricted use data
- Census of Population and Housing
- American Community Survey
- Current Population Survey
- **American Housing Survey**
- Public Use vs Restricted Use

© Warren Brown 2013, all rights reserved

American Housing Survey

Sponsor: Housing and Urban Development

Collector: Census Bureau

Purpose: Household survey of housing conditions and housing markets. More detailed housing data than Census. Analyzing the flow of households through housing

AHS: From the interview protocol…

This survey provides information about:

- The cost of housing for the American public.
- The availability of housing for different segments of the American public, such as the elderly, low income, and different minority groups.
- Trends in housing market. For example:
  - What type of housing (single family homes, condominiums, townhouses, manufactured/mobile homes, etc.) are people buying?
  - What type of fuels and appliances are being used?
  - Is the quality of housing in this country improving or getting worse for most people?
  - Is the cost of owning and maintaining a home changing?
  - Is the cost of renting changing?

© Warren Brown 2013, all rights reserved

AHS: Target Population

- All Housing Units
  - Apartments
  - Single-family homes
  - Mobile homes
  - Assisted living units
  - Vacant units
- 50 States and District of Columbia

© Warren Brown 2013, all rights reserved

AHS: Sampling Plan

- Stratified multistage probability sample of housing units
- Use Frame and Primary Sampling Units (PSU) developed for CPS
- Since 1985, use sample selected from 1980 Census
- Same core sample since 1985, providing a panel on flow of households through housing
- Additions of new housing units follows CPS procedures

© Warren Brown 2013, all rights reserved

AHS: Sample Characteristics

- National sample had covered on average 55,000 housing units
- In 2011 sample increased to 186,000 units
- Each metropolitan area sample now covers 4,500 housing units
- Weights derived by controlling survey to independent Census Bureau estimates of housing units, based on Census 2000

© Warren Brown 2013, all rights reserved

AHS: Sample Characteristics pre-2011

National Sample

- Odd-numbered years
- Throughout the country
- (Includes the 6 largest cities -- every other time)
- Chicago, Detroit, Los Angeles, New York, Northern New Jersey, Philadelphia
- About 50,000 homes

41 Metro Areas

- Even-numbered years
- on a rotating basis
- about 13 each year
- About 3,000 homes per metro

© Warren Brown 2013, all rights reserved

AHS: Sample Characteristics 2011

- National and Metro Area Samples Combined
- 186,000 housing units
- Odd-numbered years
- Over-samples for 29 metro areas
- Separate weights for national and metro area estimates

© Warren Brown 2013, all rights reserved

AHS: Coverage Issues

- Units constructed without permits in permit-issuing areas
  - New Structure
  - Conversions/Mergers
- Mobile Homes

© Warren Brown 2013, all rights reserved

AHS: Data Collection pre-2011

- National data - odd years
- Metro data – even years
  - 47 Metropolitan Areas
  - 6 year rotation
  - 6 largest metro areas every 4 years
- Data collection begins late spring
- Lasts 3-7 months

© Warren Brown 2013, all rights reserved

AHS: Data Collection

- Personal visit or telephone interview using laptops.
- Questions are asked of any knowledgeable adult (16+ years) household member
- If vacant then questions asked of landlord, rental agent or knowledgeable neighbor

© Warren Brown 2013, all rights reserved

AHS: Data Collection

- Content areas include:
  - Household
  - Housing Unit
  - Equipment and Facilities
  - Housing costs
  - Housing and Neighborhood Quality
  - Previous residence for recent movers
  - Alterations and Replacements

© Warren Brown 2013, all rights reserved

Subjects Included in the American Housing Survey

Housing Units and Householders

- Citizenship of householder
- Cooperatives and condominiums
- Duration of vacancy
- Hispanic origin
- Housing units
- Last used as a permanent residence
- Occupied housing units
- Owner or manager on property
- Population in housing units
- Previous occupancy
- Race
- Rental vacancy rate
- Seasonal units
- Suitability for year−round use
- Tenure
- Vacancy status
- Vacant housing units
- Year householder immigrated to the United States
- Year householder moved into unit
- Year−round housing units

Household Characteristics

- Adults and single children under 18 years old
- Age of householder
- Elderly (age 65 and over)
- Family or primary individual
- Household composition by age of householder
- Household moves and formation
- Nativity
- Nonrelative
- Other relative of householder
- Own never−married children under 18 years old
- Persons other than spouse or children
- Single children under 18 years old
- Subfamily
- Years of school completed by householder

© Warren Brown 2013, all rights reserved

Housing Units Occupied by Recent Movers

- Change in housing costs
- Choice of present home and home search
- Choice of present neighborhood and neighborhood search
- Location of previous unit
- Person's previous residence
- Present and previous units
- Previous home owned or rented by someone who moved here
- Reasons for leaving previous unit
- Recent mover comparison to previous home
- Recent mover comparison to previous neighborhood
- Recent movers
- Structure type of previous residence
- Tenure of previous unit

Plumbing Characteristics

- Complete bathrooms
- Flush toilet and flush toilet breakdowns
- Plumbing facilities
- Sewage disposal and sewage disposal breakdowns
- Source of water and water supply stoppage

Utilization Characteristics

- Bedrooms
- Lot size
- Persons per room
- Persons
- Rooms
- Square feet per person
- Square footage of unit

Structural Characteristics

- Common stairways
- Elevator on floor
- External building conditions
- Foundation
- Light fixtures in public halls
- New construction
- Site placement
- Stories in structure
- Stories between main and apartment entrances
- Units in structure
- Water leakage during last 12 months
- Year structure built

© Warren Brown 2013, all rights reserved

Equipment and Fuels

- Cooling degree days
- Electric fuses and circuit breakers
- Fuels
- Heating degree days
- Heating equipment and heating equipment breakdowns
- Kitchen appliances

Housing and Neighborhood Quality

- Age of other residential buildings within 300 ft
- Bars on windows of buildings
- Cars and trucks available
- Condition of streets
- Description of area within 300 feet
- Gated communities
- Mobile homes in group
- Moderate physical problems
- Nearby buildings vandalized or with interior exposed
- Neighborhood conditions
- Overall opinion of structure
- Overall opinion of neighborhood
- Selected deficiencies
- Selected amenities
- Severe physical problems
- Trash, litter, or junk on streets or on any properties

Financial Characteristics

- Amount of savings and investments
- Annual taxes paid per $1,000 value
- Cash received in primary mortgage
- Condominium and cooperative fee
- Cost and ownership sharing
- Current line−of−credit
- Current total loan as percentage of value
- Current interest rate
- First−time owners
- Food stamps
- Households with home equity loans
- Income
- Items included in primary mortgage payment
- Lenders of primary and secondary mortgages
- Lower cost State and local mortgages
- Major source of down payment
- Median monthly housing costs for owners
- Monthly housing costs
- Monthly payment for principal and interest
- Monthly housing costs as percentage of income
- Mortgage origination
- Mortgages currently on property
- Other activities on property
- Other housing costs per month
- Payment plans of primary and secondary mortgages

© Warren Brown 2013, all rights reserved

Financial Characteristics

- Primary mortgage
- Property insurance
- Purchase price
- Real estate taxes
- Reason refinanced
- Remaining years mortgaged
- Rent reductions
- Rent paid by lodgers
- Repairs, improvements, alterations in last 2 yrs
- Routine maintenance last year
- Term of primary mortgage at origination or assumption
- Total outstanding principal amount
- Type of primary mortgage
- Use of cash from refinancing
- Value
- Value−income ratio
- Year unit acquired
- Year primary mortgage originated

Alterations and Replacements

- Adding or replacing: roof, siding, doors, windows, insulation, pipes, plumbing, fixtures, wiring
- Additions: inside bathrooms, kitchens, bedrooms, other
- Changing unfinished areas to finished rooms: bedrooms, baths, other
- Installing: carpet, flooring, paneling
- Remodeling: bathrooms and kitchens
- Renovations: bedrooms, bathrooms, other
- Repairs due to a major disaster
- Replacing or adding outside: driveway, fence, patio, pool, shed
- Replacing or adding inside: central air, heating equipment, septic, water heater, dishwasher, disposal

© Warren Brown 2013, all rights reserved

Neighborhood Quality and Problems

- 68.7 % of households rate their neighborhood 8 or better, out of 10.
- 6.7% live in gated communities (walls or fences).
- Whether neighborhood has crime, odors, noise, litter, and whether it is bothersome.
- Satisfaction with public transportation, shopping, schools, police.
- Neighborhood buildings abandoned or with bars on windows.

© Warren Brown 2013, all rights reserved

AHS: Non-Response

- Non-Interviews
  - Type A – eligible but refused, not home or language (10 % of sample)
  - Type B – not eligible for interview but may be in future. Currently non-residential, vacant mobile home site, unit under construction
  - Type C – not eligible for sample. Demolished, relocated, building permit but never built
- Use variable “STATUS” to identify

© Warren Brown 2013, all rights reserved

AHS: Non-Response

- Item Non-Response
  - Answer not known by respondent
  - Respondent refuses to answer

© Warren Brown 2013, all rights reserved

AHS: Missing Data

- Critical items imputed using the cold or hot deck procedure
- Non-critical items coded with reason missing

© Warren Brown 2013, all rights reserved

AHS: “Cold Deck” Allocation

Tenure

- Randomly distributed to meet known distribution

© Warren Brown 2013, all rights reserved

AHS: “Hot Deck” Allocation

- Number of rooms is imputed from

  Tenure, vacancy, and # of floors

- Allocation Flag

  0 = unchanged

  1 = edited

  2 = allocated

© Warren Brown 2013, all rights reserved

AHS: Weighting

- Housing Unit weights for National Sample
- Product of 9 factors dealing with
  - Probability of selection (initial sampling rate)
  - Adjust for non-response
  - Adjust for non-self-representing PSUs
  - Adjust for new and old construction
  - Control to Census Bureau HU Estimates

© Warren Brown 2013, all rights reserved

AHS: Geographic Areas

United States

- Regions
- Metropolitan statistical areas
  - Central cities
  - Suburbs
- Urbanized areas
- Rural areas
- Places grouped by size
- IUF has id's for Tracts

© Warren Brown 2013, all rights reserved

Working with Longitudinally Linked AHS Data

- Same as CPS, the physical housing unit, and not the household is the object of the sampling plan
- A public use control number, scrambled from the master file control number, can be used to match record of housing unit over time

© Warren Brown 2013, all rights reserved

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/portal/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 (*till 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 PUFs 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: PUF / IUF

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

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

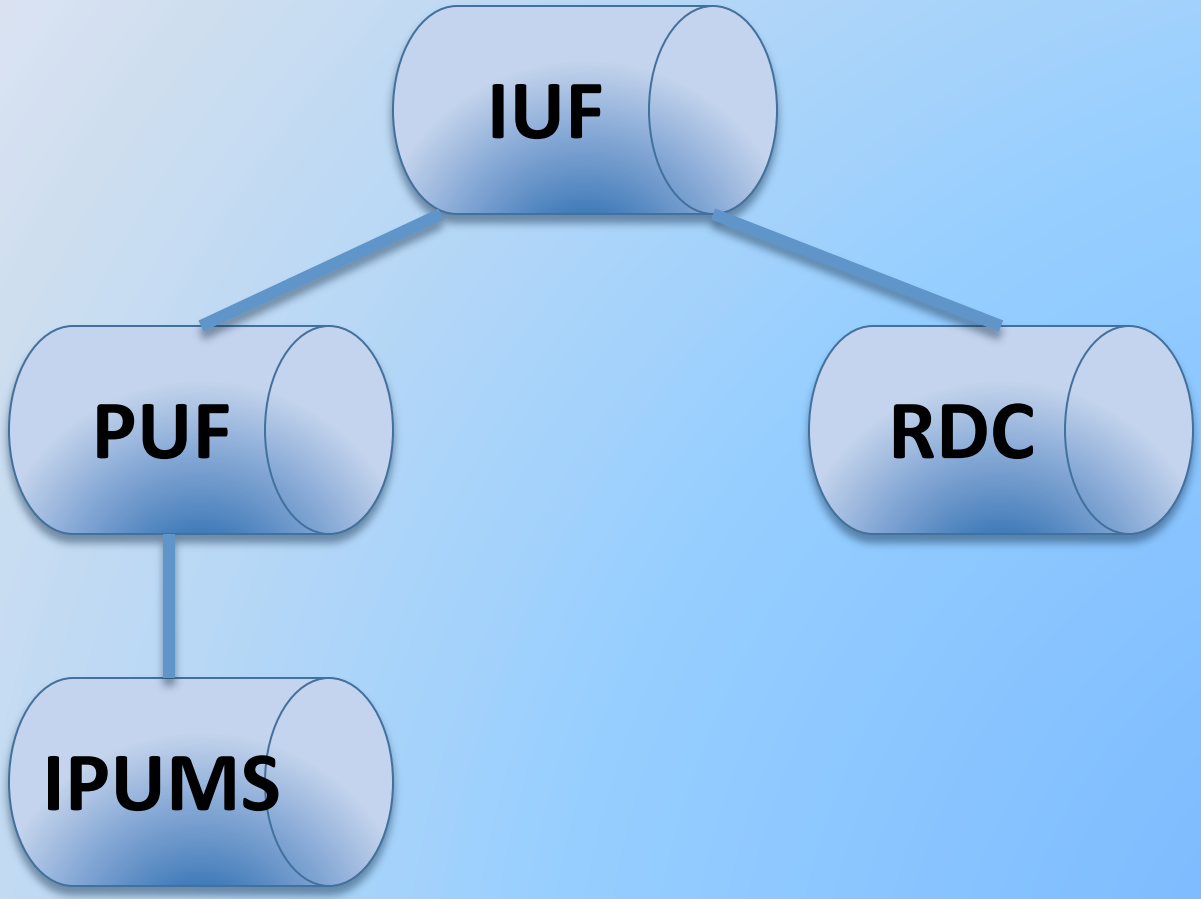
© Warren Brown 2013, all rights reserved

Outline

- Surveys and microdata files
- Research questions and access to restricted use data
- Census of Population and Housing
- American Community Survey
- Current Population Survey
- American Housing Survey
- **Public Use vs Restricted Use**

© Warren Brown 2013, all rights reserved

ACS Microdata Files Provenance



© Warren Brown 2013, all rights reserved

ACS PUF / IUF-RDC

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

- 2.5% Sample
- Block/Tract/Place
- Full Distribution
- Full Distribution
- Actual Ages
- Details of Collection
- Swapping (usually)

© Warren Brown 2013, all rights reserved

Census: PUF / IUF

- 1%, 5% Sample
- PUMA
- Top Coding
- Collapsed Categories
- Data Swapping
- Perturbation
- No Administrative

- 16 % Sample
- Block/Tract/Place
- Full Distribution
- Full Distribution
- Data Swapping (usually)
- Actual Ages
- Details of Collection

© Warren Brown 2013, all rights reserved

Census: PUF / IUF

Top Coding	Full Distribution
AGE	
0 to 90+ years	0 to 115+ years
INCOME	
$0 to $175,000+	$0 to $999,999+

© Warren Brown 2013, all rights reserved

Census: PUF / IUF

Collapsed Categories	Full Distribution
REAL ESTATE TAXES PAID	
$0 to $9,150+	$0 to $22,500+
In 67 categories	In single dollars

© Warren Brown 2013, all rights reserved

Census: PUF / IUF

No Administrative

RESPONSE COLLECTION MODE

N/A

Details of Collection

1 Paper from mail/enumerator

2 CATI from TQA

3 IDC (Internet)

4 CEFU

5 Transcribed from Experimental

© Warren Brown 2013, all rights reserved