

Household Samples

John M. Abowd
March 2005

Outline

- The Current Population Survey
- The American Housing Survey

CPS Overview

- Nature of the survey
- Sampling frame and plan
- Sample characteristics
- Working with the monthly data
- Working with longitudinally linked CPS data
- Toolkits, prepared extracts, etc.
- RDC files and use

Nature of the CPS

- Jointly sponsored and managed by the Bureau of Labor Statistics and the Census Bureau
- Official household survey for measuring the unemployment rate
- Produces more than 25,000 time series generically called the “Labor Force Statistics” from the CPS
- Produces more than 250,000 household-based time series with characteristics of the households and persons in the U.S. civilian non-institutional population

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
- Independent samples in each state (two independent samples in NY and CA)
- Multistage probability sample

Sampling Plan

- Defining the Primary Sampling Units (PSU)
 - Metropolitan Statistical Areas, County, Group of contiguous Counties
- Stratification of the PSUs
 - Self-representing (sampled with probability 1)
 - Non-self-representing (sampled with probability proportional to estimated population)
- Ultimate sampling units (USU)
 - Clusters of approximately 4 households sampled from housing lists based on the Decennial Census refreshed by building permits

Rotation Groups

- A sampled household is interviewed in 4 consecutive months, not interviewed for 8 consecutive months, then interviewed for 4 consecutive months
- Every sampled household is interviewed in the same calendar months of two consecutive years
- 75% of the sample households are common between two consecutive months
- The households interviewed in the 4th and 8th month in sample are referred to as the “outgoing rotation groups”

Sample Characteristics

- Approximately 60,000 sampled households each month (varies according to budget)
- Approximately 150,000 individuals age 16+
- Weights are provided that force the CPS estimates to match the Census Bureau intercensal population estimates

Questionnaire

- Computer-assisted personal interview (CAPI)
- 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

Working With the Monthly Data

- Time series from the BLS web site
- Micro data from a variety of collections
 - Official BLS/Census web site
 - ICPSR
 - NBER
 - University of California (Berkeley, David Card)
 - Unicon

Working with Longitudinally Linked CPS Data

- The physical domicile, and not the persons found there-in, is the object of the sampling plan
- The same household (physical domicile) is interviewed on the 4-8-4 sampling plan
- When there are new residents in the domicile, they are in-scope
- When the same residents are found in the domicile, they are in-scope and constitute a longitudinal sample

Linking the CPS records

- The public use sample identifier allows the creation of longitudinal links across the households
- Person links are trickier
 - Edited line number refers to the same person (since 1994 re-design)
 - If there are new residents in the household the edited line number may be re-used
 - Use age, sex and other demographic characteristics to verify

Tools

- NBER merged outgoing rotation groups
- ICPSR and Card extractor programs
- Unicon CPS support tools

RDC Files and Use

- March internal files available
- Full geography, no top coding
- Internal identifier and individual identifier linking possible
- Same naming structure as the public use file

AHS Overview

- Nature of the AHS
- Sampling plan
- Sample characteristics
- Questionnaire
- Working with longitudinally linked CPS data
- Tools



American Housing Survey

Nature of the AHS

- Conducted by the Bureau of the Census for the Department of Housing and Urban Development (HUD) Official
- Household survey of housing conditions and housing markets
- National data are collected in odd numbered years
- Data for each of 47 selected Metropolitan Areas are collected currently about every six years
- Data on 6 largest metro areas collected every four years

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

Sample Characteristics

- National sample covers on average 55,000 housing units
- Each metropolitan area sample covers 4,100 or more housing units
- Weights derived by controlling survey to independent estimates of housing units, based on Census 2000

Questionnaire

- Computer-assisted personal interview (CAPI) and telephone interview (CATI)
- Questions are asked of any knowledgeable adult (16+ years) household member
- Content areas include:
 - Household
 - Housing Unit
 - Equipment and Facilities
 - Housing costs

Missing Data

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

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

Tools

- HUDUSER for public use data, documentation, and “file flattener”
- DataFerrett extracts and tabulations

RDC Files and Use

- Many AHS samples available on the RDC network
- No suppressions
- Full geography
- Files very similar to public use files