Census Data for Transportation Planning

Transitioning to the
American Community Survey

May 11, 2005
Irvine, CA

Design Origins and Early Proposals

- Concept of "rolling sample design"

- Mid-decade census

- Proposed "Decade Census Program"

- Continuous measurement alternatives to the Census 2000 long form (ACS Design)

Consultations

- Congressional briefings
- Technical workshops
- User meetings (COMSIS '94)
- Advisory committee and oversight group meetings
- Information meetings with federal government agencies (DOT first agency)

Consultations

- Resulted in additional changes to basic design

- Cost and accuracy tradeoffs debated (Benefit of current data versus slightly higher variances)

Full implementation

- Expanded to full sample in 2005

- ACS is in every county stateside

- ACS is in every municipio in Puerto Rico

- Plan to delay expansion until 2006 to include population living in places other than housing units (group quarters population)

- Critical for the ACS to continue to monitor quality and operational efficiency

Current Design

- Sample design
- Data collection and capture
- Data processing
- Weighting and estimation
- Data products

Sample Design

- Survey designed to include
  - U.S. Stateside and Puerto Rico
  - Population in both housing units and group quarters (group quarters delayed until 2006)

- Survey designed to produce annually updated single-year and multi-year estimates

Sample Design

- Initial sampling rate
    - about 2.5% each year
    - about 12.5% over 5-year period

- Results in an initial sample of about
    - 250,000 addresses each month
    - 3 million addresses each year
    - 15 million addresses over 5-year period

Sample Design

- Variable sampling rates are used to ensure sufficient sample sizes in the smallest governmental units

- Initial sampling rates range from
  - about 1.7% to about 10% each year
  - about 8.5% to about 50% over a 5-year period

USCENSUSBUREAU
Helping You Make Informed Decisions

9

Sample Design

- About 2.2 million interviewed units expected per year due to
  - subsampling prior to personal visit interviewing
  - elimination of commercial or nonexistent addresses from initial address sample
  - noninterviews

Sample Design

- Sample is cumulated over TIME to produce lowest levels of geographic detail to replace census sample

- Five years of data are required for areas with less than 20,000 population such as Traffic Analysis Zones

Sample Design

- Sample is cumulated over SPACE to produce:

    – 3-year estimates for areas with populations of 20,000 or greater

    – single-year estimates for areas with populations of 65,000 or greater

USCENSUSBUREAU
Helping You Make Informed Decisions

12

Sample Design

- We project that the estimates of sampling error for the 5-year ACS estimates will be about 1/3 higher than those from decennial census sample estimates

- Deemed a reasonable trade-off relative to the functional gain associated with obtaining annually updated estimates throughout the decade

Sample Design
Frame

- Sample cases selected from an updated Master Address File (MAF)

- MAF updated through the use of…
  – Postal Service updates in most areas
  – Special field updating in more rural areas

USCENSUSBUREAU
Helping You Make Informed Decisions

ACS Content

- Identical to the Census 2000 long form

- Formal process working with OMB and all Federal agencies to revise ACS content for 2008-2013 period

- DOT decided not to recommend changes to the Journey to work questions for 2008

Transportation Section

24 **At what location did this person work LAST WEEK?** *If this person worked at more than one location, print where he or she worked most last week.*

a. Address (Number and street name)

If the exact address is not known, give a description of the location such as the building name or the nearest street or intersection.

b. Name of city, town, or post office

c. Is the work location inside the limits of that city or town?

☐ Yes
☐ No, outside the city/town limits

d. Name of county

e. Name of U.S. state or foreign country

f. ZIP Code

25 **How did this person usually get to work LAST WEEK?** *If this person usually used more than one method of transportation during the trip, mark (X) the box of the one used for most of the distance.*

☐ Car, truck, or van
☐ Bus or trolley bus
☐ Streetcar or trolley car
☐ Subway or elevated
☐ Railroad
☐ Ferryboat
☐ Taxicab

☐ Motorcycle
☐ Bicycle
☐ Walked
☐ Worked at home → *SKIP to question 33*
☐ Other method

I *Answer question 26 ONLY IF you marked "Car, truck, or van" in question 25. Otherwise, SKIP to question 27.*

26 **How many people, including this person, usually rode to work in the car, truck, or van LAST WEEK?**

Person(s)

27 **What time did this person usually leave home to go to work LAST WEEK?**

Hour   Minute
☐ a.m.
☐ p.m.

28 **How many minutes did it usually take this person to get from home to work LAST WEEK?**

Minutes

USCENSUSBUREAU
Helping You Make Informed Decisions

Data Collection and Capture

- Data are collected as of the date of interview using a current residence rule

- Data are collected throughout the entire year

- Survey participation is Mandatory

Data Collection and Capture

- Methodology based on best practices from decennial census and demographic surveys

- Monthly samples using overlapping multi-mode data collection methods
  - Mail
  - Telephone
  - Personal Visit

Data Collection
Sample Panels

Sample Panel	Calendar Month				
	Jan 2005	Feb 2005	Mar 2005	Apr 2005	May 2005
Nov 2004	Personal Visit				
Dec 2004	Phone	Personal Visit			
Jan 2005	Mail	Phone	Personal Visit		
Feb 2005		Mail	Phone	Personal Visit	
Mar 2005			Mail	Phone	Personal Visit

Data Collection
Monthly

Sample Panel	Calendar Month				
	Jan 2005	Feb 2005	Mar 2005	Apr 2005	May 2005
Nov 2004	Personal Visit				
Dec 2004	Phone	Personal Visit			
Jan 2005	Mail	Phone	Personal Visit		
Feb 2005		Mail	Phone	Personal Visit	
Mar 2005			Mail	Phone	Personal Visit

Data Collection and Capture
Mail

- Four mailings used to maximize mail response

- Over 95% of sample universe eligible for mailout

- Mail response rates average about 50% each month (Similar to Census 2000 mail response for long form)

USCENSUSBUREAU
Helping You Make Informed Decisions

Data Collection and Capture
Mail

- Mailout in one language
  - U.S. stateside English
  - Puerto Rico Spanish

- Language forms available upon request

- Telephone assistance provided in English and Spanish

Data Collection and Capture
Mail

- Current capture method involves keying data from paper questionnaires

- Research planned to convert to image capture and key-from-image technology

- Data are reviewed for completeness with telephone follow-up to resolve missing and inconsistent responses

23

Data Collection and Capture
Telephone

- About 5 weeks after the initial mailout, most mail returns have been received

- The nonresponse workload is identified for telephone follow-up

- Commercial vendors provide telephone numbers

USCENSUSBUREAU
Helping You Make Informed Decisions

Data Collection and Capture
Telephone

- 3 call centers conduct interviews using computer-assisted methods

- Telephone follow-up lasts about four weeks

- Survey instruments in English and Spanish

Data Collection and Capture
Personal Visit

- Two universes for personal visit followup
  - nonrespondents to the mailout and telephone
  - cases ineligible for mailout

- Subsample selected

- Personal visit followup conducted out of our 12 regional offices

Data Collection and Capture
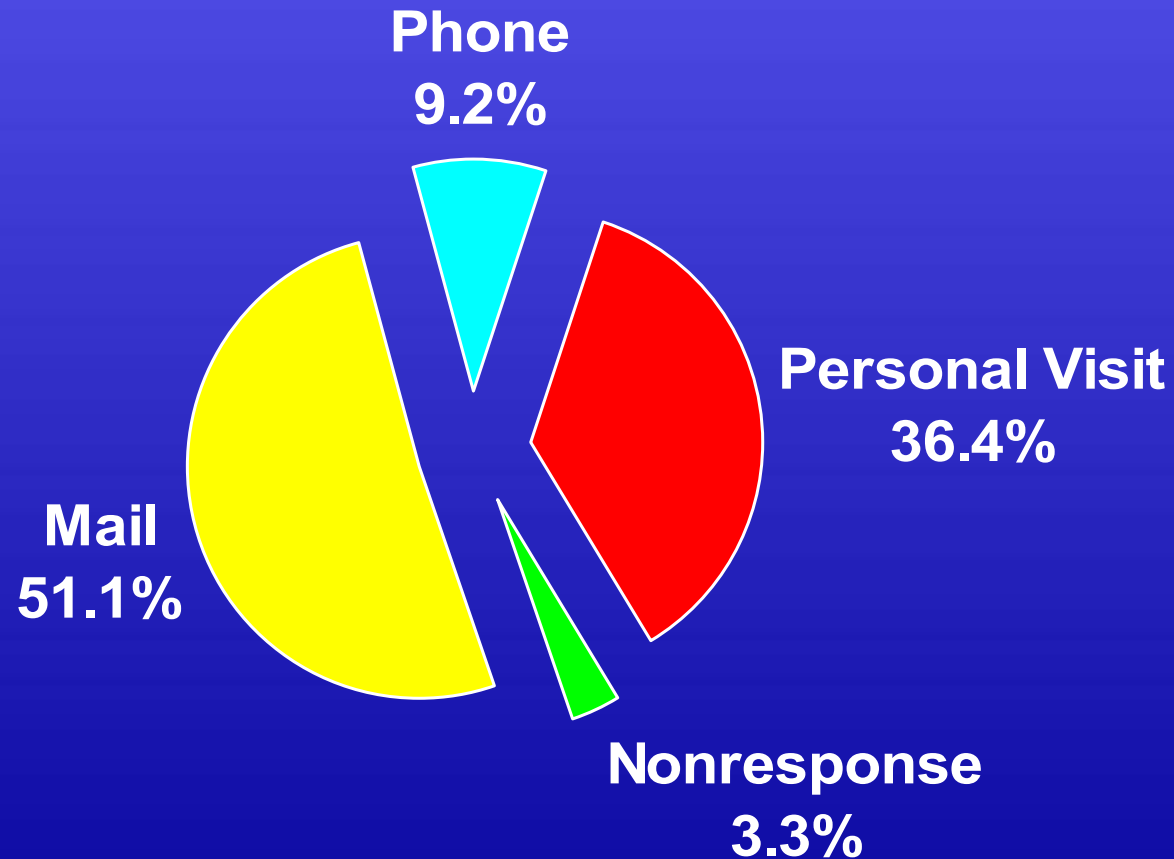## Personal Visit

- Interviewers are experienced, continuously employed report to regional offices

- Supervised by Senior Interviewers with 3+ years experience

- Interviewers use laptops with English and Spanish translations

- Regional offices recruit bilingual staff to ensure data collection from non-English speaking households

Data Collection
Response Rates

- These three sequential modes of data collection have been successful
  - response to the ACS is very high
  - true nationally and for all states

- ACS national-level response rates for 2000 - 2004 have all exceeded 95%

- Full implementation since January continues to achieve response rates above 95%

U S C E N S U S B U R E A U
Helping You Make Informed Decisions

Data Collection
2001 interview results by mode

Phone
9.2%

Personal Visit
36.4%

Mail
51.1%

Nonresponse
3.3%

Group Quarters Population

- Developing best methods for adding group quarters population to the survey
  - Sampling
  - data collection
  - estimation

Data Processing
Annual accumulation

- All data collected in a given calendar year are used to produce the ACS estimates for that year

- Sample used for estimation is not the sum of the 12 sample panels for a given year

Data Processing Annually on Collection Months

Sample Panel	Calendar Month				
	Jan 2005	Feb 2005	Mar 2005	Apr 2005	May 2005
Nov 2004	Personal Visit				
Dec 2004	Phone	Personal Visit			
Jan 2005	Mail	Phone	Personal Visit		
Feb 2005		Mail	Phone	Personal Visit	
Mar 2005			Mail	Phone	Personal Visit

Data Processing
Annual Processing

- Coding

- Editing

- Imputation

Data Processing
Coding

- Automated and clerical coding used for write-in entries such as
  - Race, Hispanic origin
  - Language
  - Place of work
  - Ancestry
  - Industry, occupation and class of worker

Data Processing
Editing

- First step involves distinguishing between interviews and noninterviews
  - only interviews continue into edit
  - noninterviews dealt with during weighting

- For interviews, identify inconsistent and missing answers requiring imputation

U S C E N S U S B U R E A U
Helping You Make Informed Decisions

Data Processing
Imputation

- Assignments
  - Rule based
  - Uses other reported information from the data record

- Allocations
  - Nearest neighbor or hot-deck methods
  - Uses data from other data records

Imputation Rates of Interest

Imputation Rates for Items of Branch Interest: 2003, 2002, 2001,and 2000 ACS, and Census 2000					
Note: Data are limited to the household population and exclude the population living in institutions, college dormitories, and other group quarters in the ACS tabulations. However, the Census 2000 data include these persons.					
	Imputation Rates*				
Description	2003 ACS	2002 ACS	2001 ACS	2000 (C2SS)	Census 2000
Number of vehicles available	1.0	1.1	1.3	1.6	6.2
Place of birth	6.2	4.4	4.6	6.4	10.1
Citizenship	0.4	0.4	0.4	0.5	0.8
Previous residence** --					
Mobility status	2.2	2.5	2.6	4.0	6.9
Previous residence geography (one or more parts)	5.9	6.0	7.3	14.9	11.0
Employment status recode	3.4	3.5	3.8	6.0	10.9
Place of work geography (one or more parts)	5.2	4.9	5.3	9.9	10.7
Means of transportation to work	3.1	3.0	3.1	4.6	7.6
Private vehicle occupancy (carpooling)	4.1	3.9	4.1	5.8	10.0
Time leaving home to go to work	9.6	9.2	9.9	11.3	15.0
Travel time to work (minutes)	7.0	6.9	7.2	8.7	11.8
* Base to the imputation rate is the population at risk for the characteristic. For example, the imputation rate for "Travel time to work" is based on "workers 16 years and over who did not work at home."					
** Previous residence is for a 1 year interval in ACS and for a 5 year interval in Census					
Source: Data are based on 2003, 2002, 2001, and C2SS data from the American Community Survey detailed tabulations and the Summary File 3 from the Census 2000 detailed tabulations.					

USCENSUSBUREAU

Helping You Make Informed Decisions

Weighting and Estimation

- Initial weights reflect the probability of selection

- Adjust weights of interviewed households to account for noninterviews

- Adjust weights to independent housing unit and population estimates (controls)

Weighting and Estimation
Ratio Adjustments to Controls

- Post-census estimates are produced by updating the previous census results using various administrative records data

- In a multi-stage process, housing unit and population adjustment ratios are applied to the weights

- Applied at the county (or group of counties) level by race/ethnicity and age/sex groups

Weighting and Estimation
Single-year Estimation

- Estimates include

  – population estimates

  – rates

  – medians

USCENSUSBUREAU
Helping You Make Informed Decisions

Weighting and Estimation
Single-year Estimates

- Percent of workers who carpool to work in year 1:

$$P_1 = \text{Percent Who Carpool} =$$

$$\frac{\text{Number Who Carpool}}{\text{Total Worker Population}} = \frac{N_1}{T_1}$$

Weighting and Estimation
Multi-year Estimation

- Most multi-year estimates are generated by computing an average based on combining each years' estimates

Weighting and Estimation
Multi-year Estimates

- Three-year estimate of percent who carpool in years 1-3:

$$P_{1-3} = \frac{N_1 + N_2 + N_3}{T_1 + T_2 + T_3}$$

Weighting and Estimation
Multi-year Estimation

- Medians are produced using combined data records from all years, not by averaging each year's median

Data Review, Acceptance and Release

- Automated review tool

- Data released within 8 months of completion of data collection
  - August of year following data collection