

American Community Survey Technical Training

Mark E Asiala
CaSDC General Update Meeting
June 27, 2013



U.S. Department of Commerce
Economic and Statistics Administration
U.S. CENSUS BUREAU



Overview

- Sample Design
- Sample Disposition
- Weighting and Controls
- Migration Flow Data



U.S. Department of Commerce
Economic and Statistics Administration
U.S. CENSUS BUREAU



SAMPLE DESIGN

Starting Points of Sample Design

- We have a fixed budget to design an annual survey to satisfy the needs of decennial long form
- Driven by reliability parameter called coefficient of variation which measures variability compared to size of estimate.
- Target is the 5-year estimates
- 1- and 3-year estimates are by-products

Goals of Sample Design

- Produce reliable estimates for:
 - Small areas including census tracts, governmental units
 - Small populations as a by-product
- Reduce the disparity in reliability across size of area
- Reduce response burden on the public



U.S. Department of Commerce
Economic and Statistics Administration
U.S. CENSUS BUREAU



Response Burden

- Early in the history of the ACS, there was concern about the burden on the public
- Thus, it is built into the sample design that an address is only eligible to be in sample once in a 5-year period



U.S. Department of Commerce
Economic and Statistics Administration
U.S. CENSUS BUREAU



Ensuring Reliability for Small Areas

- Important concept:
 - The reliability of an estimate is based on its sample size not its sampling rate
- Thus, large area can be sampled at a lower rate than smaller areas to get the same level of reliability
- This allows us to use different sampling rates to improve reliability for small areas

History of Varying sampling rates

- 1980 Census had 2 sampling rates
- 1990 Census had 3 sampling rates
- 2000 Census had 4 sampling rates
- 2005-2010 ACS had 7 sampling rates
- 2011+ ACS has 16 sampling rates

Why So Many Rates?

- More rate categories allows greater fine tuning of the sample
 - Less disparity in reliability across size classes
 - Less sudden jumps for areas near the boundaries of the size classes
- However, there is a trade-off
 - Improves reliability for target areas
 - May decrease reliability in higher geo areas
- Optimum is a question of balance

How are our rates defined?

- Items to note:
 - Total sample size is fixed not sampling rate
 - Sampling rates for smallest areas are fixed
 - Remainder are parameterized as a function of our base rate (BR)
 - Rates are dependent on measure of size (MOS) for the area and response rate
 - Sampling rate for a census block is based on smallest governmental unit or tract that contains it

Sampling Entities

- Governmental Units
 - Counties
 - Minor Civil Divisions
 - Places
 - School Districts
 - American Indian Areas including Chapters
- Tracts

Sampling Rates for Blocks based on their Governmental Units Measure of Size (Gov't Unit MOS)

STRATUM	BLOCK MOS CRITERIA	Sampling Rates
1	$0 < \text{Gov't Unit MOS} \leq 200$	15% (fixed)
2	$200 < \text{Gov't Unit MOS} \leq 400$	10% (fixed)
3	$400 < \text{Gov't Unit MOS} \leq 800$	7% (fixed)
4	$800 < \text{Gov't Unit MOS} \leq 1,200$	$2.8 \times \text{Base Rate}$

- Measure of size (MOS) is defined as estimated occupied housing units for area
- Fixed rates for strata 1-3, only stratum 4 has a parameterized sampling rate
- If Gov't Unit MOS > 1,200 use next slide

Sampling Rates for Blocks based on their Tract Measure of Size (Tract MOS)

Stratum	Block MOS Criteria	Sampling Rates
5	0 < Tract MOS ≤ 400	3.5×Base Rate
6	0 < Tract MOS ≤ 400 High Resp.	0.92×3.5×Base Rate
7	400 < Tract MOS ≤ 1,000	2.8×Base Rate
8	400 < Tract MOS ≤ 1,000 High Resp.	0.92×2.8×Base Rate
9	1,000 < Tract MOS ≤ 2,000	1.7×Base Rate
10	1,000 < Tract MOS ≤ 2,000 High Resp.	0.92×1.7×Base Rate
11	2,000 < Tract MOS ≤ 4,000	Base Rate
12	2,000 < Tract MOS ≤ 4,000 High Resp.	0.92×Base Rate
13	4,000 < Tract MOS ≤ 6,000	0.6×Base Rate
14	4,000 < Tract MOS ≤ 6,000 High Resp.	0.92×0.6×Base Rate
15	6,000 < Tract MOS	0.35×Base Rate
16	6,000 < Tract MOS High Resp.	0.92×0.35×Base Rate



U.S. Department of Commerce
Economic and Statistics Administration
U.S. CENSUS BUREAU



Simulated Reliability of Tract-Level Estimates by Size Class

Tract Size Category	Average Tract Size	CV* Using 7 Rates	CV* Using 16 Rates
0 – 400	291	59%	35%
401 – 1,000	766	36%	25%
1,001 – 2,000	1,485	26%	25%
2,000 – 4,000	2,636	23%	25%
4,000 – 6,000	4,684	17%	25%
6,000 +	8,337	13%	25%

CV = coefficient of variation (standard error / estimate)
for simulated 10% poverty estimate



U.S. Department of Commerce
Economic and Statistics Administration
U.S. CENSUS BUREAU



Final Comments on Sampling

- The addition of more summary rates allows us to achieve more equitable reliability across small areas
- The impact on the reliability of larger areas has been minimal
- The combination of increasing the number of sampling rates and the sample expansion in 2011 should help small area data substantially



U.S. Department of Commerce
Economic and Statistics Administration
U.S. CENSUS BUREAU



SAMPLE DISPOSITION



U.S. Department of Commerce
Economic and Statistics Administration
U.S. CENSUS BUREAU



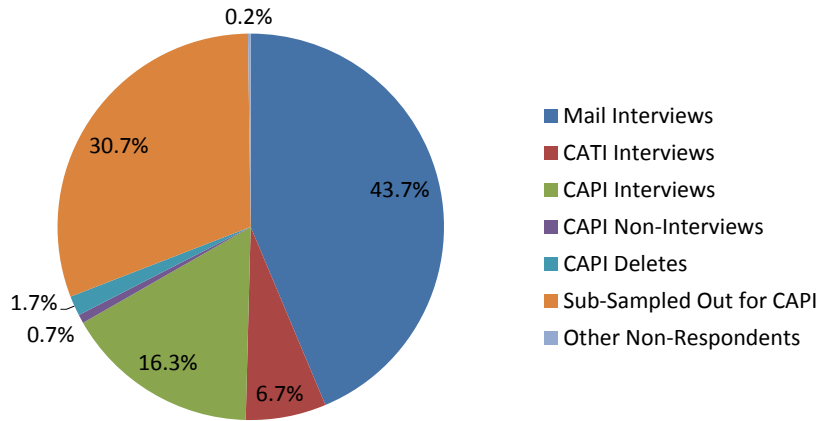
What is Sample Disposition?

- What happens to our sample once it is sent to the field?
- Modes of data collection
 - Internet (2013+)
 - Mail
 - Computer Assisted Telephone Interview
 - Computer Assisted Personal Interview

Classifications for Sample Disposition (2011 Data)

- Interviews
 - Mail Interview
 - CATI Interview
 - CAPI Interview
- Noninterviews
 - CAPI Non-Interview
 - Other Nonrespondents
- Other
 - CAPI Delete
 - Sampled Out for CAPI

Sample Disposition for U.S.

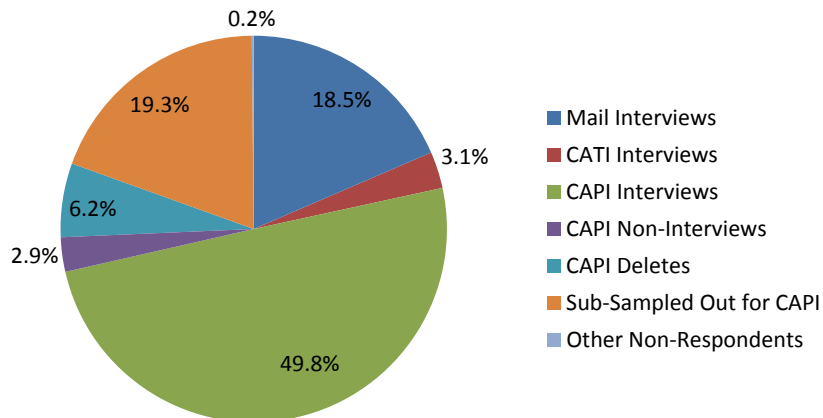


United States
Census
Bureau

U.S. Department of Commerce
Economic and Statistics Administration
U.S. CENSUS BUREAU



Sample Disposition for Alaska

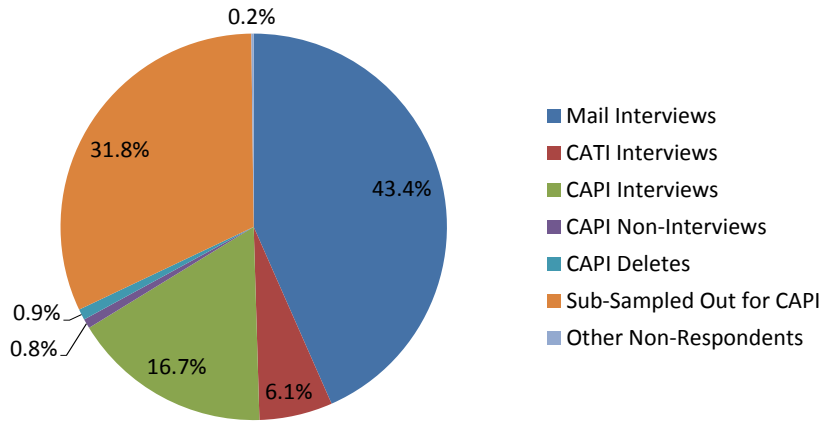


United States
Census
Bureau

U.S. Department of Commerce
Economic and Statistics Administration
U.S. CENSUS BUREAU



Sample Disposition for California

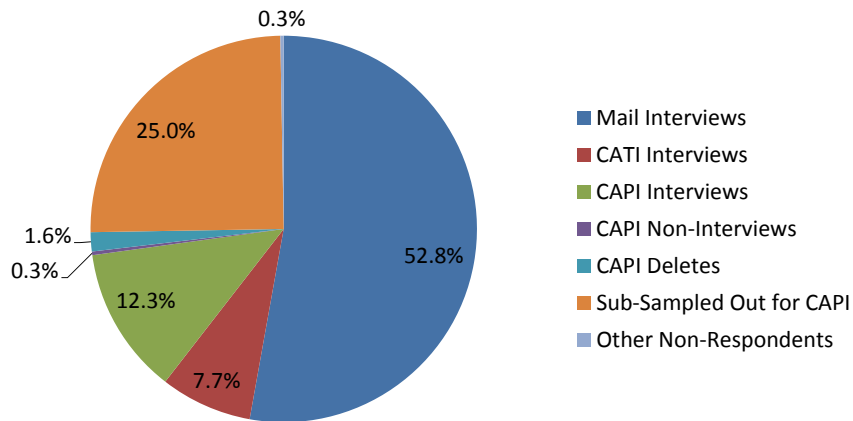


United States
Census
Bureau

U.S. Department of Commerce
Economic and Statistics Administration
U.S. CENSUS BUREAU



Sample Disposition for Minnesota



United States
Census
Bureau

U.S. Department of Commerce
Economic and Statistics Administration
U.S. CENSUS BUREAU



WEIGHTING AND CONTROLS

Purpose of Weighting

- We sample only a portion of universe of housing and population
- Without weights, estimates from the sample would represent only the interviewed sample
- Process of weighting allow estimates to reflect the original universe of housing units and population

Starting Points for Weighting

- Unique housing unit weight is assigned to every sample housing unit to estimate
 - Housing characteristics
 - Household characteristics
- Unique person weight is assigned to each person to estimate
 - Householder characteristics
 - Population characteristics



U.S. Department of Commerce
Economic and Statistics Administration
U.S. CENSUS BUREAU



Focus for Today

- General overview of housing unit and household person weighting
- Role of independent estimates of housing units and population by demographics as survey controls



U.S. Department of Commerce
Economic and Statistics Administration
U.S. CENSUS BUREAU



Basic HU Weighting Summary

- Assign base weights account for probability of selection
 - Adjust for nonresponse
 - Adjust for total housing unit coverage
- ...go off and do household person weighting...
- Adjust for differential coverage of housing units based on demographics of householder

Base Weights

- Two sample selections occur for the ACS
- Initial sample selection
 - Affects all sample records
 - Rates vary from 0.5% to 15%
- Sample selection of nonrespondents for personal visit (CAPI)
 - Affects only housing units that do not respond to prior to personal visit
 - Rates vary from 1-in-3 to full followup (100%)

Examples: Base Weights

- Weights are the inverse of the sampling rates
 - Sampling rate = 5%, weight is $1/0.05 = 20$
 - Sampling rate = 10%, weight is $1/0.10 = 10$
- If a unit does not respond via mail, internet, or telephone, it may be sampled further before being selected for personal visit (CAPI)
 - CAPI sampling rate is 1-in-3, then the weight above is multiplied by 3

Adjustment for Nonresponse

- Key assumptions of nonresponse adjustment
 - All vacant units are properly identified in the field
 - All noninterviews are assumed to be occupied
- Based on these assumptions, the nonresponse adjustment
 - Adjusts the weights of the occupied interviews to account for the noninterviews
 - Does not adjust the weights of vacant interviews
 - Weights of noninterviews are set to zero

Examples: Adjust for Nonresponse

- Consider the following:
 - Weighted occupied interviews = 81
 - Weighted noninterviews = 9
 - Weighted vacant interviews = 10
- After adjustment [factor occ. int. = $(81+9)/81$]
 - Weighted occupied interviews = 90
 - Weighted noninterviews = 0
 - Weighted vacant interviews = 10
- Note weighted vacant interviews is fixed

Adjustment for HU Coverage

- The adjustment for coverage uses independent estimates of housing units (HUs)
 - Independent estimate is for total HUs only
 - No breakdown of occupied / vacant is available
- Thus, this adjustment has the following impact
 - Estimate of occupied and vacant HUs are impacted proportionally
 - Vacancy rate for HUs is not impacted

Example: Coverage Adjustment

- Consider the following:
 - Weighted occupied interviews = 90
 - Weighted vacant interviews = 10
 - Total weighted HUs = 100, Independent Est = 110
- After adjustment (factor = $110 / 100 = 1.1$)
 - Weighted occupied interviews = 99
 - Weighted vacant interviews = 11
 - Total weighted HUs = 110 = independent estimate
- Note that the vacancy rate:
start $10/100 =$ finish $11 / 110 = 90\%$

Household Person Weighting

- Initial person weight is the housing unit weight after the HU coverage adjustment
- Already incorporated
 - Adjusting nonresponse
 - Adjusting for whole household / HU coverage
- What remains?
 - Within household coverage
 - Differential coverage by demographics

Goals of Person Weighting (Iterative Process)

- Adjust for subcounty total population coverage
- Ensure consistency between person and housing unit weights when estimating
 - Occupied housing units and householders
 - Married-spouse-present households and estimates spouse (also unmarried partners)
- Adjust for differential coverage by demographics

Subcounty Coverage

- Subcounty area (cannot cross county)
- Area must have the following population
 - 24,000 for 1-year ACS
 - 8,000 for 3-year ACS
 - 2,500 for 5-year ACS
- Collapse areas as necessary
- Must have 2x population above for controlling subcounty areas to be possible
- Entire county is the fall back method

Spouse / Householder Equalization

- Adjust person weights so that the following hold true:
 - Make estimated householders equal to estimated occupied housing units
 - Make estimated spouses + unmarried partners equal to households of same type
 - Preserve total population
- Without this step, estimates made with HU weights vs person weights would be inconsistent

Differential Person Coverage

- These adjustment are performed within a weighting area (1 or more counties)
- Adjust for coverage at weighting area by
 - Hispanic + Non-Hispanic by 5 Race (6 total)
 - Age / Sex (13 x 2 = 26 categories)
- Collapse categories as necessary
 - Maximum is 156 combinations of above
 - Realistically, collapse to approximately half

Iterative Process

- We can't achieve each of the 3 goals simultaneously
- We iterate the three steps up to 40 times
- Criteria for stopping
 - Achieve goals within 0.01% in fewer than 20 iterations (e.g., +/- 1 for subcounty pop = 10k)
 - Achieve goals within 0.1% in fewer than 40 iterations (e.g., +/- 10 for subcounty pop = 10k)
 - Reach 40 iterations



U.S. Department of Commerce
Economic and Statistics Administration
U.S. CENSUS BUREAU



Adjustment for Differential Householder Coverage

- Applied only to occupied HUs
- Total occupied housing units
 - Changes very little at county level
 - Can change more at the subcounty level
- Total vacant housing units
 - Unchanged at any geographic level
- Vacancy rate
 - Impacted by any change in estimated occupied
 - Impacted subcounty but very little at county level



U.S. Department of Commerce
Economic and Statistics Administration
U.S. CENSUS BUREAU



Example: Differential Coverage Adjustment

- Consider the following area total:
 - Estimated HUs with White Householder = 55
 - Estimated HUs with non-White householder = 44
 - Estimated vacant HUs = 11, Est. total HUs = 110
- After adjustment
 - Estimated HUs with White Householder = 53
 - Estimated HUs with non-White householder = 47
 - Estimated vacant HUs = 11, Est. total HUs = 111
- Note: estimated vacant HUs is unchanged
- Note: est. vacancy rate changes because total +1



U.S. Department of Commerce
Economic and Statistics Administration
U.S. CENSUS BUREAU



In summary

- Weighting assigns person and HU weights
- Weighting accounts for sampling, nonresponse, and coverage adjustments
- Use population estimates as survey controls
- Some subcounty areas (places / minor civil divisions) may be (nearly) controlled
- Published counties typically are controlled for
 - Total population
 - Possibly Hispanic / non-Hispanic, some age / sex categories
- Do not expect county race totals to match population estimates



U.S. Department of Commerce
Economic and Statistics Administration
U.S. CENSUS BUREAU



MIGRATION FLOW DATA



U.S. Department of Commerce
Economic and Statistics Administration
U.S. CENSUS BUREAU



County-to-County Migration Flows

- Derived from residence one year ago question on ACS
- In-flow estimate would list estimated population coming from all 3,142 other counties to the county of interest
- Out-flow would be the reverse of that
- Poses significant challenges to standard American Fact Finder system



U.S. Department of Commerce
Economic and Statistics Administration
U.S. CENSUS BUREAU



Release of Migration Flow Data

- Data are released as downloadable tables

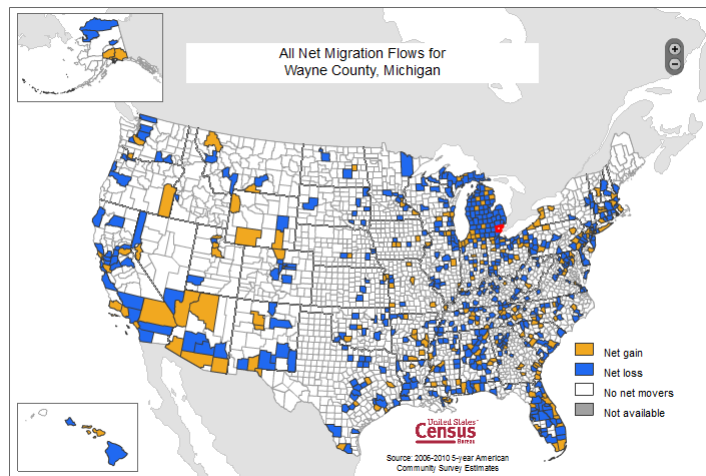
<http://www.census.gov/hhes/migration/data/acs/county-to-county.html>

- Data are also available on an interactive mapping tool

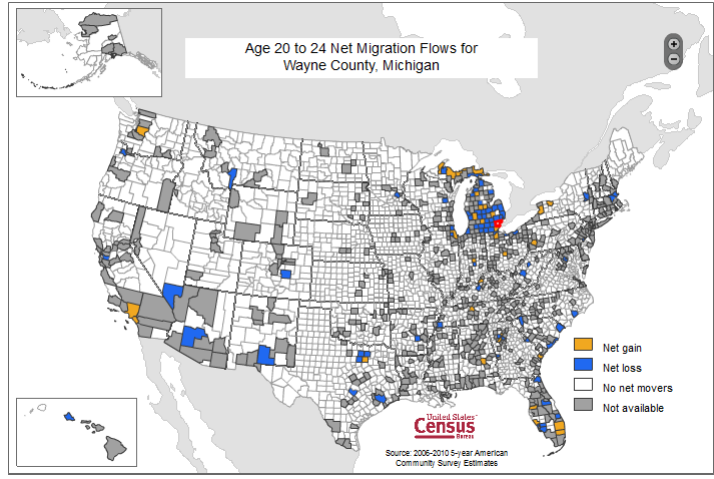
<http://flowsmapper.geo.census.gov/flowsmapper/flowsmapper.html>



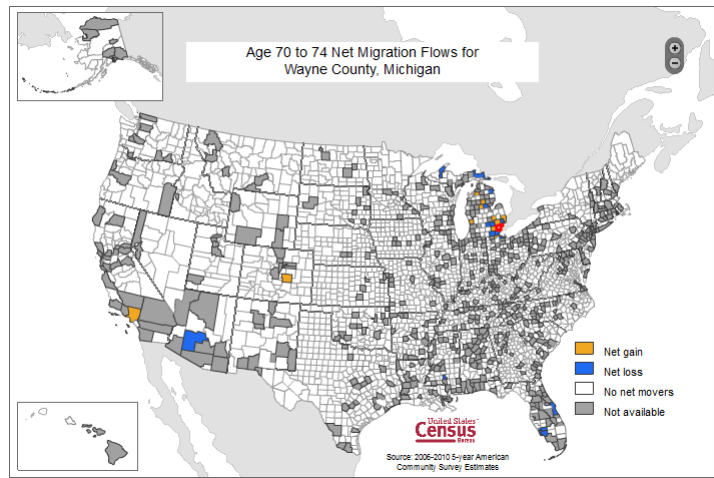
Wayne County, MI – Net Flow



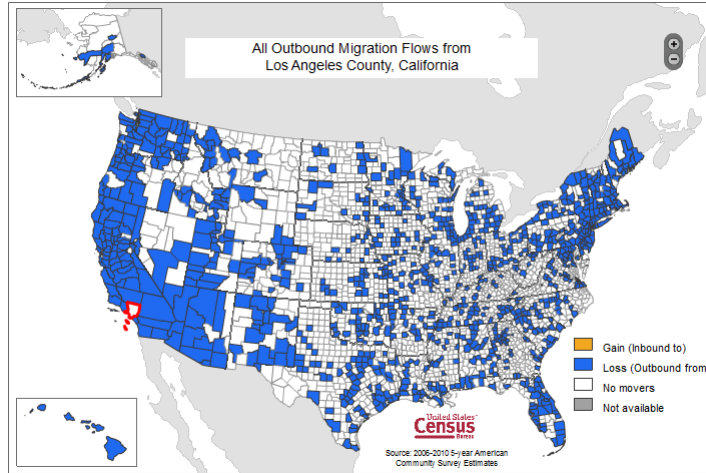
Wayne County, MI – Net Age 20-24



Wayne County, MI – Net Age 70-74



Los Angeles - Outflow

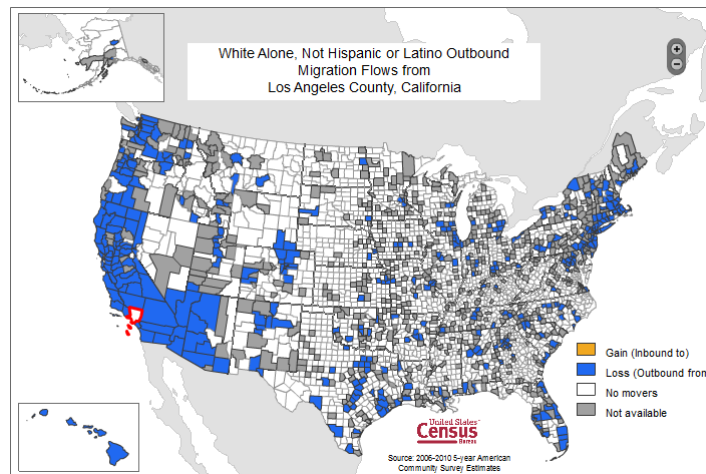


United States
Census
Bureau

U.S. Department of Commerce
Economic and Statistics Administration
U.S. CENSUS BUREAU



Los Angeles – Outflow White Alone, Not Hispanic

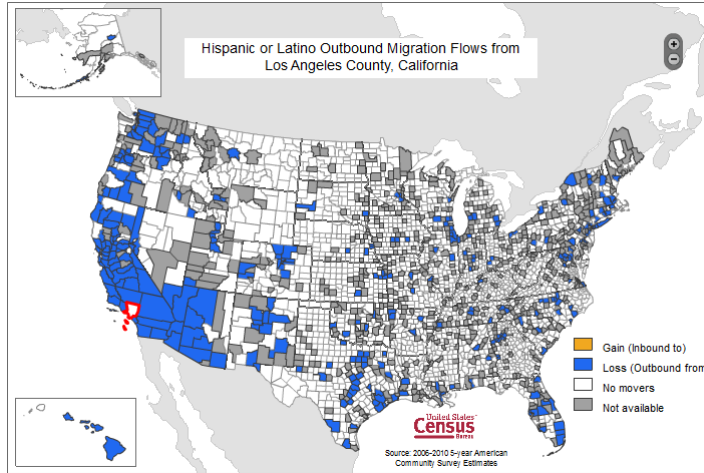


United States
Census
Bureau

U.S. Department of Commerce
Economic and Statistics Administration
U.S. CENSUS BUREAU



Los Angeles – Outflow Hispanic



U.S. Department of Commerce
Economic and Statistics Administration
U.S. CENSUS BUREAU



What to Expect

Year	2012	Late 2012	2013	2014	2015	2016	2017
ACS Data set	05-09	06-10	07-11	08-12	09-13	10-14	11-15
File:							
County-to-County	Released	Released	X	X	X	X	X
County/MCD-to-County/MCD	Released	Released	X	X	X	X	X
...by Age		Released					X
...by Sex		Released					X
...by Race		Released					X
...by Hispanic Origin		Released					X
...by Marital Status			?				
...by Place of Birth			?				
...by Nativity by Education				?			
...by Labor Force Status				?			
...by Industry Group					?		
...by Occupation Group					?		
...by Labor Force Status					?		
...by Poverty Status						?	
...by Tenure						?	
...by Individual Income							?



U.S. Department of Commerce
Economic and Statistics Administration
U.S. CENSUS BUREAU



Contact Information

- Mark.E.Asiala@census.gov
- Useful links:
 - ACS webpage:
<http://www.census.gov/acs/>
 - Design and Methodology Report:
http://www.census.gov/acs/www/methodology/methodology_main/
(working on update for 2013 and internet)



U.S. Department of Commerce
Economic and Statistics Administration
U.S. CENSUS BUREAU

