Planning for Production: Design and Sampling for the 2016 National Survey of Children's Health

Jason Fields, Scott Albrecht, Keith Finlay, Antoinette Lubich
U.S. Census Bureau[1]

2016 AAPOR
May 12-15, 2016
Austin, TX

[1] This work is released to inform interested parties of ongoing research and to encourage discussion of work in progress. Any views or opinions expressed in the paper are the authors' own and do not necessarily reflect the views or opinions of the U.S. Census Bureau or the Maternal and Child Health Bureau.

Outline

- Short Background
- Sample Design
- Administrative and Auxiliary Data Flags
  - Child Present Flags for Stratum Designation
    - Flag Evaluation Against 2014 ACS
  - Internet Likelihood
- Experimental Evaluations
- 2016 NSCH Implementation Schedule

Acknowledgements

Sponsor:

U.S. Department of Health and Human Services
Health Resources and Services Administration
Maternal and Child Health Bureau
Office of Epidemiology and Research

Michael D. Kogan, PhD
Reem M. Ghandour, DrPH, MPA
Catherine J. Vladutiu, PhD, MPH

Survey team at Census:

- Leah Meyer
- Antoinette Lubich
- and many others …
- Elizabeth Sinclair
- Tracy Mattingly
- Scott Albrecht
- Keith Finlay

Background

NSCH: To produce national and state-based estimates on the health and well-being of children, their families, and their communities.

Fielded: 2003, 2007, 2011-12

NS-CSHCN: To assess the prevalence and impact of special health care needs among children in the U.S., and to evaluate change over time.

Fielded: 2001, 2005-06, 2009-10

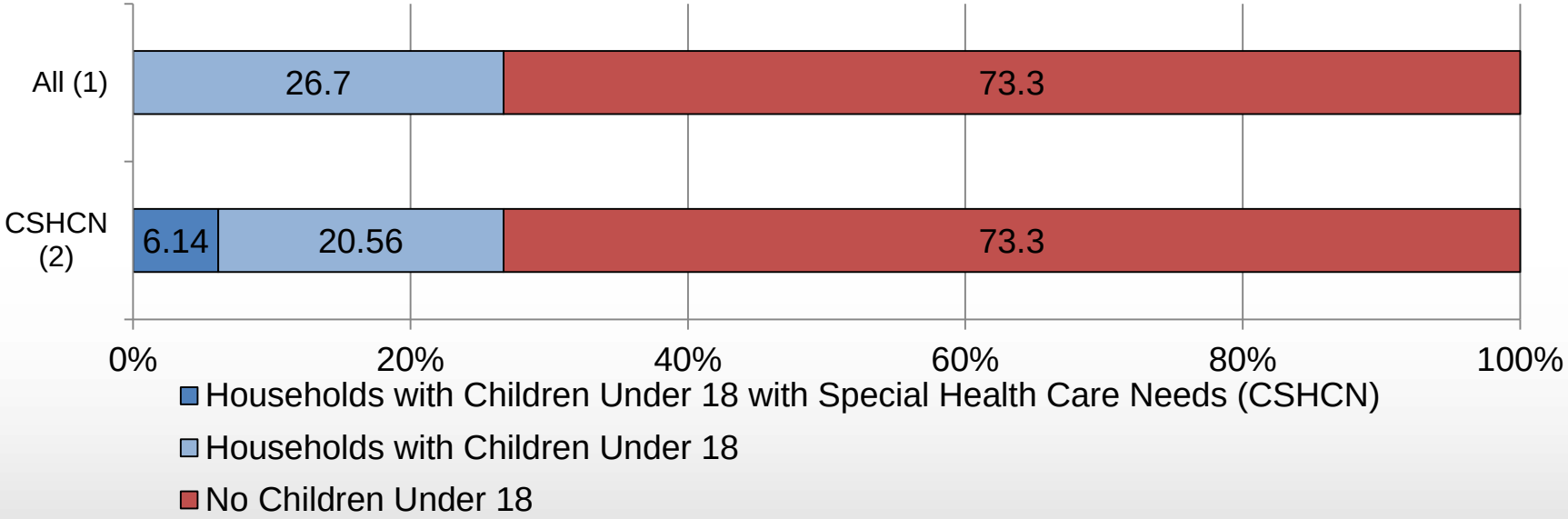
Common Elements:

- Historically directed and funded by HRSA MCHB and fielded by the CDC/NCHS as a module of SLAITS as a RDD telephone survey (landline + cell-phone samples);
- Produces both national and state-level estimates;
- All data are parent/care-giver reported.

Sample Design

- Nationally and state representative address-based annual survey
- Interviews completed through self-administered web or paper modes
- Sampling objective: To ensure an adequate sample which will provide reliable and statistically sound estimates for states and children with special health care needs
- Two-stage interview with screener procedure to identify households with children and subsampling process to select a single reference child for topical questions

Presence of Children in Households



[1] Anticipated household characteristics based on ACS audit

[2] Distribution based on 2009/10 National Survey of Children with Special Health Care Needs (23% of Households with children have 1 or more CSHCN)

Administrative Records Supplementation of the MAF

Motivation:

- Improved sampling efficiency and reduction of survey costs

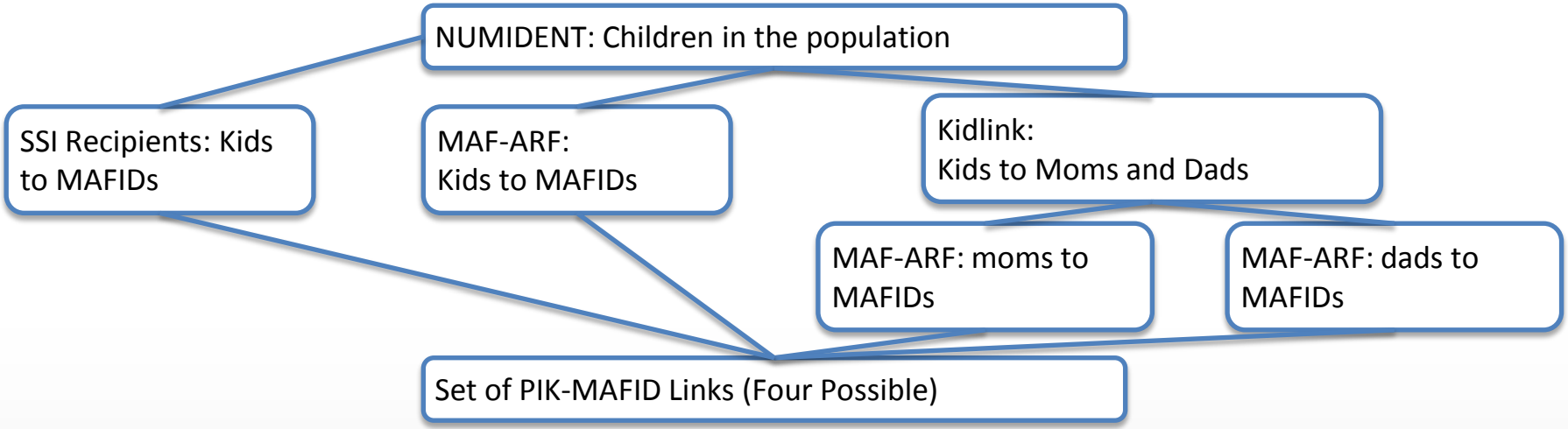
Method:

- Utilize the Title 13 Census Master Address File (MAF) as a sampling foundation and leverage administrative records to indicate the presence of children at a MAF-ID

Primary Information Sources:

- Numident: a list of Social Security Number applicants
- CARRA Kidlink file: a prototype linkage between children and parents based on Census and administrative records
- Master Address File Auxiliary Reference File (MAF-ARF): a file that links person identifiers with the latest location updates from a variety of administrative data

Sample frame construction



- Input data for Kidlink and MAF-ARF
  - * MAF-ARF: Master Address File – Auxiliary Reference File
  - * SSA: Numident
  - * Census: 2010 Census Unedited File
  - * IRS: 1040 and 1099 files
  - * CMS: Medicare Enrollment Database
  - * Indian Health Service database
  - * Selective Service System
  - * HUD: Public and Indian Housing (PIC) and Tenant Rental Assistance Certification System (TRACS) data

Administrative Records Supplementation of the MAF

File	Total Records	Matched to a MAFID
Children in the Numident	75,156,219	59,841,686
Kidlink-matched Mothers	65,529,375	60,031,595
Kidlink-matched Fathers	53,936,285	49,767,120

- Social Security Agency's Supplemental Security Income program (SSI) direct matching of records for children under 18 years
- Resulting in a sampling frame file of
  - 196,507,103 valid MAFIDs
  - 36,609,700 include child flags

Strata Development

- The sample size was allocated based on:
  - The budget
  - Relative sizes of Stratum 1 (households flagged as having children under 18 present) and Stratum 2 (households expected to have no children under 18 present)
  - Prevalence of households with children in each stratum
  - Expected eligibility and response rates
- State-level samples are allocated to produce equally-sized final sets of completed interviews in each state
- Sampling is designed for an initial sample size of 364,153 households nationwide to yield:
  - at least 1500 households with children per state and
  - include approximately 300 children with special health care needs in that state
- Approximately 61 percent of the sample is will be drawn from Stratum 1
- Average oversampling ratio is about 5.2:1 for Stratum 1 versus Stratum 2
- We would need to sample 703,653 addresses (1.93x the current sample) to confidently get the minimum number of completed interviews in the absence of the flags

Estimated NSCH 2016 Sample

Estimated Households with Children by Strata	Stratum 1 Households flagged as having children present	Stratum 2 Households not flagged as having children present	Total
Total	222,751	141,402	364,153
Estimated Number of Households with Children	165,963	11,730	177,693
Estimated Number of Screener Only (no children) Households	56,788	129,672	186,460

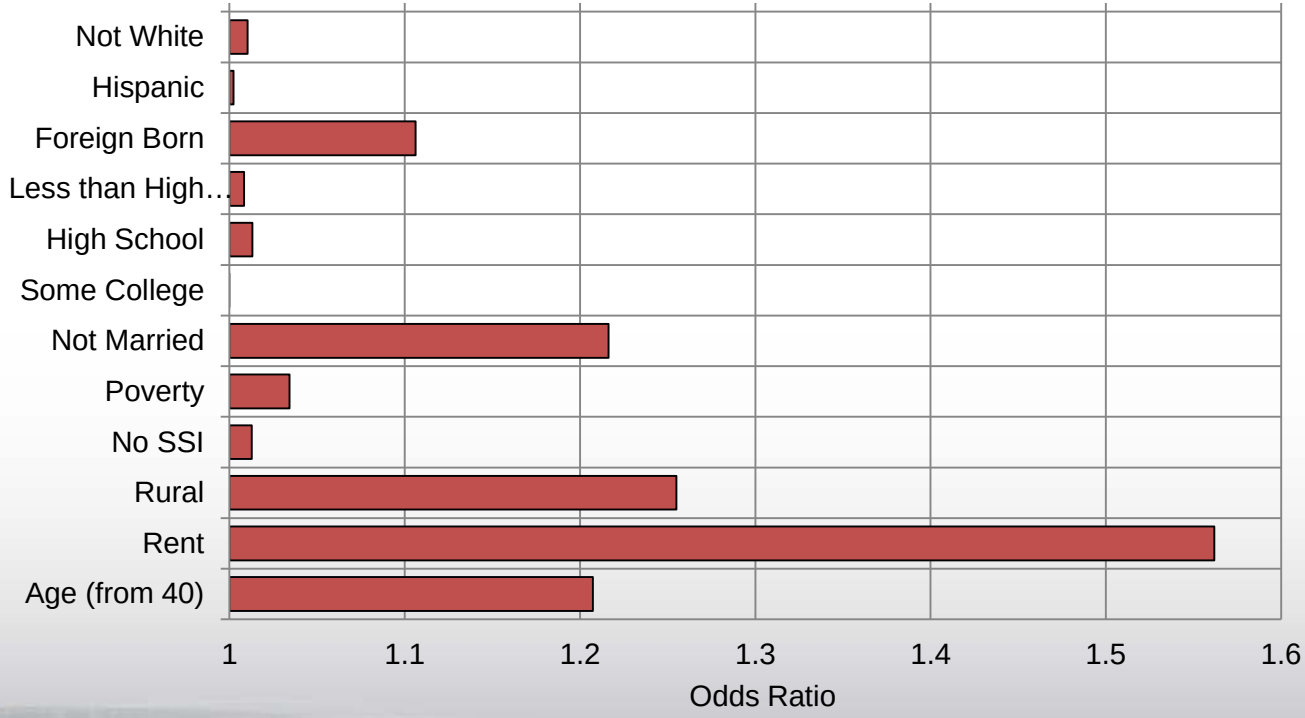
Characteristics of Households with Children by Sampling Stratum

Household Characteristics	Stratum I	Stratum II
Reference Person		
Hispanic	19.5%	23.9%
Not Hispanic		
White Alone	59.6%	50.7%
Black Alone	12.8%	16.1%
Asian Alone	5.6%	5.5%
Other/Two or More Alone	2.5%	3.7%
Foreign Born	21.3%	25.3%
Education		
Less than High School	11.4%	16.3%
High School	21.2%	25.9%
Some College	32.1%	32.7%
Bachelor's or More	35.3%	25.1%
Household		
Poverty	16.1%	27.3%
Rural	17.0%	20.3%
Own Residence	65.1%	39.2%
Married (Any person)	72.1%	54.8%

American Community Survey 2014, children ages 0 to 15 in 2014: n = 549,642

Probability a Household w/ Children is not Flagged

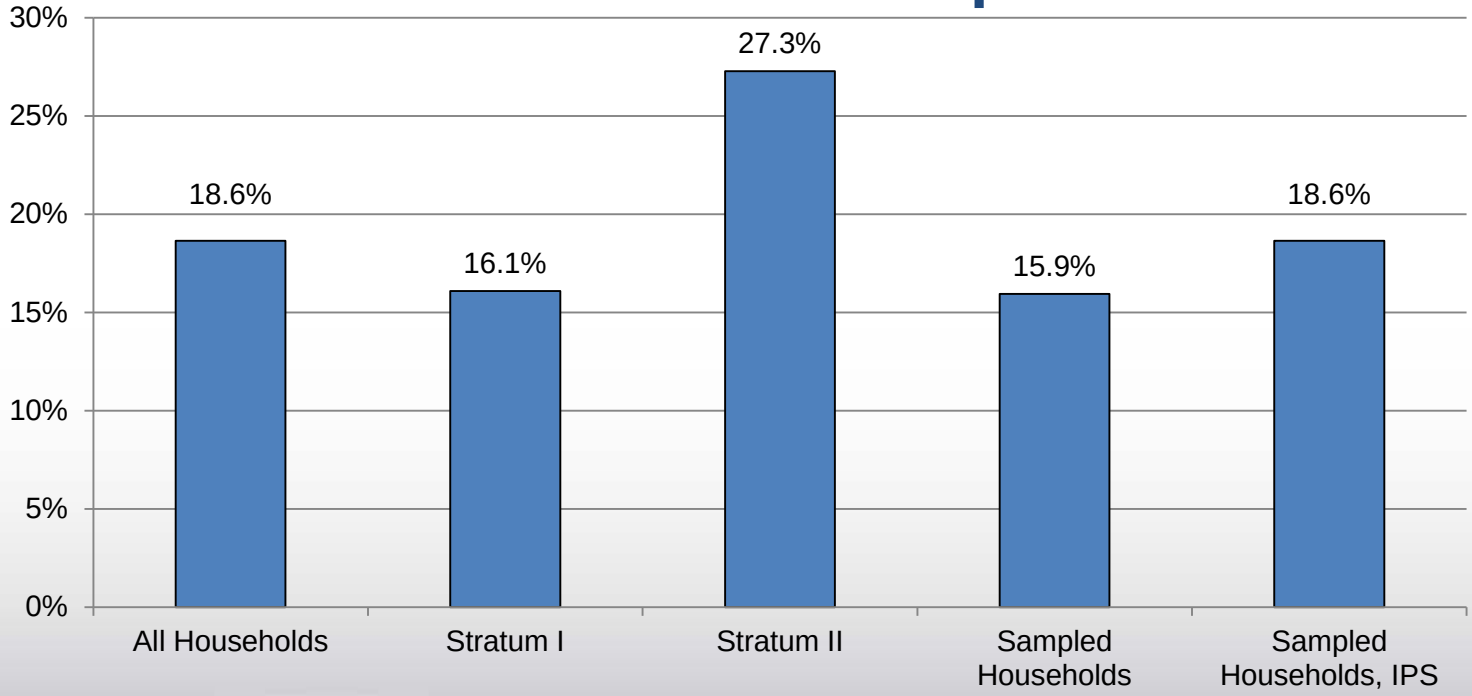
Logistic Regression, Standardized Values



American Community Survey 2014, children ages 0 to 15 in 2014

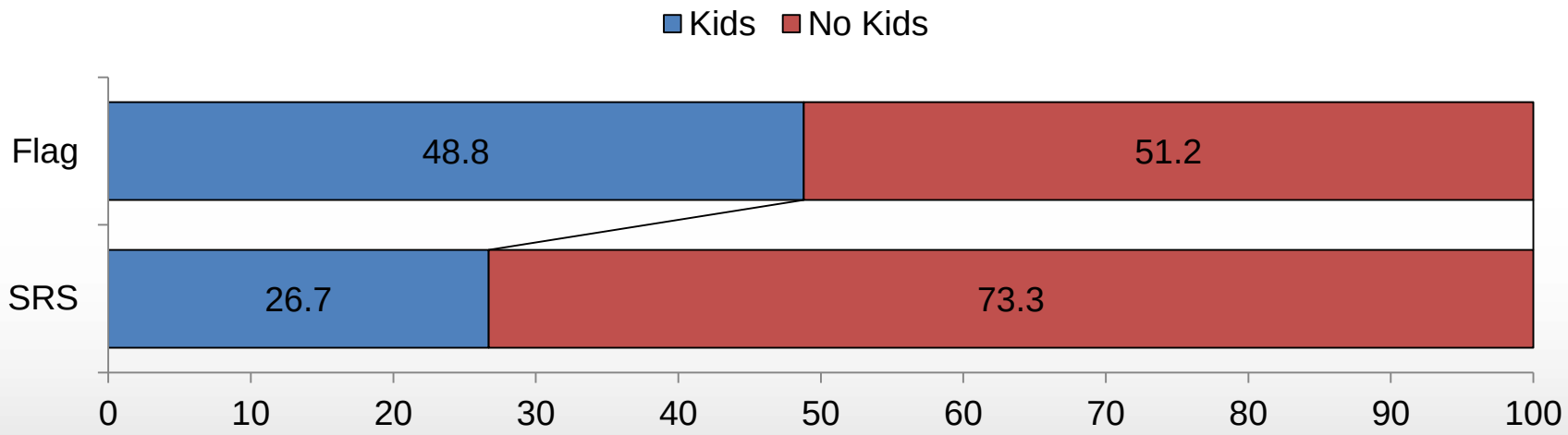
N = 549,642
Psuedo R2 = .0642

Poverty Status for Households with Children by Stratum and Sample



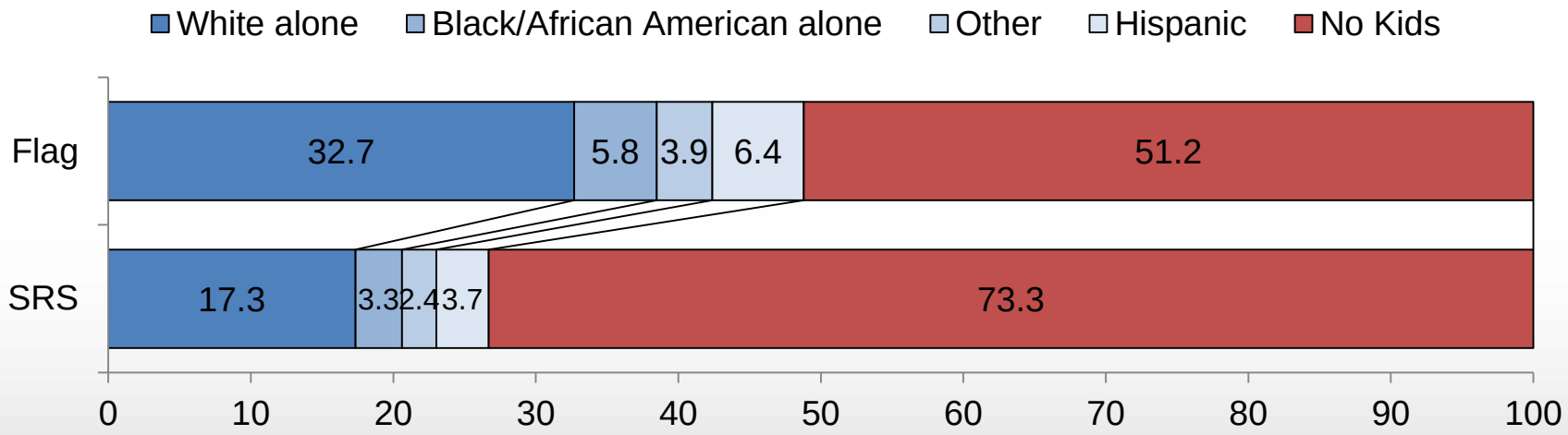
Children ages 0 to 15 in ACS 2014: n = 549,642

Characteristics of NSCH Sample, Kid Flag versus Simple Random Sample



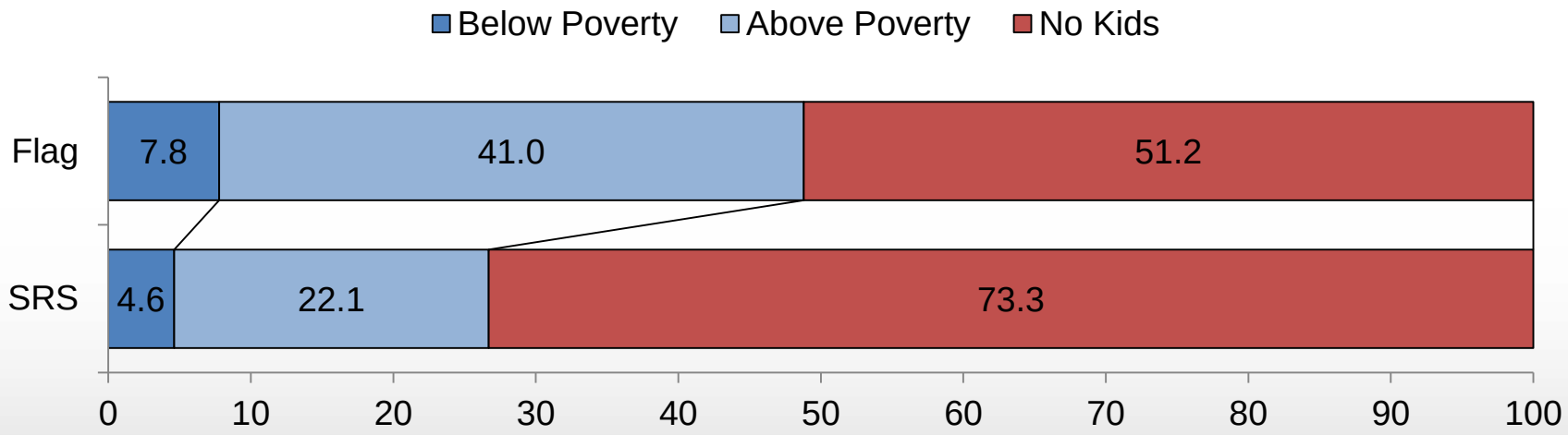
Anticipated household characteristics based on ACS audit and the kid flag sampling methodology

Characteristics of NSCH Sample, Kid Flag versus Simple Random Sample



Anticipated household characteristics based on ACS audit and the kid flag sampling methodology

Characteristics of NSCH Sample, Kid Flag versus Simple Random Sample



Anticipated household characteristics based on ACS audit and the kid flag sampling methodology

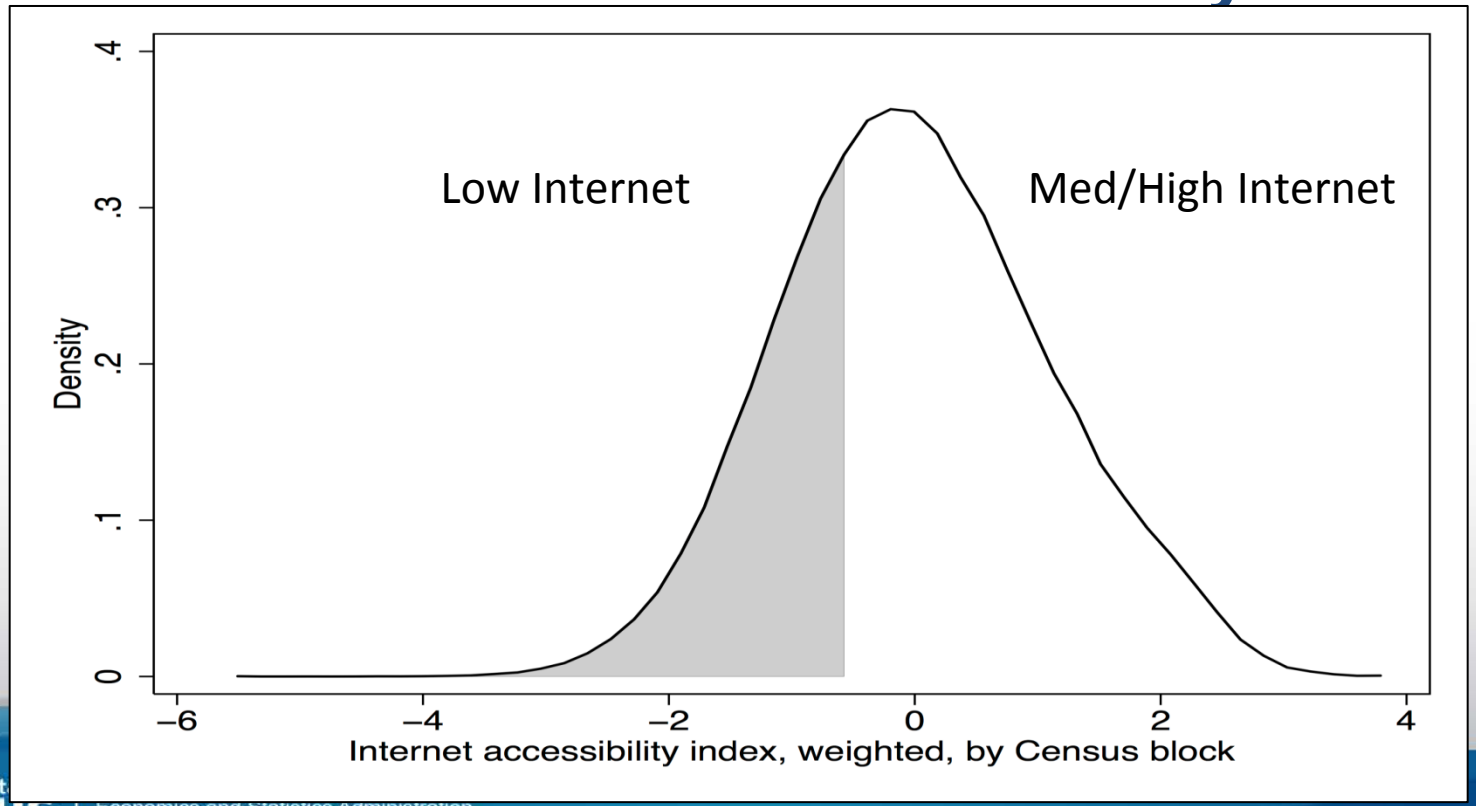
Small-Area Internet Access Index

- Two data sources
  - ACS paradata on whether respondents used Internet submission (tract-level, 2013-2014 survey years)
  - IRS 1040 data on whether households file electronically but without a paid preparer (block-level, 2014 tax year)

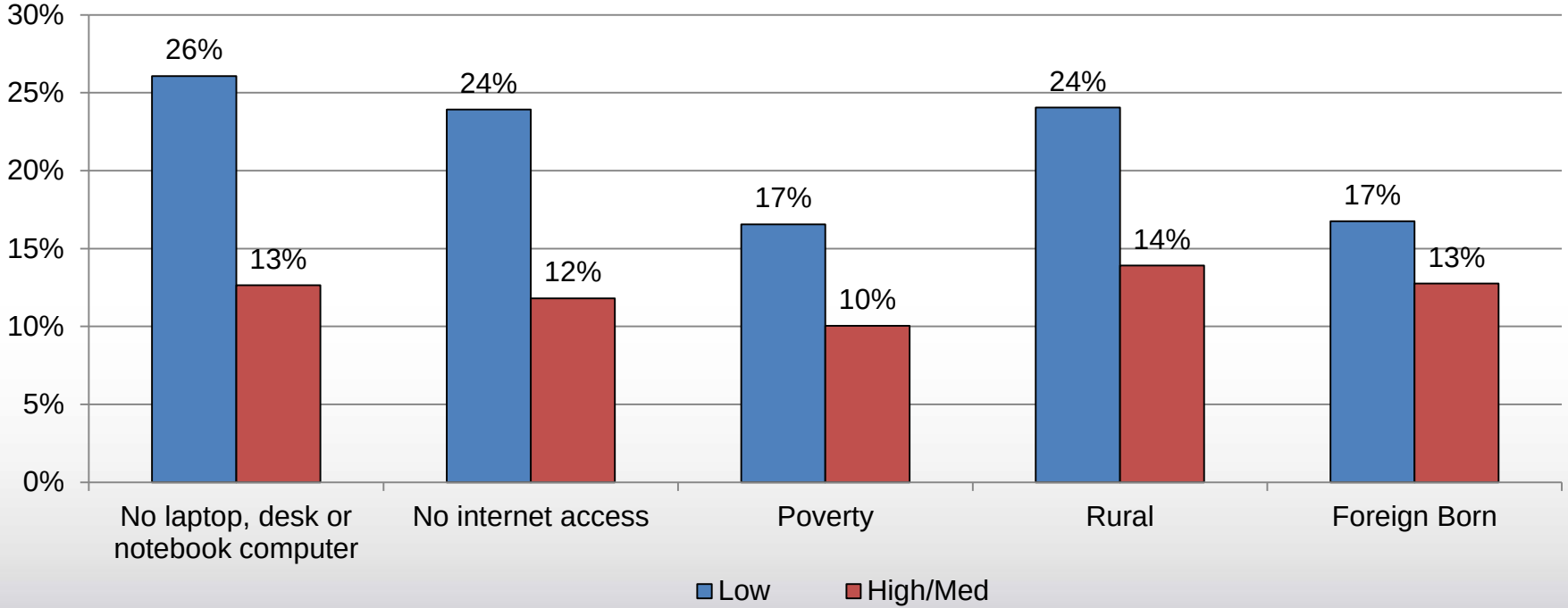
- Require scalar index
  - Use principal components analysis
  - Find the (standardized) scalar variable that maximizes the variation of linear combinations of the two data sources

- Low-Internet-accessibility flag
  - Census blocks with access index below 30th percentile of the access index distribution

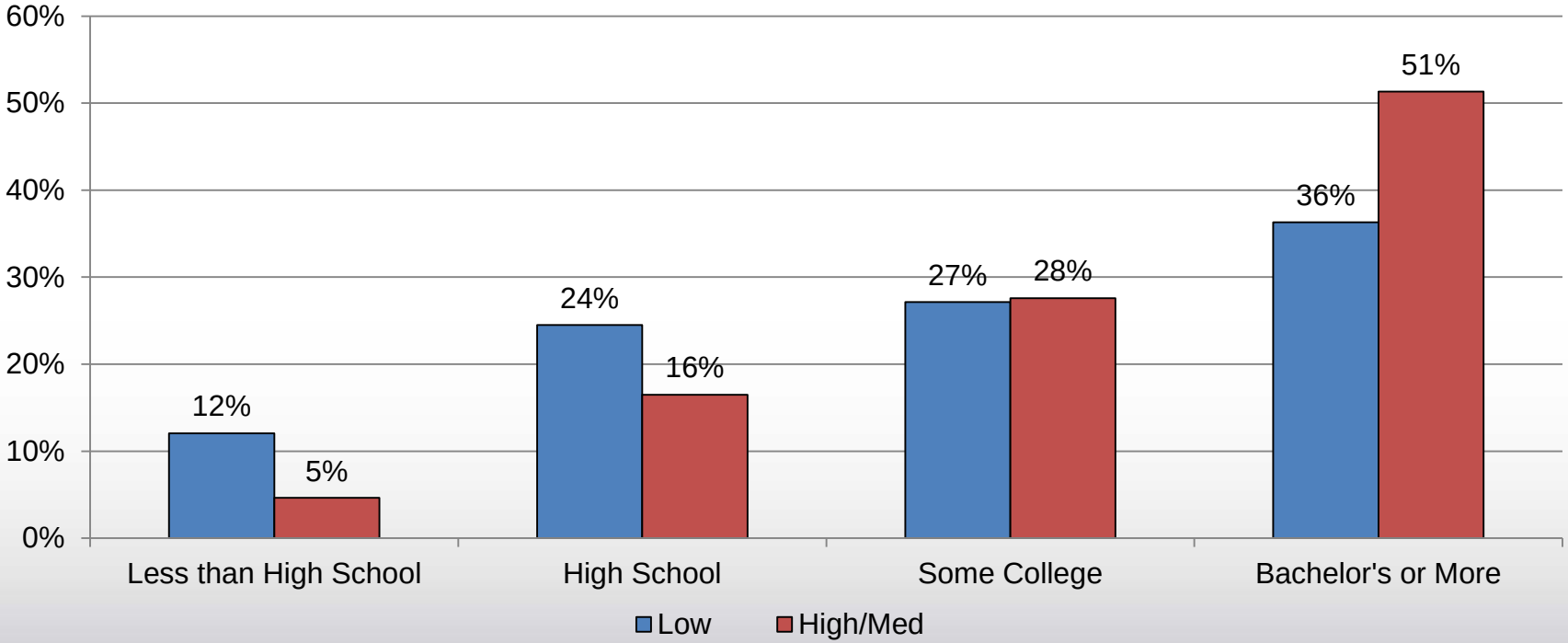
Kernel-Smoothed Probability Distribution Function of Internet Accessibility Index



High/Med and Low Internet Access Groups



High/Med and Low Internet Access Groups



Experiments

- The 2016 NSCH includes three experiments to evaluate potential areas to create efficiencies in the data collection process
- The treatment groups will be assigned within the sampling processes

1. Use of an unconditional cash incentive to reduce non-response bias and minimize follow-up costs.
    - Control group receiving no incentive
    - One-third of the sample receiving a $2 cash incentive in their initial web invitation mailing
    - One-third of the sample receiving a $5 cash incentive in their initial web invitation mailing
2. Test to evaluate whether an alternative HRSA MCHB branding improves response for the NSCH over the Census Bureau's standard branding
    - The second follow-up web-invitation mailing is split for the two different reminders
3. Evaluation of the efficacy of the internet likelihood procedure and response
    - Potential to apply more aggressive differences in methodology protocols based on this type of administrative records inputs

Experimental Samples

Incentive	Initial Cases	Mailing	Maximum Cases for Mailing Comparison	Internet Likelihood	Maximum Cases for Internet Comparison	Treatment Groups (TG)
$0 Control	121,385	Census	60,693	Low	18,208	1
				Med/High	42,485	2
		HRSA MCHB	60,692	Low	18,208	3
				Med/High	42,484	4
$2	121,384	Census	60,692	Low	18,208	5
				Med/High	42,484	6
		HRSA MCHB	60,692	Low	18,208	7
				Med/High	42,484	8
$5	121,384	Census	60,692	Low	18,208	9
				Med/High	42,484	10
		HRSA MCHB	60,692	Low	18,208	11
				Med/High	42,484	12

NSCH 2016 Data Collection

NSCH Web Invite and Screener Mail Schedule

	Group 1	Group 2
Initial Mailing Incentive = $0	10-Jun	24-Jun
Initial Mailing Incentive = $2		
Initial Mailing Incentive = $5		
First Follow-up Census	8-Jul	22-Jul
First Follow-up HRSA MCHB		
Second Follow-up High Web Letter Only	5-Aug	19-Aug
Second Follow-up Low Web Letter & 1st Screener		
Third Follow-up High Web Letter & 1st Screener	2-Sep	16-Sep
Third Follow-up Low Web Letter & 2nd Screener		
Fourth Follow-up High Web Letter & 2nd Screener	30-Sep	14-Oct
Fourth Follow-up Low Web Letter & 3rd Screener		

Estimated survey closeout
January 20, 2017

- Web Invite and Screener mailings are split into two equal groups to stagger mailing load
- All new Paper Screener returns are subsampled and batched to be included with Topical non-response in the Topical mailout schedule
- Additional Screener and Topical mailings are available as a contingency for non-response pending resources and schedule allowances

NSCH Paper Topical Mail Schedule

Mailing	Rec'd thru	Mail	Initial Mailing	1st Follow-up	2nd Follow-up	3rd Follow-up
Mailing 1	14-Sep	7-Oct	Group A			
Mailing 2	12-Oct	4-Nov	Group B	Group A		
Mailing 3	9-Nov	2-Dec	Group C	Group B	Group A	
Mailing 4	7-Dec	30-Dec	Group D	Group C	Group B	Group A

NSCH Production Timeline

- Data Collection NSCH 2016 - Early June 2016 – January 2017
- Data Processing - Begins in October 2016
- Initial File to MCHB - March 2017
- Final NSCH 2016 File for Public Use - May 2017
- OMB Submission for NSCH 2017 - January 2017
- Materials updates for NSCH 2017 - October 2016 – April 2017
- Collection of NSCH 2017 - Early June 2017 – January 2018

THANK YOU!

Jason Fields
Survey Director
National Survey of Children's Health
Survey of Income and Program Participation
Jason.M.Fields@census.gov

www.census.gov