



Oregon Health Insurance Survey

Methodology report prepared by

Robyn Rapoport, Erin Czyzewicz,
Kathy Langdale, and Margie Strickland
53 West Baltimore Pike
Suite 300
Media, PA 19063



Table of Contents

Table of Contents	2
INTRODUCTION	3
OVERVIEW OF THE OHIS	3
SAMPLE DESIGN	4
Overview	4
Sample Stratification	6
Disproportionate Incentives – Targeting Hispanics, African Americans and Lower Income HHs.....	11
Envelope Experiment	12
INSTRUMENT DEVELOPMENT	12
Questionnaire Development/Changes to the 2015 Instrument	12
Pretesting	12
CATI Pretest	13
Web Pretest	13
Mail Survey Pretest	13
DATA COLLECTION	13
Overview	13
Timeline	14
Training Materials and Interviewer Training	15
Call Rules for the CATI Interviews	15
Refusal Avoidance and Conversion Strategies	16
Caller ID	16
DATA PROCESSING AND PREPARATION	16
Overview	16
Data Cleaning, Editing and Coding Procedures	17
COMPLETED INTERVIEWS	17
Overview	17
Completed Interviews by Telephone Status	18
RESPONSE RATE	19
Defining the Response Rate	19
Final Response Rates	21
SURVEY WEIGHTS AND VARIANCE ESTIMATION	23
Survey Weights	23
Constructing the Base Weights	23
Constructing the Target Weights	25
Constructing the Person Weights	29
Variance Estimation and the Average Design Effect	29
IMPUTATION	30
Income Imputation	30
APPENDICES	30
APPENDIX A: Pretest Memos	30
APPENDIX B: Letters/Mailings	30
APPENDIX C: Web/CATI Survey Instrument	30
APPENDIX D: Paper Survey Instrument	30
APPENDIX E: Data Cleaning Guidelines and Editing/Coding Procedures	30
APPENDIX F: Final Update	30

INTRODUCTION

The Oregon Health Authority (OHA) contracted with SSRS to conduct the 2015 Oregon Health Insurance Survey (OHIS). The goal of the OHIS is to document health insurance coverage and access to and use of health care for the non-institutionalized population in Oregon. This report provides information about the sampling procedures and the methods used to collect, process, and weight data for OHIS 2015 as well as changes made to the methodology and survey instruments since the 2013 survey was fielded.

OVERVIEW OF THE OHIS

The goal of the OHIS is to collect information about health insurance coverage, health care costs, and access to and use of health care for the non-institutionalized population in Oregon. The study is designed to (1) attain adequate representation for each of the 15 regions within Oregon; and, (2) ensure the proportionality of ethnic and racial minority populations as much as possible. The OHIS sample is representative of Oregon's non-institutionalized population living in households.

The study utilizes an address-based (AB) sample design. Because the sample is address-based, data collection methods differ from traditional telephone samples. The OHIS 2015 study executed a data collection strategy designed to attain the highest response rate possible. This design combines telephone, web, and paper survey options. Telephone and web interviews were conducted in English and Spanish.

The OHIS questionnaire is comprised of questions asked about the household, a randomly selected "target" member of the household, and of all other members of the household. Basic information (including demographic and socioeconomic characteristics and insurance status) are collected on all of the members of each selected household¹, with more detailed information collected for the target person. Table 1 summarizes the topic areas covered in the OHIS for all members of the household and for the household member selected as the target individual for the survey. Further details on data collection are provided in later in this report.

¹ For online and telephone modes, up to ten household members' data are collected. The paper survey only collects information for up to six household members.

Table 1: Summary of Topic Areas Covered in the 2015 OHIS, by Household Members

Topics	Survey Respondent	All Household Members	Target Household Member	Target's Spouse (if present) and Parents (if present and Target age<26)
Demographic characteristics (age, race/ethnicity, gender, marital status, education)	X	X	X	X
Nativity and languages spoken			X	
Length of residency in the Oregon			X	
Health insurance coverage	X	X	X	X
Health Literacy			X	
Detailed employment questions			X	X
Availability of employer sponsored insurance			X	X
Health and disability status			X	
Access to and use of health care			X	
Family income			X	
Homeownership	X			
Household telephone status	X			

SAMPLE DESIGN

Overview

Historically, RDD telephone interviewing has been the method of choice for many survey data collection efforts given the strength of its randomization method (random digit dialing), the ease of administering complex questionnaires using computerized interviewing systems, excellent coverage of the overall population (given that only about 3% of Americans live in a household without telephone service), and relatively low cost. Survey coverage refers to the extent to which the sample frame for a survey includes all members of the target population. A survey design with a gap in coverage raises the possibility of bias if the individuals missing from the sample frame (e.g., households without landline telephones) differ from those in the sample frame.

Cell phone-only households are increasing rapidly in the United States, with 45.4% of households estimated to be cell phone-only in the second half of 2014, as compared to 41.0% in late 2013 (Blumberg & Luke, 2015). Researchers are also faced with increasing challenges in terms of being able to cover the entire population at the state level and for small geographies. This is due, in part, to the fact that a significant proportion of the cell phone only (CPO) households have area codes outside of the state or smaller geography that the household is located in. In addition, researchers believe that coverage error in standard list-assisted landline frames is increasing due to the exclusion of numbers in 100 blocks with no directory-assisted numbers (“zero banks”).²

² See: Pew Research Center, October 2015, “Advancements in Telephone Survey Sampling”; John Boyle, Michael Bucuvalas, Linda Piekarski and Andy Weiss. 2009. “Zero Banks: Coverage Error and Bias in RDD Samples Based on Hundred Banks with

While there are limited data available on the share of cell phone-only households within each state, a recent model-based approach (combining survey data and synthetic estimates) was used to generate state-level estimates of cell phone-only households using the National Health Interview Survey (NHIS). As of the end of 2013, the National Health Institute Survey (NHIS) estimated that 42.4% of adults in Oregon lived in households that were wireless only. Projecting to the present day, it is likely that figure is close to 45 percent, similar to the national average.

In order to capture cell phone-only households in the sample frame for the OHIS, the decision was made in 2011, the first wave of the OHIS, to utilize an address-based-sample (AB sample) for the survey. An address-based design allows for the best possible sample coverage, given that the sample source is the U.S. Postal Service's Delivery Sequence File (DSF), a database that is considered to cover at least 98 percent of all households in the U.S. Address-based designs also allow for stratification at the Census Block Group level. One limitation of both AB sample and RDD sample is that they both miss homeless persons in the state. This is estimated to be less than 1% of the population.

In order to achieve the State's objectives, the sample design incorporated stratification by both region and the ethnic populations. The stratification accommodated diverse clustering of ethnic groups, specifically African Americans, Asians, and Native Americans, within these regions.

The sample was developed in the following steps:

1. A file was generated of all Oregon residential addresses currently in use based on the United States Postal Service Delivery Sequence File (DSF). The DSF is a computerized file that contains information on all delivery addresses serviced by the USPS, with the exception of general delivery.³ The DSF is updated weekly and contains home and apartment addresses as well as Post Office boxes and other types of residential addresses for mail delivery.
2. That address file was run against databases from InfoUSA and Experian that include all landline telephone numbers in the state to identify addresses with a matched telephone number.
3. In order to facilitate the fielding of the survey (discussed below), the final AB sample was divided into two segments: addresses with a matched landline telephone number and addresses without a matched landline telephone number.

The overall sampling design included several dimensions that can be described in terms of sample stratification, household selection criteria, and within household selection criteria. These are summarized below:

Sample stratification

- Interview targets for state-specified 15 Oregon regions.
- Within regions, a creation of strata of addresses by listed Asian and Hispanic surnames.
- Stratification of residual households (those without an Asian or Hispanic surname) by Census block group aggregate incidence of Hispanicity, and percent African American, percent Native American, and percent Asian.

Household-level selection

- The respondent's primary residence is in Oregon.
 - Confirmation of Oregon residency via county of residence or zip code

Listed Numbers." Public Opinion Quarterly Vol. 73, No. 6, pp. 729-750; Mansour Fahimi, Dale Kulp, and J. Michael Brick. 2009. "Reassessment of List-Assisted RDD Methodology" Public Opinion Quarterly Vol. 73, No. 4. Pp. 751-760
 3 See <http://pe.usps.gov/text/dmm300/509.htm>.

- The respondent, age 18 or older who can answer questions about the health insurance coverage of the members of the household. If the person on the phone is younger than 18, interviewer asks for another household member who is 18 or older.
- Screening households where every person in the household is age 65 or older:
 - If a household with a matched telephone number contained only adults ages 65 and older, the telephone or web interview initially terminated in 75 percent of such instances when OHIS launched in April; later in the field (as of June 1, 2015), this was changed to 100%. This screening procedure was designed to address the fact that such households more readily respond to surveys compared to other households.
 - Households with unmatched telephone numbers and mail survey respondents are excluded from this criteria

Person selection

- A *target*(or *focus person* in the mail survey) is randomly selected from all household members
 - If it is a single-person household, that person is the target.
 - If it is a household with two or more people, a random selection of household members is performed by the CATI/web program; in the hard copy instrument, the household member with the most recent birthday is selected as the focus person.
 - In 2013 and again at the start of 2015 children under age 18 were weighted 50% more than others in the household in order to increase the probability of selecting a child target.
 - Mid-way through the 2015 field period (June 1, 2015) the selection process was changed to preferentially choose a target 18-44 years old in order to increase the proportion of interviews completed with target respondents in this age range.

Sample Stratification

For the 2015 Oregon Health Insurance Survey (OHIS), SSRS executed a sample design that was very similar to the designs executed in 2011 and 2013. Specifically, the AB sample was stratified to the Census Block Group level, with the goal of obtaining minimum numbers of interviews in the the 15 health regions of the state, and oversampling ethnic populations of interest. As in the past, the plan was designed in order to complete a minimum of 300 to 400 interviews within each of the 15 health regions.⁴ For the 2015 plan, however, the team at the Office of Health Analytics in the Oregon Health Authority (OHA) requested that efforts would be made to increase the proportion of Hispanics in order to be as close as possible to the proportion of Hispanics in Oregon, according to the 2013 American Community Survey (ACS) population estimate of 9.8%. In order to more closely approximate the ACS estimates, the plan for 2015 included an increased proportion of interviews in the High Hispanic and Hispanic Surname strata.

The sample design needed to address the fact that there is little variance in household minority status by area, as indicated below:

⁴ Efforts were made in the final (Wave 3) sample release to help ensure that the final proportion of completed interviews in each region would be similar to original projections.

Table 2: Household minority status by region

Region	Percent of All Households	Percent African American	Percent Native American	Percent Asian	Percent Hispanic
1	3.1	0.4	2.2	0.7	11.2
2	2.9	0.4	3.2	0.8	11.9
3	2.8	0.6	2.4	1.1	11.2
4	4.4	0.3	0.6	0.8	5.0
5	4.3	0.5	1.6	0.8	4.4
6	5.2	0.6	0.9	2.7	5.2
7	9.5	0.9	0.9	2.4	5.4
8	4.6	0.3	1.5	0.7	4.3
9	5.5	0.6	0.9	1.0	7.6
10	2.9	0.4	1.5	0.7	3.5
11	7.4	0.8	1.0	1.7	16.5
12	4.1	0.4	1.5	1.2	9.2
13	20.3	4.8	0.7	5.3	7.2
14	9.7	0.8	0.6	3.1	5.5
15	13.5	1.8	0.5	8.2	10.7
TOTAL	100	1.6	1.0	3.3	7.9

Notably, the only option to effectively increase African American and other ethnic interviews within the 15-region plan would be to oversample region 13 (Multnomah county). Area 13 is also the most populous region of the state; therefore, from an overall perspective, there is no need to oversample this stratum. In addition, Native Americans, which represent the smallest of the racial/ethnic groups, do not cluster well by region. Whereas their incidence is higher in areas 1, 2, and 3 compared to all other areas of the state, their incidence, on average, in these three areas is still only 2.6 percent. Finally, Asians also do not cluster well; and similar to the African Americans, they tend to cluster in regions 13 and 15, areas that did not need to be oversampled to attain the adequate number of interviews.

In order to achieve the State's objectives, the sample design therefore incorporates stratification by both region and the ethnic populations. The design was developed by first extracting all households with a Hispanic or Asian surname and placing them into surname strata. Next, any block group that is ten percent or more African American, Asian, or Native American, or 30 percent or more Hispanic, was flagged. These groups were then placed into separate high incidence strata. This resulted in 21 strata, including the 15 regions, two surname strata, and four high incidence strata. The 21-strata design allowed for the attainment of a minimum of 300 interviews in each of the 15 regions. The study was also designed to attain proportional representation, at a state level, of African Americans, Hispanics, Native Americans and Asians. In aggregate, the sample plan was designed to attain a minimum of 9,000 interviews across the State.

Table 3 below show final strata for OHIS 2015, based on household counts from the 2015 Claritas.⁵ This design clusters minority interviews into the additional strata, as illustrated by the incidence figures below:

⁵ Claritas (through Nielsen) provides modeled household estimates, by quarter, based on the 2010 US Census and annual Census surveys such as the ACS and the Current Population Survey (CPS).

Table 3: Household race by strata

Strata	HH	Caucasian	African American	Native American	Asian	Hispanic
1	40,121	35,933	115	549	162	2,504
2	39,982	35,049	114	504	150	3,395
3	38,439	33,287	211	807	287	2,896
4	68,195	62,907	165	435	393	3,137
5	65,301	59,881	242	833	252	2,489
6	76,920	69,109	379	680	1,281	3,646
7	142,589	127,037	1,054	1,305	2,126	6,928
8	71,572	65,391	183	1,101	250	2,621
9	84,937	75,286	410	763	629	5,834
10	44,752	41,154	138	696	179	1,325
11	95,062	80,420	615	952	1,297	9,420
12	61,045	53,375	201	706	538	4,903
13	233,941	197,085	6,657	1,441	8,190	13,392
14	140,404	126,386	977	818	2,511	6,793
15	139,470	115,428	2,076	708	5,661	11,921
High African American	36,630	23,769	6,084	336	1,124	3,766
High Hispanic	55,338	31,077	618	537	2,125	19,735
High Asian	107,765	75,994	3,643	590	17,050	6,921
High Native American	6,598	3,949	22	2,073	0	219
Hispanic Surname	22,369	8,053	447	22	224	13,421
Asian Surname	13,753	5,501	275	14	7,564	275
TOTAL	1,585,184	1,326,071	24,624	15,869	51,991	125,542

Table 4 shows the estimated proportion of African American, Native American, Asian, and Hispanic households in each strata (i.e., the estimated number of households for each ethnic group/estimated number of total households in each strata based on the 2015 Claritas numbers provided in the table above).

Table 4: Stratification, Percent by Ethnicity

Strata	African American	Native American	Asian	Hispanic
1	0.3%	1.4%	0.4%	6.2%
2	0.3%	1.3%	0.4%	8.5%
3	0.5%	2.1%	0.7%	7.5%
4	0.2%	0.6%	0.6%	4.6%
5	0.4%	1.3%	0.4%	3.8%
6	0.5%	0.9%	1.7%	4.7%
7	0.7%	0.9%	1.5%	4.9%
8	0.3%	1.5%	0.3%	3.7%
9	0.5%	0.9%	0.7%	6.9%
10	0.3%	1.6%	0.4%	3.0%
11	0.6%	1.0%	1.4%	9.9%
12	0.3%	1.2%	0.9%	8.0%
13	2.8%	0.6%	3.5%	5.7%
14	0.7%	0.6%	1.8%	4.8%
15	1.5%	0.5%	4.1%	8.5%
High AA	16.6%	0.9%	3.1%	10.3%
High Hispanic	1.1%	1.0%	3.8%	35.7%
High Asian	3.4%	0.5%	15.8%	6.4%
High Native American	0.3%	31.4%	0.0%	3.3%
Hispanic Surname	2.0%	0.1%	1.0%	60.0%
Asian Surname	2.0%	0.1%	55.0%	2.0%

The final sampling plan, updated in July 2015,⁶ is provided below in Table 5:

⁶ The sampling plan for the 2015 OHIS was revised in July because the survey was generating fewer telephone and web interviews than expected and additional sample needed to be released.

Table 5: Stratification, Expected Interviews – Modeled⁷

Strata	% of Households	% of Interviews	Interviews	AA interviews	NA interviews	Asian interviews	Hispanic interviews	Weight	Region Interviews
1	2.5%	2.5%	208	1	3	1	13	1.0	376
2	2.5%	2.7%	223	1	3	1	19	1.1	322
3	2.4%	3.4%	277	2	6	2	21	1.4	419
4	4.3%	3.6%	298	1	2	2	14	0.8	311
5	4.1%	3.4%	275	1	4	1	10	0.8	298
6	4.9%	3.7%	300	1	3	5	14	0.8	317
7	9.0%	3.9%	322	2	3	5	16	0.4	391
8	4.5%	4.1%	335	1	5	1	12	0.9	349
9	5.4%	4.8%	392	2	4	3	27	0.9	415
10	2.8%	3.9%	318	1	5	1	9	1.4	342
11	6.0%	4.0%	324	2	3	4	32	0.7	724
12	3.9%	4.2%	344	1	4	3	28	1.1	455
13	14.8%	8.3%	681	19	4	24	39	0.6	1648
14	8.9%	5.8%	472	3	3	8	23	0.6	599
15	8.8%	5.9%	487	7	2	20	42	0.7	1234
AA	2.3%	6.4%	526	87	5	16	54	2.8	-
Hispanic	3.5%	8.4%	691	8	7	27	246	2.4	-
Asian	6.8%	9.2%	754	25	4	119	48	1.4	-
NA	0.4%	2.0%	162	1	51	0	5	4.7	-
Hispanic Sur.	1.4%	6.0%	489	10	0	5	293	4.2	-
Asian Sur.	0.9%	3.9%	322	6	0	177	6	4.5	-
TOTAL	100.0%	100.0%	8,201	183	120	425	973		8,201
Design Effect				1.26	1.44	1.37	1.33		
Increase from Proportionality				143%	146%	158%	150%		

Allocations were designed to maximize the number of minority interviews while maintaining a design effect under 2.0.

The right-most column denotes the number of interviews by area, once interviews from the high incidence and surname strata are folded back into the 15 OHIS regions. The sample plan was designed to allow a minimum of 275 interviews to be completed in each region. While the design “on paper” was expected to attain between interviews between 143 percent and 158 percent above proportionality among the minority groups, in practice, we expected the design to attain interviews ranging from 100 percent to 115 percent above proportionality, due to systematic nonresponse whereby ethnic minorities tend to cooperate in surveys less frequently than Caucasians, on average. We anticipated, however, that the differences between estimated and actual numbers of completes would be somewhat larger for some of the ethnic groups (especially Hispanics) in 2015 given that a larger proportion of the interviews were completed via paper surveys, compared with past years.

⁷ The modeled estimates assumed that the final sample size would be at least 8,200.

The final step in pulling the sample for the survey was the extraction of age-listed sample for which the household contains only persons ages 65 and older. For the first two sample releases, 75 percent of all age-listed sample for which the household only contained persons ages 65 and older were extracted. Before releasing Wave 3 sample, 100% of all age-listed sample for which the household contains only person 65 and older was extracted.⁸ Nationally, 82% of all adults ages 65 and older are found on these lists. As such, removing 75% of age-listed sample for households including only people age 65 or older is an effective way to reduce the number of interviews with 65+ households.

Disproportionate Incentives – Targeting Hispanics, African Americans and Lower Income HHs

In 2013, SSRS and OHA conducted an experiment in order to determine how ABS designs, in general, and the OHIS, in particular, can be enhanced to increase representativeness among the final sample of completed interviews.

ABS designs allow for the best possible coverage of the population since virtually all target population elements are included in the DSF (Delivery Sequence File) sample frame. At the same time, researchers have noted disproportionate non-response among lower SES unmatched households— those for which a matched telephone number is not available -- in ABS designs offering multiple modes to complete the survey. For example, in their 2011 POQ article “Cell-Phone-Only Households and Problems of Differential Nonresponse Using an Address-Based Sampling Design,” Michael Link and Jennie Lai point to the “key limiting factor” of ABS designs – that households without a matched phone number to match their address can only be contacted by mail.

In 2013, SSRS and OHA included both a sample-based experiment and a limited number of respondent questions. Based on that research, it was theorized that providing post-incentives to a portion of unmatched sample would boost the completion rate among unmatched sample receiving these incentives.⁹

For 2015, the incentive plan was designed to reduce disproportionate non-response among unmatched households at lower socioeconomic (SES) levels in terms of household income and education, in particular (using auxiliary information such as average household income by block group) in our provision of post-incentives:

Table 6: Incentive plans

Incentive Plan	
\$2 pre-incentives and \$10 post incentives	All unmatched HHs in block groups with ~11.5% households or more below 100% FPL for regional strata, and High NA, High Asian, and Asian Surname strata (about 20% of households in regional strata)
\$2 pre-incentives and \$10 post incentives	All unmatched HHs in High AA, High Hispanic, and Hispanic Surname strata
\$2 pre-incentives and no post incentives	Unmatched HHs in block groups within the High Native American, High Asian, and Asian Surname strata and within regional strata with fewer than ~11.5% of households in block groups that are below 100% FPL
No incentives	All Matched HHs

⁸ This change was made for Wave 3 since OHA requested that 100% of incoming matched households with 65 plus only respondents would be terminated, as of June 1, 2015.

⁹ The SSRS team is reviewing the response rates and other related data to gain greater insight into the advantages and disadvantages of using disproportionate incentives in future waves of OHIS 2015.

Envelope Experiment

For 2015, SSRS conducted another experiment aimed at improving overall response in the OHIS survey. Drawing on research conducted by Don Dillman and others regarding the benefit of using a “Tailored Design Method,”¹⁰ the experiment aimed at helping to further establish the legitimacy of the study for Oregonians and increasing the rate of completion. Specifically, for 50% of the Wave 1 sample (N=10,899), the slogan on the mailing envelopes: “Help improve health care in Oregon” was included on the outside of the mailing envelope, in addition to the standard OHA envelope that contained the state seal. The remaining 50% of the Wave 1 sample (N=10,895) received the standard OHA mailing envelope with the state seal and no slogan. OHA gave approval on experiment materials and design.

This experiment was not continued for Wave 2 and 3 sample releases because no difference was noted in the response rates for sample that received the experimental envelope compared with the control envelope. Sample mailed in Waves 2 and 3 all used the standard envelope design, the envelope that included only the state seal and no additional slogan.

INSTRUMENT DEVELOPMENT

Questionnaire Development/Changes to the 2015 Instrument

The 2015 survey changes included implementing new questions, modifying some of the 2013 questions, and removing a few previously-asked questions. All of these changes were made in the telephone and online modes. Substantive changes made to the 2015 survey are noted below:

- The follow-up question “Was health insurance coverage obtained through the state health insurance marketplace, also known as Cover Oregon or healthcare.gov” was asked for target and household members with OHP/Healthy kids/Independent insurance coverage
- All targets were asked: “Was there a time in the past 12 months when (you/target) looked for health insurance using the state health insurance marketplace? This may have been through Cover Oregon or healthcare.gov.”
- Four Health Literacy questions (HL1 to HL4) were added and asked of all targets
- Targets with health insurance for the past 12 months/since birth were asked: “(Have you/has this person) had a change in insurance carrier or coverage in (the past 12 months/since they were born?)”
- Adult targets (age 18 and older) were asked: “Do (you/target) have an Advanced Directive?” This question was not asked in 2013; wording for OHIS 2015 is similar the Advanced Directive question asked in OHIS 2011.
- All Targets were asked: “Do you/target’s parent/target’s parent or guardian have a Credit Card or Bank Account?” (F16a/b) in an effort to capture how respondents might be able to pay for health insurance coverage through the marketplace
- All Targets were asked: Db1 “Does any disability, handicap, or chronic disease keep (you/target) from participating fully in work, school, housework, or other activities?”

Pretesting

Pretest interviews are conducted in order to insure that proper wording, question sequencing, and informational objectives are being met. CATI pretest interviews also provide an opportunity to: (1) get feedback from interviewers and supervisors regarding the clarity of the instrument (including issues and concerns raised by respondents); and (2)

¹⁰ <https://sesrc.wsu.edu/about/total-design-method/>

monitor interviewers and make modifications to interviewer training procedures and materials. Separate pretests were conducted for the CATI, web, and mail survey instruments.

CATI Pretest

A total of 26 CATI pretest interviews were conducted on March 2 and March 3, 2015. Overall, the instrument worked well and respondents appeared to be engaged in the interview. The interviews averaged 23.2 minutes, which was in line with the estimated survey length. During the pretest SSRS and the Oregon project team identified a few questions that either seemed to be confusing to respondents or could benefit from further clarification in the instrument. Specifically, some edits were made to clarify how the new Health Literacy HL1) and bank card/checking account (F16) questions were being asked to promote better understanding. Post CATI-pretest adjustments made to the questionnaire and interviewing training protocol are noted in the CATI Pretest Memo included in Appendix A.

Web Pretest

A total of seven web pretest interviews were conducted between March 6 and March 13, 2015. Three respondents were recruited from households in Oregon, using age-targeted landline and cell phone sample (to reduce the likelihood of reaching a household where all household members were age 65 years of age or older). Four respondents were recruited by OHA to take part in the web pretest. SSRS project staff monitored over the phone as the respondents completed the survey and noted potential problem areas. After the respondents finished the survey, SSRS project staff asked the respondents for feedback about the instrument using a series of structured questions. The Web Pretest Memo is included in Appendix A.

Mail Survey Pretest

Internal review of the paper survey began in February, 2015 and continued through March 17, 2015 when recruiting for external pretest respondents began. The internal review was conducted with: (1) members of the SSRS staff who were previously unfamiliar with instrument; and (2) those that had been involved in testing the survey in either the web or CATI modes.

A near-final version of the paper survey was sent to OHA for review and mailed to recruited respondents in Oregon who agreed to fill out the paper survey and provide feedback. As a result, SSRS and the Oregon project team identified a few key areas that could benefit from further revision. Specifically, It was noted by both an external pretest respondent and a member of the OHA team that the Employment section could benefit by being displayed at the top of the page instead under a previous question. Due to space constraints, this was not possible; however the heading was widened and lengthened, so it would not appear to be embedded under the previous questions. In addition, before finalizing the paper survey OHA simplified the language on the first page (instructions).

As part of the mail survey pretest process, several adjustments were made to the paper survey instrument, and are noted in the Mail Pretest Memo included in Appendix A.

DATA COLLECTION

Overview

Data collection relied on three interview modes: telephone, web, and mail. Advance letters, reminder letters, and reminder post cards were mailed to all members in the sample, offering the option to complete the survey by telephone or online. In addition, for sample associated with matched telephone numbers, traditional telephone interviewing methods were used.

The specific steps for the data collection process were as follows.

1. Advance letters were sent to all households. The advance letter invited the household to participate in the study and offered the option of calling in to the survey center using a toll-free telephone number or completing a web-based survey. Letters sent to households associated with a matched telephone number also notified people that they would be receiving a call within the next few weeks to complete the survey.
2. Telephone interviews were attempted with all households for which we had a telephone number. The initial calls commenced one to three weeks after the mailing of the advance letters.
3. Reminder post cards were mailed approximately one week after the mailing of the advance letters.
4. A final reminder letter was sent to all non-responding households in a packet that included a copy of the mail questionnaire and a business reply envelope.¹¹

The advance letters and reminder letters were printed on The Oregon Health Authority letterhead and signed by Leslie Clement, Chief of Policy and Program at the Oregon Health Authority. Letters and reminder postcards included a 1-800 toll-free number that the respondent could call for additional information or to complete the survey by telephone.

Timeline

The study timeline was as follows:

Table 7: Timeline

Milestone	Date(s)
Draft Instrument Received by SSRS	January 9, 2015
CATI Survey Pretest Interviews	March 2-3, 2015
Web Survey Pretest Interviews	March 6 – March 13, 2015
Mail Survey Pretest Interviews	March 27-April 8, 2015
Sample Batch 1 Advance Letters Mailed (3 Waves)	April 8, 9, & 10, 2015
Sample Batch 1 First Web Interview Completed	April 11, 2015
Sample Batch 1 First CATI Interview Completed	April 11, 2015
Sample Batch 1 Outgoing Calls Begin	April 27, 2015
Sample Batch 1 Reminder Postcards Mailed	April 15, 2015
Sample Batch 2 Advance Letters Mailed	May 27, 2015
Sample Batch 2 First Web Interview Completed	June 1, 2015
Sample Batch 2 First CATI Interview Completed	May 27, 2015
Sample Batch 2 Outgoing Calls Begin	June 8, 2015
Sample Batch 2 Reminder Postcards Mailed	June 5, 2015
Sample Batch 1 Mail QN and Suspend Letter	May 27, 2015
Sample Batch 1 Mail QN and Reminder Letter Mailed	May 29, 2015
1 st Preliminary File Delivery	July 10, 2015
Sample Batch 2 Mail QN and Reminder Letter Mailed	July 22, 2015
Sample Batch 3 Advance Letters Mailed	July 8, 2015
Sample Batch 3 First CATI Interview Completed	July 11, 2015
Sample Batch 3 First Web Interview Completed	July 11, 2015
Sample Batch 3 Outgoing Calls Begin	July 22, 2015
Sample Batch 3 Reminder Postcards Mailed	July 15, 2015
Sample Batch 3 Mail QN and Reminder Letter Mailed	August 17, 2015
Field Termination	September 14, 2015
Editing and Coding Guidelines Delivery	October 30, 2015

¹¹In May, SSRS sent letters to respondents (n=256) who had started the survey on the web but had not yet completed. These letters included a \$2 incentive for all respondents. For those eligible for a post incentive, the letter also included a reminder about the \$10 they would receive upon completion.

Final Main Data File Delivery	November 10, 2015 ¹²
Final Person File Delivery	November 30, 2015
Extraneous Text Delivery	December 2, 2015
Final Paradata Delivery	December 16, 2015
Methodology Report Delivery	December 22, 2015
Codebook/Data Dictionary Delivery	By January 8, 2015
Study Interviewer Training Manual Delivery	By January 8, 2015

Training Materials and Interviewer Training

CATI interviewers received both written materials on the survey and formal training for conducting this survey. The written materials were provided prior to the beginning of the field period and included:

1. An annotated questionnaire that contained information about the goals of the study as well as detailed explanations of why questions were being asked, the meaning and pronunciation of key terms, potential obstacles to be overcome in getting good answers to questions, and respondent problems that could be anticipated ahead of time as well as strategies for addressing them.
2. A list of frequently asked questions and the appropriate responses to those questions.
3. Contact information for project personnel.
4. A copy of the sample information for each household to aid the interviewers in finding information when respondents called in with questions or called in to complete the survey.

Interviewer training was conducted both prior to the study pretest (described below) and immediately before the survey was officially launched. Call center supervisors and interviewers were walked through each question in the questionnaire. Interviewers were given instructions to help them maximize response rates and ensure accurate data collection. They were instructed to encourage participation by emphasizing the social importance of the project and to reassure respondents that the information they provided was confidential.

Interviewers were monitored during the first several nights of interviewing and provided feedback where appropriate to improve interviewer technique and clarify survey questions. This formal interviewer monitoring process was repeated periodically during the field period.

SSRS's facilities allow Interviewing Supervisors to view a computer-assisted telephone interview as well as hear it. In this way, the Supervisor is able to both ascertain that given questions are being asked correctly, and also confirm that the answers are being recorded correctly. Throughout the study, the interviewers were monitored by field personnel and project directors.

Call Rules for the CATI Interviews

Several call rules were implemented for matched sample in order to ensure high cooperation and quality data. For matched sample, SSRS:

- Instituted an average call rule of original plus 14 callbacks before considering a sampling unit exhausted
- Varied the times of day, and the days of the week that call-backs are placed (differential call rule)
- Permitted respondents to set the schedule for a call-back; allowing them to phone back on our 800 number
- Allowed a rest period for sample after attaining six no answer calls¹³

¹² A subsequent data file with imputed income and imputed responses in the A series was sent on 11/24/2015.

¹³ In cases where sample matched to two or more telephone numbers, secondary numbers were dialed after a minimum of six attempts were made on the 'primary' number matched to the address.

- Made refusal attempts on each initial refusal, with the conversion occurring at least 14 days after the initial contact.

Refusal Avoidance and Conversion Strategies

With the increased popularity of telemarketing and the use of telephone answering machines and calling number identification (i.e., caller-ID), the problem of non-response has become acute in household telephone surveys. Similarly, the increasing prevalence of unsolicited advertising in the mail (i.e., junk mail) makes it more difficult to conduct surveys using only invitation letters as we are doing here with the AB sample without a matched telephone number. In addition to the incentives and call rules for the CATI interviews outlined above, we employed several other techniques to maximize the response rate for the survey. In the CATI interviewing, this included providing a clear and early statement that the call was not a sales call. In all three versions of the survey (telephone, web, and mail), the introduction included an explanation of the purpose of the study, the expected amount of time needed to complete the survey, and a discussion of the incentives. A toll-free number was provided to all respondents.

In an effort to maximize the response rate in the interview phase, respondents were given every opportunity to complete the interview at their convenience. For instance, those refusing to continue at the initiation of or during the course of the telephone interview were offered the opportunity to be re-contacted at a more convenient time to complete the interview. They were also offered the opportunity to complete the survey on-line or to call into the 1-800 toll-free telephone number to complete the survey. Those completing the interview on the web were able to complete the survey at their own speed and stop and re-start as needed.

A key way to increase responses rates is through the use of refusal conversions. Though all of SSRS's interviewers regularly go through "refusal aversion" training, refusals are still a regular part of survey research. SSRS used a core group of specially-trained and highly-experienced refusal conversion interviewers to call all who initially refused the survey in an attempt to persuade respondents to complete the survey.

Caller ID

In an effort to maximize the response rate, a caller ID number was set up, such that when SSRS made outgoing calls, respondents saw the number on the caller ID. For OHIS, respondents with Caller ID capabilities saw: "Oregon Health (971) 269-0561". If a respondent missed the call or wanted to check that it is a valid number, he or she could call the number back. For this reason, call forwarding was set up so that SSRS could disposition the calls and respond to questions related to the survey.

DATA PROCESSING AND PREPARATION

Overview

Data file preparation began soon after the study entered the field. CATI and web range and logic checks were used to check the data during the data collection process. Additional data checks were implemented as part of the data file development work, checking for consistency across variables and family members, and developing composite measures of family and household characteristics. As the paper surveys differ from the online and telephone instruments (i.e., only collecting six household member information as opposed to ten, not being asked follow-up questions such as S6a: age range, etc.), the paper survey cleaning was tailored for that mode to ensure data entry was accurate.

Data Cleaning, Editing and Coding Procedures

As noted above, the major distinction between the telephone mode and the web and paper survey is that, in the case of the CATI interviews, a trained interviewer guides the respondent through the process, whereas the web and paper surveys are self-administered. In addition, in contrast to paper surveys, skip patterns are programmed into both the CATI and web surveys to eliminate the possibility of a respondent being administered questions they are not eligible to receive, and to ensure they are asked all appropriate questions as identified within the survey instrument. The programming also ensures that for questions which require the selection of one response that only one response is allowed to be recorded. Methods used to address problems identified during data collection and data processing web, CATI and paper surveys are detailed in the 2015 OHIS Editing and Data Cleaning Guidelines Document included in Appendix E. These include editing and coding guidelines for processing paper surveys with missing and/or contradictory data.

COMPLETED INTERVIEWS

Overview

As noted above, data were collected using multiple modes—telephone, web, and mail. Table 8 shows the number of completions for each mode of data collection with a separate category for in-bound (toll free) telephone calls from sample members requesting to complete the survey by telephone versus outbound phone interviews where a telephone interviewer called the respondent. For the most part, questions are identical for telephone, web, and mail instruments, although some modifications were made for ease of interviewing within each mode. The mail survey is a slightly condensed version of the CATI/Web instruments.¹⁴ One major distinction between the telephone mode and the web and mail modes is that, in the case of the CATI interviews, a trained interviewer guided the respondent through the process, whereas the web and mail surveys were self-administered. (However, as described below, technical support was available for those completing the web and mail surveys.)

SSRS provided bi-weekly updates to the OHA team with information on the number of completed interviews by mode, strata, and region and distributions of completed interviews by key demographics. The final update is included in Appendix F.

Table 8: Completed and Partial Interviews – Mode by Telephone Status

	Total	With Matched Landline Telephone Number	With No Matched Landline Telephone Number
Total Sample	37103	16699	20404
Total Completed Interviews	8904	3764	5140
Total Partial Interviews	123	41	82
Phone-outbound	1325	1325	0
Phone-inbound	903	188	715
Web/Internet	4185	942	3243
Mail	2614	1350	1264

Although web and mail respondents were completing the questionnaires without the direct assistance of an interviewer, all correspondence with respondents included contact information for project staff who were available to

¹⁴ The order of questions asked in the mail survey is also different than the Web and CATI instruments, to help promote usability for the paper survey respondents.

assist respondents with any problems they had completing the survey. For those completing the survey on-line, there was access to both staff telephone numbers and an email link for technical support.

Table 9: Completed and Partial Interviews – Target race by strata

Strata	Total	White	Black/AA	Native				DK/ Ref.
				American	Asian	Hispanic	Other	
1	233	215	0	4	1	2	3	8
2	245	220	1	6	5	5	0	8
3	298	258	1	12	3	13	4	7
4	342	312	0	6	6	9	1	8
5	301	273	4	10	4	5	0	5
6	313	291	1	8	5	1	1	6
7	354	312	2	6	7	13	2	12
8	374	328	2	15	8	7	4	10
9	437	385	4	14	7	13	2	12
10	338	297	1	13	3	9	1	14
11	372	329	3	3	13	10	3	11
12	370	332	2	10	6	9	3	8
13	758	656	23	11	24	15	9	20
14	528	474	4	8	15	9	5	13
15	537	474	3	11	22	12	5	10
High AA	543	403	66	15	17	27	8	7
High Hispanic	762	599	16	17	21	90	7	12
High Asian	806	654	25	11	74	15	6	21
High NA	179	117	2	44	5	4	1	6
Hispanic Surname	576	155	4	6	8	382	9	12
Asian Surname	361	157	3	2	181	11	2	5
TOTAL	9,027	7,241	167	232	435	661	76	215

The OHIS was administered in two languages-English and Spanish. Mailings to High Hispanic and Hispanic Surname strata included bilingual materials (English and Spanish). All Hispanic strata telephone interviewing was conducted by bilingual interviewers. In addition, any “language barriers” that were encountered in other strata were called back by bilingual interviewers. A total of 141 interviews were conducted in Spanish.

Completed Interviews by Telephone Status

Table 10 below shows the number of completed interview done in households that had only a cell phone, only a landline phone, both a landline and cell phone, and the residual categories for no telephone or telephone status unknown. We completed surveys with 3,734 cell phone-only households, 4,398 dual landline and cell phone households, 588 landline-only households, 71 non-telephone households and 236 households that did not report their telephone status.

Table 10: Completed interviews – Telephone type by telephone status

	Total	With Matched Landline Telephone Number	With No Matched Landline Telephone Number
Total Interviews	9,027	3,805	5,222
Cell phone-only	3,734	509	3,225
Landline phone-only	588	411	177
Cell phone and landline phone	4,398	2,747	1,651
No telephone	71	12	59
Phone status unknown	236	126	110

RESPONSE RATE

Defining the Response Rate

Response rates are one method used to assess the quality of a survey, as they provide a measure of how successfully the survey obtained responses from the sample. The American Association of Public Opinion Research (AAPOR) has established standardized methods for calculating response rates (AAPOR, 2008). This survey uses AAPOR's response rate definition RR4, with an AAPOR-approved alternative method of addressing ineligible households (described below). We report the overall response rates achieved for the sample. Before presenting those estimates, we describe our methods for calculating the response rates.

SSRS calculates response rates in accordance to AAPOR RR4 calculations. However, the AAPOR Standard Definitions manual does not provide explicit guidelines for ABS designs, nor does it provide more than general guidance for screener surveys. This memo documents the SSRS approach to both of these issues.

Screener Studies

Generally, screener surveys are different than general population surveys in that there are two levels of eligibility: household and screener. That is, a sample record is "household eligible" if it is determined that the record reaches a valid household. Screener eligible refers to whether known household-eligible records are eligible to in fact complete the full survey. In the case of the Oregon Health Insurance Survey, screener eligibility refers to whether a matched household has a member under the age of 65, for those surveys in which such criteria are mandatory. In addition, households must be located within the geographic target area of the study.

The standard AAPOR RR4 formula is as follows:

$$I + P$$

$$I + P + R + NR + [UNR + UR]e$$

Where:

I: Completed Interview

P: Partial Interview

R: Known Eligible Refusal/ Breakoff

NR: Known Eligible Non-Respondent

UR: Household, Unknown if Screener Eligible

UNR: Unknown if Household

e = Estimated Percent of Eligibility

At issue with this calculation for screener surveys is that it does not distinguish the two separate eligibility requirements: UNR and UR and both multiplied by an overall “e” that incorporates any and all eligibility criteria. An alternative RR4 calculation utilized by a large number of health researchers and academicians simply divides “e” into two separate numbers, one for household eligibility and one for screener eligibility:

$$\frac{I + P}{I + P + R + NR + [(UNR)e2 + (UR)]e1}$$

Where:

e2 = Estimated Percent of Household Eligibility

e1 = Estimated Percent of Screener Eligibility

“E” calculations are completed via the standard “proportional representation” method dictated by AAPOR. In short, e2 is all identified household / (all identified households + all identified non-households) and e1 = all identified households eligible to do the full survey / (all identified households known to be eligible to do the full survey + all identified households known to not be eligible to do the full survey).

ABS Studies

ABS studies are particularly challenging for response rate calculations given that they are typical multi-modal. That is, while the frame is address-based, the method of interviewing is often web and/or telephone as well as address. Therefore, the question is how to treat telephone dispositions when the frame is based on address. Prior studies (specifically, MA HIS 2011) show that over 95% of the time, completed interviews via phone were completed by a person at the address sampled. Given this high “hit rate,” it is our opinion that all sample records determined by phone to be an occupied household should be considered a successful match between phone and address. This is important, because it therefore means any eligible refusal should in fact be treated as an eligible refusal (meaning, we assume that the phone matched the address and therefore it is a refusal from a valid sample record). That said, any non-working, fax, and business disposition is, by its nature, proven to be an unsuccessful match between phone and address. Therefore, any such records should not be treated as ineligible, but in fact UNR, a sample record for which household eligibility has not yet been established.

By definition then, a large percent of sample records will end up as UNR. The vast majority of unmatched sample, however, will be considered a “no answer,” given that invitations to participate were mailed, without any response whatsoever. And as mentioned, within matched sample, all non-working/fax/business telephone dispositions for which there is no more “important” web or mail dispositions (like a completed interview) will be considered UNR as well.

Final Response Rates

Table 11: Final Dispositions and Response Rates

	Total	1	2	3	4	5	6	7	8	9	10	11
<u>INELIGIBLE</u>												
Undeliverable	4675	118	147	213	164	186	106	171	208	193	177	175
No Eligible Respondent	1169	32	25	36	37	52	43	62	60	85	53	54
Total Ineligible Cases	5844	150	172	249	201	238	149	233	268	278	230	229
<u>ELIGIBLE</u>												
Refusal	78	3	0	3	4	3	1	4	6	3	8	3
Breakoffs	317	5	9	11	12	16	8	18	11	18	9	10
Answering Machine/Other	85	1	1	2	1	3	4	2	6	4	2	1
Partial Interview	123	2	2	2	5	4	5	4	3	9	5	7
Interview Complete	8904	231	243	296	337	297	308	350	371	428	333	365
Total Eligible Cases	9507	242	255	314	359	323	326	378	397	462	357	386
<u>Unkown Eligibility</u>												
Refusal	346	7	4	11	12	11	7	22	21	16	20	18
Breakoffs	1359	32	31	52	27	43	33	74	65	94	58	73
Answering Machine	1622	27	22	41	27	54	40	59	53	87	41	52
Language Unable	191	0	1	2	0	2	1	0	0	1	1	5
Busy	40	0	2	1	0	2	1	0	0	1	0	5
No Answer	14006	328	395	544	526	405	381	510	464	672	507	606
Call Blocking	2	0	0	0	0	0	0	0	0	0	0	0
Bad Number, but Mail Delivered	3903	65	82	115	80	108	80	137	129	201	118	139
No Screener Completed	146	6	2	3	4	9	3	2	7	16	5	3
Physically/Mentally Unable/Group Quarters	137	5	2	1	1	5	7	7	8	11	6	3
Total Unknown Eligible Cases	21752	470	541	770	677	639	553	811	747	1099	756	904
TOTAL CASES	37103	862	968	1333	1237	1200	1028	1422	1412	1839	1343	1519
RR4	37.0%	41.6%	40.2%	37.4%	41.6%	41.1%	43.1%	37.9%	41.7%	35.8%	39.1%	36.7%

Table 11: Final Dispositions and Response Rates (Cont'd)

	Total	12	13	14	15	High AA	High Hisp	High Asian	High NA	Hisp Sur	Asian Sur
<u>INELIGIBLE</u>											
Undeliverable	4675	132	322	184	231	266	548	468	109	413	144
No Eligible Respondent	1169	58	82	96	76	41	76	99	18	42	42
Total Ineligible Cases	5844	190	404	280	307	307	624	567	127	455	186
<u>ELIGIBLE</u>											
Refusal	78	3	3	2	4	3	8	8	2	5	2
Breakoffs	317	8	18	11	14	20	24	26	10	42	17
Answering Machine/Other	85	2	11	5	7	4	7	10	1	9	2
Partial Interview	123	6	14	8	4	4	8	15	3	10	3
Interview Complete	8904	364	744	520	533	539	754	791	176	566	358
Total Eligible Cases	9507	383	790	546	562	570	801	850	192	632	382
<u>Unkown Eligibility</u>											
Refusal	346	16	26	28	14	17	22	28	6	22	18
Breakoffs	1359	44	86	111	97	37	90	158	27	83	44
Answering Machine	1622	62	139	112	133	75	113	227	15	142	101
Language Unable	191	0	15	3	10	15	24	30	0	22	59
Busy	40	4	4	3	1	0	3	6	3	3	1
No Answer	14006	491	1114	701	623	745	1217	1351	434	1396	596
Call Blocking	2	0	0	0	0	0	1	0	0	1	0
Bad Number, but Mail Delivered	3903	138	265	265	296	171	388	509	72	358	187
No Screener Completed	146	5	8	8	7	5	11	26	0	9	7
Physically/Mentally Unable/Group Quarters	137	2	13	12	6	8	8	13	4	9	6
Total Unknown Eligible Cases	21752	762	1670	1243	1187	1073	1877	2348	561	2045	1019
TOTAL CASES	37103	1335	2864	2069	2056	1950	3302	3765	880	3132	1587
RR4	37.0%	40.0%	38.2%	36.7%	38.3%	41.0%	38.3%	32.9%	32.0%	29.2%	31.6%

SURVEY WEIGHTS AND VARIANCE ESTIMATION

Survey Weights

Survey data are weighted to adjust for differential sampling probabilities, to reduce any biases that may arise because of differences between respondents and non-respondents (i.e., nonresponse bias), and to address gaps in coverage in the survey frame (i.e., coverage bias). Survey weights, when properly applied in surveys can reduce the effect of nonresponse and coverage gaps on the reliability of the survey results (Keeter et al. 2000, Groves 2006).

We constructed analytical survey weights for the Oregon Health Survey using standard procedures. That is, separate weights are created for all persons and for the target-person in the household.

Constructing the Base Weights

The first step in the weighting process for each sample is to create a household weight for each completed survey. That household weight is used to construct weights for each person in the household and for the target-person in each household.

We first adjusted the base weight so that all the households attain the same probability of selection. This adjustment corrects for the over-sampling of addresses in some strata in comparison to others:

$$(f = n_{\text{interviews}} / N_{\text{frame}})$$

Table 12 shows the stratification weights.

Table 12: Base weights by strata

STRATA	Percent of Households	Percent of Interviews	Stratification Weight
1	2.5	2.6	0.99
2	2.5	2.7	0.93
3	2.4	3.3	0.73
4	4.3	3.8	1.14
5	4.1	3.3	1.24
6	4.9	3.4	1.40
7	9.0	3.9	2.29
8	4.5	4.1	1.09
9	5.4	4.8	1.11
10	2.8	3.7	0.75
11	6.0	4.1	1.46
12	3.9	4.1	0.94
13	14.8	8.4	1.76
14	8.9	5.9	1.51
15	8.8	6.0	1.48
High AA	2.3	6.0	0.38
High Hispanic	3.5	8.4	0.41
High Asian	6.8	8.9	0.76
High NA	0.4	2.0	0.21
Hispanic Surname	1.4	6.5	0.22
Asian Surname	0.9	4.0	0.22

This step has the additional feature of correcting for non-response¹⁵ as well, since the percent of interviews, rather than total sample, in each strata, is matched to the percent of households in each strata. Therefore, non-response and stratification are corrected in one step (compared to matching sample to households, and then correcting the number of interviews to sample as a separate non-response correction).

A second adjustment corrected for non-response between matched and unmatched sample. Since matched sample has the benefit of containing a phone number, response rate is often much higher in that frame compared to unmatched sample. However, this was not the case in Oregon, given a robust web response in the unmatched sample. Nevertheless, the adjustment was made:

Table 13: Base weight of Matched and Unmatched sample

Sample Type	Percent of Households	Percent of Interviews	Stratification Weight
Matched	40.9%	42.3%	0.97
Unmatched	58.9%	57.7%	1.02

¹⁵ Non-response creates biases survey estimates because the characteristics of those interviewed differ from those who were not interviewed. The size of the bias is based on this difference and the response rate (Groves, 1989). Non-response adjustments are designed to reduce this bias. A weighting class adjustment (Brick and Kalton, 1996) method is the type of nonresponse adjustment procedure typically used in most survey research, and is utilized here as a function of the stratification adjustment as described in the body of the text.

The base weight also corrected for the termination of approximately 50% of households in which there were no adults under the age of 65. This correction gave all households for which there was someone under the age of 65 a weight of 1, and those who were interviewed that did not contain a person under the age of 65 a weight of 2.

Finally, a number of adults adjustment was made, such that households with 1 adult received a base weight correction of 1, households with 2 adults received a base weight correction of 2, and households with 3 or more adults received a base weight correction of 3 (t). This correction was capped at 3 to prevent large weights for large households.

The final base weight was a product of these four corrections (strata weight * sample type weight * under 65 weight * number of adults weight).

Constructing the Target Weights

To create a target weight for each person in an interviewed household we started with the base weight and then post-stratified so that our weighted sample population totals equaled population control totals based on data for Oregon. Specifically, we aligned the sample to current Census population estimates for Oregon on age, race/ethnicity x puma, race/ethnicity x education, gender, homeownership, and foreign born status x education. We provided two weights, one balanced to sample and the other inflated to match the population totals. The demographic information and homeownership data came from the American Community Survey, 2013 and excluded those living in institutions or other group quarter living arrangements.

To handle missing data among some of the demographic variables we employed a technique called hot decking. Hot deck imputation replaces the missing values of a respondent randomly with another similar respondent without missing data. These are further determined by variables predictive of non-response that are present in the entire file. Using an SPSS macro detailed in 'Goodbye, Listwise Deletion: Presenting Hot Deck Imputation as an Easy and Effective Tool for Handling Missing Data' (Myers, 2011), we imputed missing values for home ownership, education, and race.

We examined the distribution of the resulting target weights and determined that there were some large weights. We trimmed 2% off the low- and high-end weights for both the final base and post-stratified weights. We also included an untrimmed weight as well.

We utilized an iterative proportionate fitting (IPF) procedure create the post-stratification weights. IPF is a now-ubiquitous sample balancing routine originally developed by W. Edwards Deming and Frederick F. Stephan to adjust samples in economic and social surveys on selected demographic characteristics against data obtained from the U.S. Census. The theory behind IPF is explained in Deming's book *Statistical Adjustment Of Data* (1943), available in reprint from Dover Publications. Details on the Deming-Stephan method are spelled out in Chapter VII: "Adjusting to Marginal Totals." (Werner, 2004). IPF ("raking") uses least-squares curve fitting algorithms to obtain a unique weight for each case that minimizes the root mean square error (RMSE) across multiple dimensions simultaneously. Then it applies these weights to the data and repeats the procedure using the newly obtained marginal counts to obtain yet another set of weights. This process is repeated for a specified number of iterations or until the difference in the RMSE between successive steps becomes less than a specific minimum value. This study employed an IPF procedure using the statistical software, SPSS. SPSS is an "industry standard" statistical software package used in performing several statistical techniques including sample balancing post-stratification and allows for the application of a pre-existing base weight to the input data for the sample balancing process.

Below are the control totals used and frequencies of the data, before and after the post-stratification routine. Note the adjustment of the control targets to account for missing data in the sample, a standard method of dealing with missing data for weighting procedures:

Table 14: Pre-Rake/Post-Rake frequencies

Gender	Population	Percent	Pre-Rake Sample Percent	Post-Rake Sample Percent
Male	1894077	49.3	44.8	48.7
Female	1948301	50.7	55.2	51.3
Total	3842378	100.0	100.0	100.0

Home Ownership	Population	Percent	Pre-Rake Sample Percent	Post-Rake Sample Percent
Rent	1465014	37.9	27.1	36.5
Own	2377364	62.1	72.9	63.5
Total	3842378	100.0	100.0	100.0

Age	Population	Percent	Pre-Rake Sample Percent	Post-Rake Sample Percent
0 thru 17	856465	22.3	12.1	21.5
18 thru 24	335154	8.7	5.3	8.5
25 thru 34	516998	13.5	9.3	13.2
35 thru 44	511406	13.3	12.0	13.5
45 thru 54	506456	13.2	13.2	13.2
55 thru 64	531077	13.8	26.8	14.3
65 thru 96	584822	15.2	21.3	15.9
Total	3842378	100.0	100.0	100.0

Table 14: Pre-Rake/Post-Rake frequencies cont'd

Race by Education		Population	Percent	Pre-Rake Sample Percent	Post-Rake Sample Percent
No H.S. Diploma	White/Other	176492	4.6	2.7	4.2
	Black	6983	.2	.1	.1
	Asian	21397	.6	.3	.4
	Native American	8558	.2	.2	.2
	Hispanic	108223	2.9	1.5	2.4
H.S. Diploma	White/Other	618048	16.1	15.5	16.4
	Black	11966	.3	.4	.3
	Asian	23528	.6	.6	.5
	Native American	14325	.4	.5	.4
	Hispanic	76485	2.0	1.5	1.8
Some College	White/Other	909395	23.7	27.2	24.5
	Black	22455	.6	.6	.6
	Asian	45254	1.2	1.2	1.2
	Native American	29053	.8	.9	.8
	Hispanic	70917	1.7	1.5	1.6
College Degree	White/Other	731759	19.0	29.2	19.9
	Black	9091	.2	.5	.2
	Asian	55848	1.5	2.0	1.5
	Native American	8501	.2	.7	.2
	Hispanic	34518	.9	1.0	.9
Under 18	White/Other	562535	14.8	8.4	15.1
	Black	30660	.8	.4	.8
	Asian	54596	1.4	1.0	1.4
	Native American	23265	.6	.3	.6
	Hispanic	184340	4.9	2.0	3.8
Total		3838192	100.0	100.0	100.0

Table 14: Pre-Rake/Post-Rake frequencies cont'd

Race by PUMA		Population	Percent	Pre-Rake Sample Percent	Post-Rake Sample Percent
41100	White/Other	401835	10.3	14.4	10.7
	Black	2091	.1	.1	.1
	Asian	9154	.2	.4	.2
	NA	15181	.4	.7	.4
	Hispanic	71923	1.9	1.6	1.9
41200	White/Other	598151	15.6	11.0	15.6
	Black	7744	.2	.1	.2
	Asian	23471	.6	.3	.5
	NA	17482	.5	.4	.5
	Hispanic	54397	1.4	.5	.9
41300	White/Other	407427	10.6	11.9	11.0
	Black	3955	.1	.1	.1
	Asian	8936	.2	.3	.2
	NA	15111	.4	.5	.4
	Hispanic	39855	1.0	.6	.8
41400	White/Other	347937	9.1	11.5	9.4
	Black	4422	.1	.2	.1
	Asian	12290	.3	.5	.3
	NA	11983	.3	.4	.3
	Hispanic	105156	2.7	2.1	2.5
All Other Pumas	White/Other	1242879	32.4	34.2	33.5
	Black	62943	1.6	1.5	1.6
	Asian	146772	3.8	3.5	3.7
	NA	23945	.6	.7	.6
	Hispanic	203152	5.4	2.7	4.5
Total		3838192	100.0	100.0	100.0

US Born by Education		Population	Percent	Pre-Rake Sample Percent	Post-Rake Sample Percent
US Born	Less Than High School	201797	5.2	3.2	4.9
	High School Diploma	660955	17.1	16.7	17.4
	Some College	984024	25.6	29.4	26.5
	College Degree	743861	19.4	30.3	20.2
	Under 17	818564	21.4	11.8	20.8
Foreign Born	Less Than High School	120267	3.1	1.5	2.4
	High School Diploma	84821	2.2	1.6	2.1
	Some College	94073	2.4	2.1	2.3
	College Degree	96115	2.5	3.0	2.5
	Under 17	37901	1.0	.3	.9
Total		3842378	100.0	100.0	100.0

Constructing the Person Weights

The person file required separate weighting procedures. These procedures mimic the target person weighting procedure with two exceptions. First, as respondents were allowed to refuse age and gender of household member, age and gender were not decked in the person file. Secondly, the number of adults base weight correction, which was not done since each case in the person file is a household member.

Variance Estimation and the Average Design Effect

Complex survey designs and post-data collection statistical adjustments affect variance estimates and, as a result, tests of significance and confidence intervals. Variance estimates derived from standard statistical software packages that assume simple random sampling are generally too low, which leads significance levels to be overstated and confidence intervals to be too narrow.

The impact of the survey design on variance estimates is measured by the design effect. The design effect describes the variance of the sample estimate for the survey relative to the variance of an estimate based on a hypothetical random sample of the same size. In situations where statistical software packages assume a simple random sample, the adjusted standard error of a statistic should be calculated by multiplying by the design effect. Each variable will have its own design effect. Average design effects are summarized below. In 2015, the design effect for estimates for the target person in the household was 2.47. Trimmed, this was reduced to 1.76.

Table 15: Design effects

Race/Ethnicity	Estimate	Standard Error	95% Confidence Interval		Design Effect	Unweighted Count
			Lower	Upper		
White	80.1%	.6%	78.9%	81.2%	1.909	7317
Black	2.0%	.2%	1.7%	2.5%	2.073	167
Asian	5.0%	.3%	4.4%	5.7%	2.001	435
Native American	2.3%	.2%	1.9%	2.7%	1.787	232
Hispanic	10.6%	.5%	9.7%	11.5%	1.976	661

Education	Estimate	Standard Error	95% Confidence Interval		Design Effect	Unweighted Count
			Lower	Upper		
No H.S. Diploma	7.3%	.4%	6.5%	8.1%	2.279	419
H.S. Diploma	19.5%	.6%	18.5%	20.7%	1.783	1667
Some College	28.8%	.6%	27.6%	30.0%	1.646	2826
College Degree	22.8%	.5%	21.8%	23.8%	1.308	2990
Under 18	21.6%	.6%	20.4%	22.9%	2.191	1091

Age	Estimate	Standard Error	95% Confidence Interval		Design Effect	Unweighted Count
			Lower	Upper		
0 thru 17	21.5%	.6%	20.3%	22.8%	2.192	1090
18 thru 24	8.5%	.4%	7.7%	9.4%	2.220	480
25 thru 34	13.2%	.5%	12.2%	14.2%	2.128	837
35 thru 44	13.5%	.5%	12.6%	14.4%	1.800	1086
45 thru 54	13.2%	.5%	12.3%	14.1%	1.644	1195
55 thru 64	14.3%	.4%	13.6%	15.0%	1.022	2420
65 thru 96	15.9%	.4%	15.0%	16.7%	1.282	1919

Table 15: Design effects cont'd

Gender	Estimate	Standard Error	95% Confidence Interval		Design Effect	Unweighted Count
			Lower	Upper		
Male	48.7%	.7%	47.3%	50.1%	1.771	4042
Female	51.3%	.7%	49.9%	52.7%	1.771	4985

IMPUTATION¹⁶

Income Imputation

Income is commonly imputed utilizing a range of methods, including hot deck imputation, regression imputation, and mean imputation within adjustment cells (see Brick and Kalton, 1996; Mandal and Stasny, 2004). Comparative analysis finds highly similar results by each technique. For this study, we utilized hot deck imputation, as we have done for the missing demographic variables for weighting. As noted above, hot deck imputation replaces the missing values of a respondent randomly with another similar respondent without missing data. These are further determined by variables predictive of non-response that are present in the entire file. Using an SPSS macro detailed in 'Goodbye, Listwise Deletion: Presenting Hot Deck Imputation as an Easy and Effective Tool for Handling Missing Data' (Myers, 2011), we imputed missings for income. Out of the 2,620 missing cases for income, 2600 were hotdecked with 7 variables (region, average HH education, married, home ownership, overall health, and employment, poverty_recode). The remaining 20 cases were hotdecked with 3 variables (region, average HH education, and overall health) to ensure that the same donor case was not used more than once.

APPENDICES

APPENDIX A: Pretest Memos

APPENDIX B: Letters/Mailings

APPENDIX C: Web/CATI Survey Instrument

APPENDIX D: Paper Survey Instrument

APPENDIX E: Data Cleaning Guidelines and Editing/Coding Procedures

APPENDIX F: Final Update

¹⁶ During the final data clean after field end the SSRS team noted that respondents who selected "Don't know" at A2b and A6 in the web program program were not asked A4 and A8-A9 respectively. To adjust for this programming error, A4, A8 and A9 were imputed for these cases. Due to a very small sample size A8a was not imputed.