shadac
STATE HEALTH ACCESS DATA ASSISTANCE CENTER

Technical Report for the 2017 Minnesota Health Access Survey: Survey Design and Data Processing Methodology SHADAC, March 2019

This report provides information concerning the data collection process and methodology behind the Minnesota Health Access Survey (MNHA), emphasizing the most recent administration of the survey completed in 2017.

- Section 1 provides a brief overview of the survey's purpose and history since its inception
- Section 2 describes the sampling strategy
- Section 3 describes the content of the MNHA survey
- Section 4 explains the survey administration process
- Section 5 provides information about response and cooperation rates
- Section 6 explains how the data are weighted to represent the Minnesota population
- Section 7 describes key data edits and variable construction which precedes the analysis, and
- Section 8 describes the analysis strategy

1. Overview of the Survey

The Minnesota Health Access Survey (MNHA) is a biennial telephone survey of non-institutionalized Minnesota residents. The survey collects detailed information on health insurance coverage options, access to coverage and health care services, and basic demographic data. The goal of the survey is to document trends in health insurance coverage, and access to insurance and health care at the state and regional level, as well as for select subpopulations (e.g., rural, low-income families, populations of color and American Indians). The MNHA represents a partnership between the Minnesota Department of Health (MDH) Health Economics Program and the University of Minnesota's State Health Access Data Assistance Center (SHADAC).

The MNHA data play an important role in monitoring trends in health insurance coverage, evaluating and informing health policy development in Minnesota on topics such as affordability of coverage, access to healthcare, and redesign of public program coverage. The MNHA provides precise and timely estimates on a range of coverage and access relevant questions, can be modified to be responsive to developing state health policy issues, and ensures the availability of micro-data for time sensitive research and policy analysis.

The MNHA has been conducted a number of times over the years; in 1990, 1995, 1999, 2001, 2004, and every two years beginning in 2007.[1] This technical report focuses primarily on the 2017 MNHA, providing some cumulative data in table form.[2]

2. Sampling Methodology

---

[1] MNHA is funded through a legislative appropriation to the Minnesota Department of Health since 2007 and has received additional support from the Minnesota Department of Human Services since 2011.

[2] For information about earlier versions of the Minnesota Health Access Survey contact Kathleen Thiede Call at callx001@umn.edu and the Health Economics Program at health.mnha@state.mn.us.

The design of a sample is important to ensure that estimates derived from a survey are representative of the overall population and inferences are largely unbiased. The key objectives of the MNHA surveys are to update our understanding of Minnesota's uninsured population and to monitor key trends relevant to understanding factors that influence access to health insurance and health care services in Minnesota across populations and over time. The sampling strategy is designed specifically to generate reliable health insurance coverage estimates for the state overall, several geographic regions, populations of color and American Indians in Minnesota.

The MNHA surveys are telephone surveys. Although telephone surveys have lower response rates than in-person interviews, they are more economical. Mail surveys are more economical than both telephone and in-person interviews, and recent research indicates address based sample (ABS) have response rates that are on par with or higher than random digit dial (RDD) telephone surveys.[3] The complexity of the MNHA (where people with different insurance types are asked different sets of questions) is less conducive to self-administered/mail format. Web surveys hold promise for creating a self-administered version of the MNHA in the future. However, no web-based sample frames exist that can ensure appropriate representation of the Minnesota population. Moving to an ABS frame would allow for the adoption of a mixed-mode survey format. For example, the sample would be mailed an advance letter describing the study and offering the options of calling the survey center or going to a secure website to complete the survey. Sample elements with listed phone numbers could be telephoned if there was no response to the mail survey. However, address frames may not adequately capture Minnesota's most mobile and/or under-resourced populations. We continue to weigh these data collection trade-offs as costs increase and response rates decrease.

As with all telephone-based surveys, although the goal is to learn about people, the sampling units in the study are telephone numbers. The MNHA surveys are based on stratified RDD telephone samples that oversample in rural Minnesota and areas with greater probabilities of reaching populations of color and American Indians. Thus, the sample for the MNHA surveys consists of telephone numbers grouped into geographically contiguous areas or strata. Using data provided by SSRS[4], strata are created to resemble state and sub-state geography in the areas sampled as closely as possible. Each year, data are presented for the 13 economic development regions (EDRs) and are also available for some of the more populous counties (e.g., Hennepin and Ramsey). Within each geographic stratum, each telephone number has an equal probability of selection for the survey. As a way of obtaining sufficient sample sizes among populations of color and American Indians, telephone exchanges and rate centers in certain counties that the U.S. Census Bureau and other outside data indicate have higher rates of key populations are oversampled.

Since 2009 the MNHA has employed a *dual frame design* that includes both landline and cell phone RDD frames due to the persistent increase of people living in cell phone-only households and the desire to sample from all households with at least one type of phone. The range of telephone usage households includes landline only, cell phone-only and those with both landline and cell phone service (the latter are referred to as dual usage). According to Blumberg and Luke (2018), 53.9 percent of US households were cell phone-only in the last half of 2017. In this same period, 53.3 percent of US adults 18 and over lived in

[3]AAPOR Report: Address-Based Sampling. Prepared for AAPOR Council by the Task Force on Address-based Sampling, operating under the auspices of the AAPOR Standards Committee, January 7, 2016. Available at: https://www.aapor.org/Education-Resources/Reports/Address-based-Sampling.aspx#4.8%20Survey%20Designs%20for%20Single-%20and%20Mixed-mode%20Surveys

[4] GENESYS Sampling Systems, Marketing Systems Group (MSG)

cell phone-only households compared to 61.8 percent of US children under 18. In the Midwest 55.6 percent of adults live in cell phone-only households.[5]

The cell phone-only population (those who have a cell phone but no landline telephone) is systematically different. As compared to people in landline households, those in cell phone-only households are more likely to live in poverty or near poverty, to be age 25-34, Hispanic, male, renters, and live in metropolitan statistical areas. Compared to adults in households with a landline those in cell phone-only households report better health status and more active life styles, they are also more likely to be uninsured, report barriers to health care, and have higher rates of drinking and smoking.[6] Past research indicates post-stratification adjustments are effective for reducing bias in overall population estimates but may not do enough to reduce bias in estimates within subpopulations (e.g., populations of color and young adults).[7]

While cell phone-only cases are typically included in a sample to reach those who otherwise would be excluded from landline telephone studies, including persons who can be potentially reached through both phone frames is likely to increase the representativeness of the sample. There is evidence that those who live in dual telephone usage households and use their cell phone for all or most of their calls (termed "cell mostly") may be different than landline-only and cell phone-only populations.[8] However, as the prevalence of cell phone-only households rises, moving to a single-frame cell-phone only survey may become more viable.[9]

In the 2017 MNHA we continued to allocate a higher proportion of the total sample to the cell frame (73 percent) as compared to the landline frame (27 percent). Consistent with the 2015 sample design, a portion of the cell phone sample frame was drawn from numbers associated with a prepaid cell phone flag. Numbers with this prepaid cell phone flag were oversampled at twice the rate of their prevalence in the cell phone sample. This is based on evidence that numbers with the prepaid flag are more than twice as likely to be uninsured, more likely to be low income and from a population of color or American Indian community as compared to cell phone owners paying a monthly fee (Dutwin and Malarek, 2014).[10]

The decision to conduct interviews with all cell phone respondents adds to the complexity of weighting the data. Because base weights to adjust for the probability of selection are calculated separately for each survey frame, individuals with landlines *and* cell phones – the sizable overlap population – would be double-counted in the construction of weights. Adjustments to correct for this factor are discussed in Section 6, Weighting Survey Responses.

The goal with both landline and cell is to represent the state's population demographically and geographically. Landline sample frames are, by nature, geographically bound and linked to demographic

---

[5] Blumberg SJ, Luke JV. Wireless substitution: Early release of estimates from the National Health Interview Survey, July-December 2017. National Center for Health Statistics. June 2018. Available from: https://www.cdc.gov/nchs/data/nhis/earlyrelease/wireless201806.pdf. Accessed September 2018.

[6] Blumberg SJ, Luke JV. Wireless substitution: Early release of estimates from the National Health Interview Survey, July-December 2017. National Center for Health Statistics. June 2018. Available from: https://www.cdc.gov/nchs/data/nhis/earlyrelease/wireless201806.pdf. Accessed September 2018.

[7] Call KT, Davern M, Boudreaux M, Johnson PJ, Nelson J. Can post-stratification adjustments do enough to reduce bias in telephone surveys that do not sample cell phones? It depends. Medical Care, 2011; 49(4):355-64.

[8] Lee S, Brick JM, Brown ER, Grant D. Growing cell-phone population and noncoverage bias in traditional random digit dial telephone surveys. Health Services Research. 2010; 45(4):1121-39.

[9] Ganesh N, Khare M, Ormson EA, Zeng W, Jeyarajah J, Yankey D, Zhao Z Wolter KM. Noncoverage adjustments in a single-frame cell-phone survey: Weighting approach to adjust for phoneless and landline-only households. 2014, JSM Survey Research Methods Section. Available from: https://www.amstat.org. Accessed October 2018

[10] Dutwin D, Malarek D. The use of recent activity flags to improve cellular telephone efficiency. Survey Practice, 2014; 7(1):1-10.

data whereas cell sample frames are not. For the cell phone sample we used rate centers to allocate sample strata (a rate center is the area in which cell phones and numbers are purchased). The results of an analysis of 2013 MNHA data that indicated high concordance between respondent-reported residential zip code and the rate center location within sample strata (87 percent overall compared to 97 percent concordance between zip code and exchange in the landline frame).[11]

As in past years we took advantage of techniques used to remove inactive numbers from the cell phone sample (Cell-WINS screening) and the landline sample (scrubbing).[12] Details about the 2017 MNHA landline and cell phone samples are included in Appendix A.

Age screening in both sample frames is necessary given the disproportionately high participation in surveys among those 65 years and older. In 2017, we continued to use a list-assisted elderly screen for the landline frame, removing 75 percent of cases identified with this flag. In addition, we screened out 66 percent of all landline and cell households (including prepaid cell numbers) comprised of only adults age 65 or older, however, for the cell frame this was dropped to 50 percent of households after one month in the field.

Within each randomly selected household that agrees to participate, the person most knowledgeable about household members' health insurance is asked to complete the survey. After identifying that the household is eligible to participate in the study (e.g., Minnesota is primary place of residence), the household is enumerated and gender, age and relationship information is gathered for all household members. One person is then selected at random to be the target of the survey.

In order to produce reliable estimates for Minnesota children, when selecting the target from the household, children under age 18 within the household are given a 50 percent higher probability of selection than adults in the household.

3. Survey Content

Each year the majority of MNHA survey questions ask about health insurance coverage for the randomly selected individual within the sampled household – the target. This is followed by questions about health insurance coverage for all other household members, and education and employment information for all adults in the household. Information is also collected concerning potential sources of insurance (e.g., through the target's own or a family member's employer). Those lacking insurance are asked why they (or, in the case of a child target, their parents) did not purchase coverage.

In addition, the target's health status, access to health care and dental coverage is assessed, along with details about marital status (requested for primary caregiver or wage earner if the target is a minor), county of residence, race/ethnicity, nativity, citizenship and length of time living in the US. Finally, information about family income is requested along with questions relevant to weighting the data (e.g., number of phone lines, home ownership).

---

[11] Dutwin D. Stratification of cell phones: Implications for research. Presented at 66th Annual AAPOR Conference. Phoenix, AZ. May 2011.

[12] As is common practice in survey research, the landline sample was drawn from banks of telephone exchanges that contained at least three listed household phone numbers (versus numbers assigned non-residential households). This increases the efficiency of the sample – by increasing the likelihood of reaching an eligible household, study cost are reduced. Both the landline and cell frames are "scrubbed" or prescreened to remove inactive numbers; this results in cost savings from using interviewer time efficiently on active phone lines.

Some survey content changes each year the MNHA is conducted. This flexibility to alter questions to meet policy needs is a major advantage of the MNHA over other federal sources of data. Here we summarize changes to the 2017 survey. Detailed notes listing additions, minor revisions and omissions to the 2017 MNHA survey are documented in a "Final Revisions" spreadsheet and are available upon request.[13] A comprehensive record of change in survey content over time is available in the 2015 MNHA Technical Report.[14]

Revisions

Questions asking about the gender of household members were altered slightly. To adhere to best practices for capturing non-binary gender identification,[15] the "other" response option was changed to "something else." Consistent with 2015, interviewers can record a specific response if it is provided.

Additions/Substitutions

There is continued interest in experiences post ACA implementation, with a particular interest in monitoring the affordability of insurance and health care services.

Questions about Health Savings Accounts (HSAs) and Health Reimbursement Accounts (HRAs) were extended to those who reported having individual market (direct purchase) coverage (in the past this was only asked of those with employer-sponsored insurance). This is particularly relevant as enrollment in high deductible health plans has been growing over time.[16]

To further capture concerns about affordability of care we added a follow-up question for anyone indicating they had forgone care due to costs in an earlier series. This question asks for the reason they decided they could not afford that care. Responses include (1) knew the costs of care and could not afford, (2) unsure of the cost of care but afraid they could not afford it (3) it didn't matter how much it cost, they could not afford it, and a fourth option to provide their own response.

In response to stakeholder's request for greater understanding of financial burden caused by medical care we added two additional response options to an existing question asking about problems paying medical bills: (1) taking out a mortgage or loan due to medical bills and (2) declaring bankruptcy due to medical bills.

A paid sick leave ordinance for Minneapolis and St Paul went into effect July 1, 2017. We repeated the 2015 baseline question asking about access to paid sick leave. In 2017 this question was moved to the employment section of the survey and asked of all employed adults in the household. The ordinance is restricted to those who work at least 80 hours per year in Minneapolis or St. Paul. Therefore, we added a question asking if their primary place of employment (the physical location where they work) is located in Minneapolis or St Paul for those responding earlier that they reside in surrounding and commuter counties.[17]

---

[13] Contact Kathleen Call at callx001@umn.edu or Health Economics Program at health.mnha@state.mn.us.

[14] https://www.health.state.mn.us/data/economics/hasurvey/techreport2015.pdf

[15] Steiger D, L Heaton, J Behm, C MacAllum, J Stroop. Improving the measurement of sexual orientation and gender identity among youth. Presentation at AAPOR Annual Conference, New Orleans, LO, May 20, 2017.

[16] Miller GE, JP Vistnes, F Rohde, PS Keenan. High-deductible health plan enrollment increased from 2006-2016, employer-funded accounts grew in largest firms. Health Affairs 37(8):1231-1237, 2018; America's Health Insurance Plans. 2015 Census of Health Savings Account – High Deductible Health Plans https://www.ahip.org/wp-content/uploads/2015/11/HSA_Report.pdf, accessed December 2018.

[17] Hennepin, Ramsey, Dakota, Washington, Sherburne, Anoka, Chisago, Isanti, Wright, Mcleod, Carver, Scott, Goodhue, St. Louis, Carlson, Aitkin, Pine, Kanabec, Mille Lacs, Crow Wing, Stearns, Blue Earth.

Given documented burden of health care spending among those with chronic conditions[18], stakeholders requested the addition of a question to capture the presence of a health condition (i.e., physical, behavioral, mental or developmental conditions) that has lasted a year or more or a condition that is expected to last a year or more. This item replaced the third in a series of questions asking about the number of physical and mental unhealthy days. Specifically, we omitted the item asking how many days poor physical and mental health combined prevented the target from doing their usual activities (self-care, work, recreation).

There is growing concern that restrictive provider networks will limit access to care, although, results to date are mixed in terms of reductions in office visits among those in narrow networks.[19] To monitor this we added a fourth question to an existing provider supplies series that asks if the target was not able to get an appointment with the provider they wanted because the provider was not in their insurance network. This is followed by a question to distinguish the type of provider that was out of network (e.g., primary care, specialist, dentist, behavioral or mental health care provider).

Finally, we replaced one of the health insurance literacy (HIL) questions that was terminology knowledge-based[20] for another that asks participants about their confidence using their insurance, stated as a future rather than past action. For example, confidence in (1) finding a provider in their insurance network, and figuring out (2) if a service is covered, (3) how much a visit or service costs, and (4) which costs count toward the deductible.[21] Participants with insurance were randomly assigned to this new HIL question or another that asks about proactive use of insurance coverage that was included in the 2015 survey.[22]

Omissions

In addition to the replacements for unhealthy days and HIL questions described above, we removed a series of questions related to awareness of and use of the MNSure marketplace. Because the marketplace is fully implemented and no longer novel, these questions were no longer essential to informing policy.

Order Effects

In 2015, a question was added to the survey that asked targets how satisfied they were with the financial protection provided by their insurance. The location of the question in the survey was randomized between two positions in order to assess whether the placement of the question affected responses to that or subsequent questions. In this case, the question about financial protection offered by health insurance coverage was asked of all insured respondents either before or after the set of questions about health care utilization and access to health care.

---

[18] Paez KA, Zhao L, Hwang W. Rising out-of-pocket spending for chronic conditions: A ten-year trend. Health Affairs 28(1):15-25, 2009.

[19] Atwood A, AT Lo Sasso. The effect of narrow provider networks on health care use. Journal of Health Economics 50:86-98, 2016; Guber J, R McKnight. Controlling health care costs through limited network insurance plans: Evidence from Massachusetts state employees. American Economic Journal: Economic Policy 8(2):219-250, 2016.

[20] Kenney G, Karpman M, Long SK. Uninsured adults eligible for Medicaid and health insurance literacy. Health Reform Monitoring Survey. Washington, DC: The Urban Institute. 2014.

[21] Health Reform Monitoring Survey, 2016, Quarter 1 Survey. http://hrms.urban.org/survey-instrument/hrms-quarter-1-2016-survey.pdf/ Accessed October 2018.

[22] Paez KA, Mallery CJ, Noel H, et al. Development of the Health Insurance Literacy Measure (HILM): Conceptualizing and Measuring Consumer Ability to Choose and Use Private Health Insurance. J Health Commun. 19(sup2):225-239, 2014

After administration, we assessed whether the question performed differently in the two placements. In 2015, we found that people were significantly more likely to be very satisfied when they received the question before the set of other questions. When the two question positions were combined into a single variable for analysis, the combined estimate was not significantly different from either of the separate estimates. We then conducted a differential item functioning (DIF) analysis to detect item-level biases between the randomly assigned groups. The questions in this analysis included the satisfaction question, along with the questions related to forgone care due to cost and medical bills. The results indicated that the satisfaction question exhibited non-uniform DIF. That is, if people experienced a moderate amount of financial burden or forgone care, those who received the question first had a higher probability of being satisfied than people who received the question second. Question placement did not change how people responded if they did not experience financial burden and forgone care or if they experienced extensive financial burden and forgone care. Forgone dental care was the only other financial question that exhibited DIF, and it was negligible. Based on these results, the question remained in both positions in 2017.

We reran the analysis again after the 2017 administration and found that there were no significant differences for the overall population. Although the satisfaction question continued to exhibit non-uniform DIF, the forgone dental care question no longer exhibited any.

4. Survey Administration

The 2017 MNHA was conducted by SSRS, an independent survey research company based in Pennsylvania. The study received IRB approval from MDH and the University of Minnesota. As part of the survey protocol, respondents are read a consent script and are provided telephone numbers for the University of Minnesota Human Subjects office should they have concerns about the interview, and Dr. Call, should they have additional questions or concerns about the goals of the study and use of the data. Each year records of these calls from respondents are collected and the data are coded in order to monitor the frequency and nature of respondent inquiries (see Table 1).

As shown, contacts from respondents is a relatively rare occurrence given the large sample size of the MNHA each year. Similar to 2015, in 2017 many of the contacts (53 percent) were related to verifying the legitimacy of the survey after receiving calls from the survey firm, wanting to know more about the survey and use of the data, with most respondents feeling assured and happy to participate in the survey. These respondents are directed to web information about the MNHA. In 2017, a larger percentage of respondents called to share concerns about their health, the cost and adequacy of their insurance coverage, and opinions about federal and state health policy more generally. Each year, respondents requesting information to enroll in health insurance are provided phone numbers or directed to the appropriate website. Less than 10 percent voice complaints about the interviewer or request that we not call their phone number as they are not interested in participating; this was the modal category in 2013, which decreased in 2015 and remains stable. Finally, 11 percent call to compliment the content and structure of the survey or to voice concerns and suggest improvements.

Table 1. **MNHA Frequency and Content of Contacts from Respondents**

		2013		2015		2017	
Code	**Labels**	**Count**	**%**	**Count**	**%**	**Count**	**%**
1	Concerns about health care/health policy	3	7%	1	4%	14	25%
2	Legitimacy of survey, questions about survey, use of data, happy to participate	12	29%	19	70%	29	53%
3	Question about applying for insurance	4	10%	1	4%	1	2%
4	Complaints about interview/do not call	19	46%	3	11%	5	9%

5	Comments about survey (+/-)	3	7%	3	11%	6	11%
Total comments coded		**41**	**100%**	**27**	**100%**	**54**	**100%**
Unique contacts		*35*		*27*		*43*	
Initiated by IRB		*2*		*2*		*2*	
Email versus telephone contacts		*3*		*6*		*3*	

Each year MNHA data are collected through Computer Assisted Telephone Interviews (CATI) following a brief field test (see Table 2). The CATI system was programmed and thoroughly reviewed by all partners (SSRS, MDH, and SHADAC) prior to pretesting the instrument. The review consisted of multiple iterations of analyzing the accuracy of the skip pattern logic and interviewer directions for this complex instrument. A total of 41 pretests were conducted (14 landline; 27 cell phone) in early June 2017 followed by careful review of the recorded interviews and data. A key value of the pretest is to estimate the time to complete the interview with the goal of limiting the interview to 20 minutes on average. While the pretest resulted in only minor changes to the survey and CATI program, pretests were not included in the final sample (details are available by request[23]).

Interviewer training was conducted prior to the pretest and just before the study officially entered the field using SSRS's standard training methods. In addition, SHADAC provided a question by question ("Q by Q") manual that explains the motivation behind each question and provides responses to common or potential inquiries from respondents. SHADAC also provided a recording of the correct pronunciations of Minnesota counties to help interviewers correctly identify and record the respondent's county of residence. Interviewers were instructed to emphasize the social and policy relevance of the study and to reassure respondents that the information they provided was confidential. Calls were monitored over the course of the study (live by SSRS supervisors, and via de-identified audio recording by MDH and SHADAC staff), with intermittent interviewer training provided as needed. Best practices were used to achieve the highest possible response rates.[24] Detailed data collection procedures and results authored by SSRS are available by request.[25]

Table 2. MNHA Field Period

Survey Year	Dates
2001	November 2000 - May 2001
2004, 2007	July - December
2011	September - December
2009, 2013, 2015	August - November
2017	June - October

[23] Contact Kathleen Call at callx001@umn.edu or Health Economics Program at health.mnha@state.mn.us.
[24] Best practices included a target of 8 maximum call attempts, calling at different times of day and days of week, setting callback appointments, resting sample, and using specially trained interviewers for refusal conversion.
[25] Contact Kathleen Call at callx001@umn.edu.

The actual time it takes to conduct an interview varies by household size, insurance status, telephone status, and survey language. The average length of time it takes to complete the MNHA interview has been relatively consistent over the years (see Table 3). Variation in interview length between cell and landline samples has dropped over time. Due to the complexity of translating health insurance and access terms, the surveys completed in Spanish require more time on average than English language interviews.

Table 3. MNHA Length of Survey in Minutes

Survey Year	Landline	Cell	Landline and Cell	
			Spanish*	English
2013	18	22	29	20
2015	19	23	32	22
2017	22	22	32	22

*A total of 97, 142 and 123 interviews were conducted in Spanish in 2013, 2015 and 2017 respectively. In 2001, surveys were also completed in Hmong (32) and in 2004 surveys were completed in Hmong (85) and Somali (38). Due to the high cost of translations and the low number of surveys completed in Hmong and Somali this was discontinued in 2007.

Remuneration

MNHA respondents in the cell phone sample are offered five dollars remuneration for completing the interview. This is particularly important due to the decision to oversample prepaid cell phone numbers in the cell phone frame. Contact information is acquired at the end of the interview and stored in a database separate from survey responses. Although compensation is offered, not all respondents provide the contact information necessary to receive compensation, and take-up of this offer dropped between 2015 and 2017, especially for the non-prepaid sample (Table 4).

Table 4. Percent of MNHA Cell Phone Sample Accepting Remuneration

Cell phone sample:	2017	2015	Difference
Prepaid	59%	62%	-3
Non-prepaid	46%	54%	-7
Total	48%	55%	-7

5. Response Rates and Sample Coverage

Over time response rates have dropped for all types of surveys, with some evidence of having plateaued in recent years.[26] For telephone based surveys this general trend is attributable to growth in the non-contact rate (e.g., fewer people answering their phone as a result of telephone screening devices) and small growth in refusal rates (e.g., households/individuals declining to participate in a survey perhaps due to frustration with marketing, bot calls, and survey research in general).[27]

[26] Pew Research Center. May, 2017. What low response rates mean for telephone surveys: Telephone polls still provide accurate data on a range of social, demographic and political variables, but some weaknesses persist. Available at: http://www.pewresearch.org/methods/2017/05/15/what-low-response-rates-mean-for-telephone-surveys/. Accessed October 2018.

[27] Lavrakas PJ et al., The future of U.S. general population telephone survey research. Report from the AAPOR task force, 2017. Available at: http://www.aapor.org/getattachment/Education-Resources/Reports/Future-of-Telephone-Survey-Research-Report.pdf.aspx. Accessed June 2017.

Falling response rates and the implications for data quality is the subject of intense attention and scrutiny.[28] Response rates are a commonly used indicator of the quality of a survey. However, research indicates that lower response rates are not necessarily associated with greater response bias because surveys with high and low response rates demonstrate similar levels of absolute bias.[29,30,31,32]

In general terms, the response rate is the ratio of the number of completed interviews divided by the number of eligible reporting units in a sample; the cooperation rate is the ratio of all interviewed cases to all eligible cases contacted. The response rates reported below refer to AAPOR Response Rate #3[33] from 2009 to present for the blended sample (cell and landline combined), which is the equivalent of the number of completed interviews divided by the total number of eligible phone numbers.[34] As presented in Table 5 below, consistent with other surveys, the MNHA response and cooperation rates have somewhat diminished over time; by contrast refusal rates have fluctuated over time.

---

[28] Czajka JL, Beyler A. Declining response rates in federal surveys: Trends and Implications. Mathematica Policy Research, June 15, 2016. Available at: https://aspe.hhs.gov/system/files/pdf/255531/Decliningresponserates.pdf. Accessed June 2017; Lavrakas PJ et al., The future of U.S. general population telephone survey research. Report from the AAPOR task force, 2017. Available at: http://www.aapor.org/getattachment/Education-Resources/Reports/Future-of-Telephone-Survey-Research-Report.pdf.aspx. Accessed June 2017.

[29] Groves R. Nonresponse rates and nonresponse bias in household surveys. Public Opinion Quarterly, 70(5): 646-675, 2006; Groves R, Peytcheva E. The impact of nonresponse rates on nonresponse bias: A meta-analysis. Public Opinion Quarterly, 72(2): 167-189, 2008.

[30] Keeter S, Kennedy C, Dimock M, Best J, Craighill P. Gauging the impact of growing nonresponse on estimates from a national RDD telephone survey. Public Opinion Quarterly, 70(5): 759-799, 2006.

[31] Davern M, McAlpine DD, Beebe TJ, Ziegenfuss J, Rockwood T, Call KT. Are lower response rates hazardous to your health survey? An analysis of three state telephone health surveys. Health Services Research 45(5):1324-1344, 2010.

[32] Groves R, Peytcheva E. The impact of nonresponse rates on nonresponse bias: A meta-analysis. Public Opinion Quarterly, 72(2): 167-189, 2008.

[33] The American Association for Public Opinion Research. 2011. Standard definitions: Final dispositions of case codes and outcome rates for surveys. 7th edition. Lenexa, Kansas: AAPOR. Available at: http://www.aapor.org/Content/aapor/AdvocacyandInitiatives/StandardsandEthics/StandardDefinitions/StandardDefinitions2011.pdf

[34] To estimate the number of eligible phone numbers among numbers with *unknown* eligibility (e.g., no answer), this rate applies the ratio of eligible to ineligible numbers among the numbers with *known* eligibility to the *unknown* numbers and includes the resultant number within the denominator of the response rate calculation.

Table 5. MNHA Count of Complete Surveys, Response, Cooperation and Refusal Rates				
Survey Year	Total Completes^	Response Rate*	Cooperation Rate*	Refusal Rate**
2001	27,315	67%	78%	19%
2004	13,802	59%	68%	28%
2007	9,728	43%	57%	32%
2009	12,031	45%	53%	39%
2011	11,355	44%	45%	39%
2013	11,778	48%	48%	23%
2015	11,178	35%	36%	30%
2017	12,042	29%	30%	32%

MNHA 2001-2007 represent landline sample frames; MNHA 2009 forward represent dual landline and cell phone sample frames.
^ The total count includes partial interviews. Cases were designated as partial completes if the survey was completed through the health insurance coverage (H series) (2001-2015), roster coverage and demographics, and access to coverage (COV) series where applicable in 2015/17.
* Based on AAPOR RR4 response and cooperation rates from 2001-2007; Based on AAPOR RR3 response and cooperation rates from 2009 forward which excludes partials.
** Based on AAPOR refusal rate 2 (REF2); includes estimates of eligible cases among unknown cases. For comparability with prior MNHA surveys, refusal rate calculations from 2009 forward ignored screening that occurred (e.g., excluding minors both years and over sampling of cell only households). In 2017, the refusal rate was 36.4 percent and 31.1 percent for landline and cell frames respectively.

Table 6 presents the counts of completed surveys and dispositions separately for the landline and cell phone frames from 2009 forward, when cell phone sampling was introduced. As shown, response and cooperation rates are consistently higher in the landline frame compared to the cell phone frame. As the cell phone frame becomes a larger portion of the total sample, the response and cooperation rates for the total sample are blended and skew toward the cell phone rates.

Table 6. MNHA Count of Complete Landline (LL) and Cell Phone Surveys, Response, and Cooperation Rates									
Survey Year	Interview counts			Response Rate*			Cooperation Rate*		
	Total^	LL	Cell	Total	LL	Cell	Total	LL	Cell
2009	12,031	9,811	2,220	45%	50%	31%	53%	58%	40%
2011	11,355	7,028	4,327	44%	48%	39%	45%	49%	40%
2013	11,778	4,952	6,421	48%	50%	40%	48%	51%	41%
2015	11,178	3,139	8,039	35%	41%	29%	36%	36%	30%
2017	12,042	3,291	8,751	29%	34%	25%	30%	35%	26%

^ The total count includes partial interviews. Cases were designated as partial completes if the survey was completed through the health insurance coverage (H series) (2001-2015), roster coverage and demographics, and access to coverage series where applicable in 2015/17.
LL = Landline
* Based on AAPOR RR3 response and cooperation rates which excludes partials.

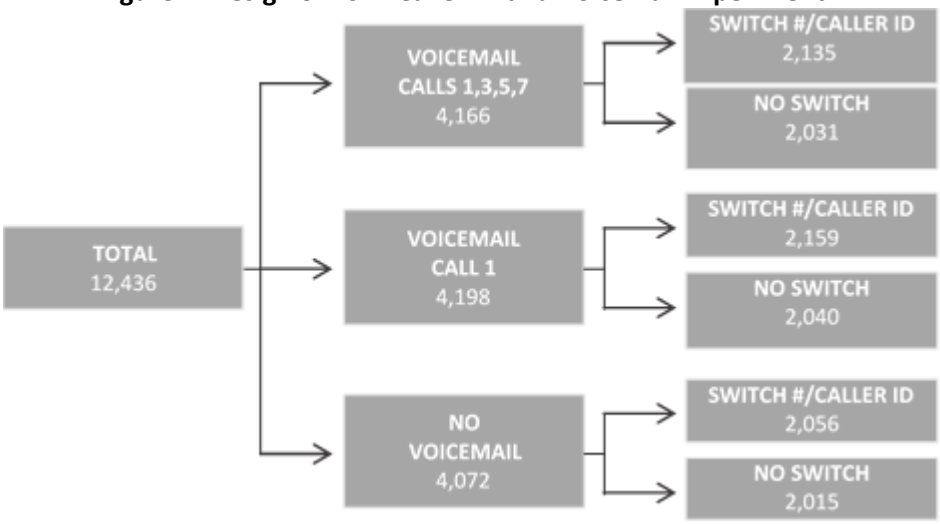
2017 Caller ID and Voicemail Experiment

In the 2017 MNHA we conducted an experiment to explore techniques for increasing the response and contact rates and reducing busy and no answer dispositions. The experiment varied two treatments:

1. Frequencies of voicemails: whether a voicemail was left on (a) the first call only, (b) the first, third, fifth and seventh call, or (c) no voicemail was left.
2. Caller ID display: whether the caller ID display (a) remained the same throughout the call protocol or (b) was switched halfway through the protocol (the fourth attempt) from the standard phone number and caller ID -- 612-626-3261 and "UNIV OF MINN" -- to a new phone number and caller ID -- 651-201-3846 and "MN DEPT HEALTH".

The landline and cell sample frames were evenly and randomly assigned one of the six conditions representing the combination of the two treatments (see Figure 1 below).

Figure 1. Design of 2017 Caller ID and Voicemail Experiment



Results in Table 7 show that both conditions independently resulted in higher response rates; more frequent voicemails and switching the caller ID display after the fourth call resulted in higher response rates. Further, we found an interaction effect between both treatments such that leaving more frequent voicemails and switching caller ID yielded higher response rates. This joint effect was amplified for the landline sample where there was a 3.3 percentage point spread in response rates between the No Voicemail/No switch condition (32.4 percent) and the Frequent voicemail/Switch condition (35.7 percent) compared to a 1.6 percentage point response rate spread for the cell sample (24.7 percent for the No voicemail/No switch condition and 26.3 percent for the Frequent Voicemail/Switch condition).

Table 7. Response rates by experimental treatment

	Caller ID Switch	
Voicemail left on	**No Switch**	**Switch after 4 calls**
Calls 1,3,5,7	28.4%	30.2%
Call 1	28.2%	29.8%
No voicemail	27.9%	28.9%

In addition, there were more busy/no answer dispositions for the No switch than the Switch condition (data not shown)[35], particularly in the landline frame where these rates are higher (54 percent) than the cell frame (16 percent). We also looked at the contact rate – the number of sample pieces called before a respondent is reached. Here again we see an interaction between conditions: the Frequent Voicemail/Switch condition was significantly more efficient (24.9 percent contract rate) than the No Voicemail/No switch condition (24.3 percent). While this difference appears small, it translates to about 3,600 fewer contacts in a sample frame of about 274,000. We found no treatment effects on the following sample dispositions: answering machine, refusal/callback and ever-refused. Finally, cost comparisons across the six experimental conditions were inconclusive.

6. Weighting of Survey Responses

The goal of weighting survey data is to adjust the results to account for sample coverage problems (the difference between respondents and non-respondents) and reduce potential bias associated with differential participation in the survey. Accounting for varying probabilities of selection and response rates through the application of weights enables the survey responses drawn from statistical samples to be representative of the entire population.

Two types of weights were generated: 1) base weights and 2) post-stratification weights. The base weight takes into consideration that each respondent's probability of selection varies by sampling stratum, the number of phone lines connected to the household (or number of cell phones accessible to adults in the case of the cell phone frame), and the number of people living in the household. The post-stratification weights adjust the base weight to account for key characteristics of the state's population. Specifically, to more accurately reflect the population, sample weights were post-stratified by region, age, education, race, nativity (US versus foreign born), home ownership, household count, and telephone usage.

In addition, because the cell phone sample frame was not limited to cell phone-only respondents, the weights account for the probability of including individuals in the sample who live in dual landline and cell phone households (referred to as overlapping sample frames).

Base Weights

Landline samples are associated with households and do not select individuals per se. This approach randomly draws landline telephone numbers associated with households within desired geographic areas (or rate centers in the case of the cell phone frame). By contrast, cell phone numbers are associated with individuals. The landline and cell RDD samples used in the MNHA were drawn from a sampling frame of Minnesota phone numbers in active area code/exchange groupings and rate centers within geographic strata.

The first component of the base weight accounts for a person's known probability of selection based on the chosen geographic strata. This is necessary because some areas of the state were oversampled relative to others. The strata adjustment is calculated by dividing the total number of telephone numbers available in each region (regardless of whether or not they are in the sample) by the total number of interviews completed in that region. This indicates how many telephone numbers are represented by each telephone number that resulted in a complete. The strata weight component also accounts for differential response rates by strata.

[35] Call KT, Hagge SL, Simon AB, Alarcón G, Turner K, Dutwin D. Won't you please pick up? How do voicemails and call number impact the likelihood of survey response? Presented at the Annual AAPOR Conference, May 17, 2018, Denver CO. Available upon request from Kathleen Call at callx001@umn.edu.

A second component of the base weight accounts for the number of people in the household. People in larger households have a smaller probability of being included than people in smaller households. Therefore, on average people in larger households receive larger base weights, correcting for their lower probability of selection. The second base weight component also illustrates that the purpose of weighting the MNHA is to develop person-level weights. That is, translating the response from randomly selected individuals in households into representative responses about Minnesotans.

Third, we adjust for the number of telephones in the household, as persons in households with more telephone lines (or cell phones in the cell sample) have a greater probability of being selected into the sample. For example, households with two telephones are twice as likely to be randomly selected as are single-telephone households; a weight of one-half appropriately adjusts for the two telephone household's greater probability of selection. In the case of households with cell phones, we account for the number of cell phones that could be answered by an adult in the household as we do not directly interview minors or cell phone numbers assigned to minors.[36] Below (see Dual Frame Weights) we describe adjustments made to account for the possibility that members of landline sample could also be captured in the cell phone sample.

A fourth adjustment accounts for the fact that even after screening respondents over age 65, they still often make up a greater portion of the survey respondents than the population overall. This adjustment is applied in a manner similar to a post-stratification adjustment. Using American Community Survey (ACS) estimates of the proportion of households comprised of members age 65 years and older (65+), we bring the number of 65+ sampled households into alignment with the population.

Additionally, we increased the probability of selecting a child as the target in households with children that required that we adjust these cases down in the base weight.

Next, we must adjust for the 65+ listed sample (see section 2, sampling methodology). This adjustment was calculated by dividing the percentage of 65+ listed sample (25 percent) by the percentage of 65+ listed completes within the sample (4 percent), resulting in an adjustment of 6.5. We then adjust by multiplying the adjustment factor of 6.5 by the previous weight for those cases with the 65+ flag.

Finally, we must adjust for the oversample of prepaid cell phones in the cell phone frame. This adjustment was calculated by dividing the percentage of prepaid sample (15 percent) by the percentage of prepaid completes within the cell sample (13 percent), resulting in an adjustment of 1.11. The adjustment factor of 1.11 was then multiplied by the 65+ weight for those cases with the prepaid flag. A similar calculation was made for the non-prepaid cell phone cases.

Dual Frame Weights

Dual frame sampling requires an additional step in the weighting process that accounts for the overlap in both frames (i.e., the portion of households that could be theoretically reached in either the landline or cell phone frame). We adhered to the dual frame weighting methodology developed and evaluated under contract with National Opinion Research Center (NORC) in 2009 (For details, see the 2009 MNHA Technical Report[37]).

---

[36] It is important to note that although the steps of the base weight calculation were the same in the landline and cell phone frame, the calculation was operationalized in separate analyses to ensure that the probability of selection was calculated specifically for each frame. In other words, the number of landline phones was considered for people who were reached in the landline frame (not the cell phones that respondents might have had access to). Similarly, in the cell frame no landlines were considered in calculating the probability of selection associated with access to working cell phones.

[37] Available by request from Kathleen Call at callx001@umn.edu or Health Economics Program at health.mnha@state.mn.us.

The common strategy of accounting for this overlap is to multiply the weights for the landline RDD interviews by a weighting adjustment factor, or λ (lambda), and multiply the weights for the cell phone interviews by 1- λ. Although some cases still have a chance of being included in either sample frame, the weights are adjusted so they are not overrepresented. Unfortunately the actual amount of overlap is not available for the relevant geographic area and time frame (i.e., Minnesota in 2017). Consistent with past years, the telephone usage status from the National Health Interview Survey (NHIS) for the Midwest Census region was used to set the value of λ for cases with different patterns of telephone use in 2017.

Table 8 demonstrates the screening we incorporated in 2011 and 2013 resulted in a larger portion of cell phone-only completes in the cell frame (23 percent and 30 percent) as compared to 2009 (7 percent), yet this was still below estimates for the Midwest of 41 percent in 2012. In 2015 the cell-phone only estimate increased to 43 percent, which is much closer to the 2014 Midwest estimate of 48 percent. Despite the weighting adjustment, the cell phone completes, generally, had larger base weights than the landline completes simply because the universe of available cell phones per sample strata is larger than is true for landline strata.

Table 8. MNHA Telephone Usage (Unweighted)

	2009	2011	2013	2015	2017
Landline Only	16%	14%	8%	4%	3%
Cell Phone-Only	7%	23%	30%	43%	47%
Cell Phone-Mostly	14%	15%	20%	19%	21%
Landline Mostly	63%	49%	42%	35%	29%
Total	100%	100%	100%	100%	100%

Post-stratification Weights

While the base weights adjust for the known unequal probability of selection, post-stratification weights adjust for ways in which the sample's demographics and the resulting completed interviews differ from what is known about the population from which the sample was drawn. For example, if 20 percent of survey respondents were 65 years of age or older (with the base weights applied) yet the census data indicate that only 12 percent of the general population was elderly, a post-stratification weight adjusts the base weight so that it represents the actual age mix in the population. This ensures that the resulting estimates more appropriately reflect the true characteristics of the population. The term *post-stratification* refers to the fact that the adjustment is conducted *after* the data are collected, and the sample is *stratified* by demographic characteristics to match the independent estimate.

Post-stratification of the MNHA surveys rely on the most current data available from the US Census Bureau's American Community Survey (ACS); typically from the year just prior to the MNHA data collection year.[38] In 2017 (consistent with past MNHA survey) several demographic characteristics were used to post-stratify the MNHA data as well as an indicator of the type of phone usage (e.g., landline only, cell phone only, etc). The first five factors represent conventional demographic adjustments; factors 6 through 9 represent less conventional adjustments informed by research on characteristics of the cell phone-only population and geared toward compensating for their omission from the sample frames in the MNHA surveys prior to 2009:

[38] Earlier MNHA surveys used the demographics from the decennial Census and estimates for the years in-between the decennial surveys.

1) Geographic region[39]
2) Race/ethnicity
3) US born or foreign born status
4) Education
5) Age
6) Age by education[40]
7) Home ownership status[41]
8) Household count
9) Phone type

In terms of the overall weighting process for the 2017 MNHA, after the base weights, the adjustment for 65+ households, the adjustment for child selection and the adjustment for prepaid cell phone were computed, and the selected λ weight adjustment was applied to the base weight. Within the process of applying the selected λ, the weights were also post-stratified using person-level telephone usage estimates (landline only, cell only, cell mostly and landline mostly). The control totals for telephone usage are derived from the NHIS for the Midwest Census region (See Table 9).[42]

Table 9. NHIS Midwest Telephone Usage Estimates Used for Weights

	2008	2009	2010	2011	2012	2014	2016
Landline Only	16%	12%	9%	7%	6%	9%	6%
Cell Phone-Only	20%	25%	30%	36%	41%	48%	53%
Cell Phone-Mostly	14%	16%	18%	19%	17%	14%	14%
Landline Mostly	50%	48%	43%	39%	37%	29%	27%
	100%	100%	100%	100%	100%	100%	100%

The post-stratification is implemented by applying a raking algorithm.[43] The goal of a raking algorithm is to lower standard errors and provide a more efficient weighting structure quickly.[44] The raking algorithm employs an iterative process that uses the base weight as the starting weight, applying each post-stratification factor one after the other, reapplying factors, and ending when a specified convergence criterion is reached. Convergence requires that each marginal total of the raked weights is within a specified tolerance level of the corresponding population control totals.[45]

For the 2017 MNHA data, the raking algorithm was set at a convergence level of .001; the data converged within 10 iterations. Trimming the weights was not necessary in 2017 as we did not have extreme outliers (this was only encountered in the 2013 weights).

---

[39] Regional analyses of the data are based on Minnesota's 13 Economic Development Regions (EDRs). For geographic weights we use a variable created by the Census Bureau called Super-PUMAs, or Super-Public Use Microdata Areas. The use of Super-PUMAs expedited the calculation of population counts necessary for weighting. EDRs and Super-PUMAs do not match up exactly by county. See "Geographic Adjustment" section below for a discussion of mismatch between (1) the geographic allocation based on the Genesys sample frame compared to (2) respondent's report of household location.
[40] This weight is a proxy for adjusting for sample coverage challenges among young adults of varying education levels.
[41] Call KT, Davern M, Boudreaux M, Johnson PJ, Nelson J. 2011. Bias in telephone surveys that do not sample cell phones: uses and limits of post-stratification methods. *Medical Care,* 49(4):355-364
[42] Special National Health Interview Survey runs for Midwest performed by JV Luke at the National Center for Health Statistics.
[43] The SAS macro RAKINGE developed by Izrael D., Abt Associates was used in the process. The adjustment factors were entered as follows: PUMA, age, race, USborn, education, age*education, home ownership, household count, phone type.
[44] The design effect is the factor by which the variance of estimates was increased due to weighting.
[45] Izrael D, Hoaglin DC, and Battaglia MP. (2004). To Rake or Not To Rake Is Not the Question Anymore with the Enhanced Raking Macro. *Proceedings of the Twenty-Ninth Annual SAS Users Group International Conference*, Cary, NC: SAS Institute Inc., Paper 207. Available from: http://www2.sas.com/proceedings/sugi29/207-29.pdf. August 25, 2010

Data Editing and Key Variable Construction

SSRS monitors data quality throughout the survey field period. As part of ongoing data quality checks, on July 24, 2017 SSRS discovered an error in the CATI programming logic that resulted in a subset of respondents not being asked the education (EDUC) or student (STUD) questions. This error affected 288 respondents at EDUC and 110 at STUD. After alerting MDH and SHADAC of this error, it was decided that any targets or parents/guardians of targets affected by this missing information would be called back to answer EDUC (a critical weighting target). SSRS utilized a separate callback program to contact the 127 respondents missing this information, and 103 were successfully re-contacted. The remaining 24 respondents who could not be reached, along with all other missing cases at EDUC and STUD, were left as blank for those questions.

SSRS created an analytical data file with all data collected during the survey field period. Data were checked using multiple methods including: (1) A "data cleaning" procedure in which data processors recreated the process of CATI variable creation (derived from skip patterns, definitions of codes and ranges specified in the questionnaire) to ensure that all variables were created correctly and had appropriate numbers of cases, and (2) the project director independently checked off all SPSS variables to confirm they were created correctly, had the correct number of cases, and were coded according to specifications.

Additional checks were performed on the composition of households. In general, household data remained as reported by the respondent. Cases with illogical household relationships were flagged for review by the research directors. If there was a clear and logical way to correct seemingly illogical household relationships, a change would be made to the data (example: A two-person household where the child is age 40 and the parent is 6 years of age).

MDH and SHADAC performed other logical edits and cleaning functions in the process of creating analytic variables. For example, if individuals reported carrying health insurance through the Indian Health Service (IHS) and no other coverage, they are coded as uninsured, because IHS is typically not considered insurance coverage. Further, logical conflicts potentially created during the imputation process were corrected.

For all variables that included response options allowing text-based entry (i.e., "Other, specify") by the interviewer (e.g., race, ethnicity, industry), respondent's answers were reviewed and data was back coded to available response options, new categories were created if appropriate, or responses were left as "other."

Income Imputations

Consistent with other surveys, income has the highest item nonresponse (i.e., respondents choose not to answer the question) of any of the survey items. Income related measures are important to the MNHA because of their association with various dimensions of health and our interest in estimating the proportion of the population that is uninsured but appears to be eligible for public health insurance or Advanced Premium Tax Credits (APTC) in the individual insurance market.

As presented in Table 10, each year a majority of cases answer the open-ended question about household income or respond to a follow-up question providing a set of income ranges for those unwilling to state their income in the first question. In 2017 about 13 percent of respondents did not respond to the income items which is consistent with the past few MNHA surveys. Excluding these cases could introduce bias to our survey estimates (non-responders may share certain income characteristics), therefore income was imputed for these respondents. A second advantage of imputation is that it

allows all respondents to be included in calculations involving income, such as uninsurance rates by poverty level and eligibility for public programs among the uninsured.

Table 10. Percentage Responding to Income Questions and Missing Data			
Survey Year	**Open-end Income**	**Income Range**	**Missing Data**
2011	77%	14%	9%
2013	77%	11%	12%
2015	76%	11%	13%
2017	77%	10%	13%

Income was imputed using a statistical procedure known as hotdeck and designed for Stata.[46] The hotdeck procedure searches for cases with complete income data (donors) based on whether they are demographically similar to cases with missing data (recipients); a donor is selected randomly from the possible set of donors.[47] Demographic variables used in this imputation include gender, age, race/ethnicity, insurance type, household size, geographic region, telephone interruption, educational achievement of target (or primary wage earner if target is a child) and use of government financial assistance programs, such as WIC, among those responding only to the categorical income question.

Age Imputations

Respondents who were not comfortable providing age data were asked a categorical age question, allowing the target to be identified as a member in one of four possible age groups: a 0-17 year old child, an 18-25 year old young adult, a 26-64 year old adult, or an adult 65 years or older. For the 1.4 percent of cases that refused the initial age question in 2017 (consistent with 2015 refusal rate), age was imputed using the categorical age question, sex, marriage status, and household relationships – specifically, if the target was listed as a parent or a child.

Geographic Assignment

Each year, respondent geography was provided by GENESYS Sampling Systems, Marketing Systems Group (MSG) in the form of county FIPS codes. Respondents were also asked to provide their county and zip code in the survey. For those cases in which the GENESYS FIPS did not match the respondent provided county or zip code, the respondent provided data were used.[48]

Computing the Primary Source of Health Insurance Coverage

MNHA respondents report all sources of health insurance coverage available to them, accurately reflecting the scenarios with primary and secondary coverage. For a range of analyses and publications, we calculate Minnesotans' primary source of health insurance coverage, meaning individuals reporting insurance are assigned only one type of coverage. The following hierarchy is used for determining the *primary* source of coverage for people who report access to multiple sources:

1. Public: Includes all state and federal public coverage and military.

[46] The software module was designed by Adrian Mander and David Clayton at the MRC Biostatics Unit of the Institute of Public Health in the University of Cambridge, UK.

[47] A hotdeck procedure was used for imputing other missing information needed for the income imputation: gender, age, homeownership, education, employment, race, country of birth, length living in the US, and phone status.

[48] For the Landline sample frame: If respondent zip code and county matched but differed from Genesys, we used respondent provided county; if respondent zip code and county did not match, we used the variable that matched Genesys; if respondent county, zip code and Genesys county did not match, we used respondent county. For the cell phone sample, respondent provided county was used.

2. Employer: Includes employer-sponsored coverage for employees and their dependents.
3. Individual: Includes all direct purchased coverage for individuals and families.
4. Uninsured: Includes those without any coverage and those who only have sources such as Indian Health Service that is not considered comprehensive health insurance coverage.

The order of the hierarchy is based on researchers' understanding of which coverage likely acts as the primary payer of health care services. For example, if an individual reports Medicare coverage and retiree coverage through an employer, then public coverage was assigned as the primary source of coverage. Beginning in 2015, we use additional questions to assign coverage. Specifically, participants who answered yes to MNsure and no to paying a monthly premium were coded as having public coverage. Participants who answered yes to MNsure and yes to paying a monthly premium were coded as having individual coverage. The different types of public coverage are not separated out in the hierarchy because respondents often experience difficulties in differentiating among the different state and federal programs.[49]

Calculation of Public Program Eligibility and Access to Employer Coverage

Questions on the MNHA related to income, household composition, age, and access to employer coverage were used to determine whether the currently uninsured were potentially eligible for public health insurance programs. Eligibility for the public programs are based on factors such as income, assets, household size, household composition, age of household members, pregnancy status, disability status, length of residence in Minnesota, immigration status, access to employer coverage, and level of employer contribution. Because the survey does not ask questions related to respondents' assets, level of employer contribution, pregnancy, disability, or immigration status, those factors could not be considered in the process of determining potential public program eligibility.

The distribution of potential access to insurance for the point-in-time or currently insured can add to more than 100 percent because some people can have access to employer coverage and still be eligible for public health insurance programs in Minnesota. Respondents with incomes low enough to qualify for Medical Assistance (MA) or MinnesotaCare who also reported having employer coverage were coded as being potentially eligible for public programs and having access to employer coverage.

Measuring Race, Ethnicity, and Country of Origin

The MNHA survey contains a series of questions that are used to identify a respondent's race and ethnicity. Collection and aggregation of this data has changed slightly over time to maintain consistency with guidelines established by the U.S. Office of Management and Budget.[50] To determine ethnicity, each respondent is first asked, *"Is the target person Mexican, Puerto Rican, Cuban, or from another Hispanic or Latino group?"* This is followed by a question about race, asking the respondent to choose one or more races that they consider the target person to be. With the exception of the weighting process, which required a mutually exclusive race variable, race and ethnic groups are generally defined using the Census Bureau's "any race" construct.[51] An individual is considered to belong to a specific racial or ethnic group if they report that race or ethnicity either alone or in combination with another race or ethnicity. Individuals reporting more than one race or ethnic identity are included in all of these

[49] Call KT, Davern ME, Klerman JA, Lynch V. 2012. Comparing errors in Medicaid reporting across surveys: Evidence to date. *Health Services Research,* Apr;48(2 Pt 1):652-64.

[50] Office of Management and Budget, 2003, *Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity.* Available at: http://minorityhealth.hhs.gov/templates/browse.aspx?lvlID=172

[51] Source: U.S. Census Bureau, 2003. *U.S. Census Bureau Guidance on the Presentation and Comparison of Race and Hispanic Origin Data.* Available at: http://www.census.gov/population/www/socdemo/compraceho.html

reported groups.[52] For this reason, counts obtained from the "any race" construct will total more than the population total for the state and percentages will sum to more than 100 percent.

Country of origin, length of time in the United States (for non-U.S. born individuals) is collected for all targets. Beginning in 2015, citizenship was also collected for non-U.S. born targets. This information was also gathered for the parents of targets under 18. Due to changes over time in the collection of length of time in US, for consistency, these data are only reported for individuals age 3 and older.

Measuring Employment

The structure of the employment questions has been consistent since 2011. Information about employment status, employment at more than one job, and total hours worked per week at all jobs is collected for all adults in the household. In addition, a separate student status question is asked of all adults under 65. This design allows respondents to more appropriately identify themselves as employed, as well as students, when that is the case.

7. Data Analysis

As in earlier years, because of the survey's complex sampling design, data were analyzed using Stata statistical software.[53] This software was chosen for its ability to obtain unbiased estimates of standard errors and confidence intervals in the face of the MNHA's multistage sampling. The analysis specifies survey weights and strata, and, in the case when information on members of given households are used, primary sampling units (PSUs). Differences between groups and changes over time are considered statistically significant when the p-value is less than 0.05. Generally, comparisons of estimates are made between years or within year with the population total serving as the reference group.

8. Availability of Research Findings

Research results from the MNHA are made available in multiple formats including:

- Short issue briefs on a variety of topics, including an overview of key results;
- Presentation slides; and
- An interactive data reporting system that allows users to query survey results with great flexibility. [54]

[52] In 2017, 4.6 percent of target persons were reported to have more than one race. In 2015, 2.3 percent of target persons were reported to have more than one race as compared to 1.8 and 1.5 percent in 2009 and 2007 respectively.

[53] StataCorp. 2015. *Stata Statistical Software: Release 14.* College Station, TX: StatCorp LP; StataCorp. 2013. *Stata Statistical Software: Release 13.* College Station, TX: StatCorp LP.

[54] Issue briefs are available online at the Health Economics Program's (HEP) home page: https://www.health.state.mn.us/healtheconomics The presentation slides can be obtained from the Health Economics Program's Chartbook series: https://www.health.state.mn.us/data/economics/chartbook/index.html. The data reporting system can be accessed at: https://mnha.web.health.state.mn.us/Welcome.action.

Appendix A: Sample Frames

Table A-1. Landline Sample Design, 2017

Strata	Region	EDR	All Households	Sample Ratio	Target Completes	African American	Native American/ American Indian	Asian	Hispanic /Latino	Residual (prim. Whites)
1	Northwest	1	25,344	0.43%	108	1	3	1	4	99
2	Headwaters	2	21,966	0.37%	81	0	4	1	2	74
3	Headwaters (Beltrami, Becker, Mahnomen, Cass)	2, 4, 5	9,835	3.40%	334	1	115	1	9	208
4	Arrowhead	3	86,548	0.17%	147	2	5	1	2	137
5	West Central, East Central, North Central	4, 5, 8	150,946	0.24%	369	3	5	3	9	350
7	Mid-Minnesota, Southwest	6, 10	63,165	0.35%	224	4	2	4	15	201
8	Upper MN Valley	7	13,497	1.10%	148	1	2	1	6	139
9	Central	9	92,136	0.17%	157	4	1	2	4	146
10	South Central	11	51,433	0.21%	107	2	0	1	5	98
11	Southeast	12	114,152	0.18%	209	0	1	6	10	193
12	Twin Cities	13	501,930	0.18%	922	66	5	53	46	752
13	Twin Cities (Hennepin & Ramsey)	13	112,419	0.42%	467	68	3	64	31	301
Total			1,271,521	0.26%	3,273	152	144	137	144	2,697

Table A-2. Cell Phone Sample Design, 2017

Strata	Region	All Households	Sample Ratio	Target Completes	African American	Native Americans American Indian	Asian	Hispanic /Latino	Residual (prim. Whites)
14	Northwest	173,805	0.16%	286	2	3	2	4	274
16	Headwaters	156,921	0.35%	554	6	56	2	19	471
17	Arrowhead	445,204	0.07%	314	3	12	3	13	284
18	West Central	250,621	0.08%	206	3	3	1	9	190
19	North Central	224,645	0.14%	316	3	12	1	3	297
20	Mid-Minnesota	187,382	0.11%	199	1	3	0	12	182
21	Upper MN Valley	51,898	0.41%	214	1	1	3	4	205
22	East Central	138,595	0.15%	212	4	4	2	7	194
23	Central	394,500	0.06%	232	10	1	5	6	211
24	Southwest (Nobles)	37,817	0.22%	85	1	1	0	19	64
25	Southwest (other)	213,782	0.11%	226	2	5	0	5	215
26	South Central	344,933	0.08%	284	3	0	5	10	266
27	Southeast (Olmstead)	261,728	0.07%	176	4	0	7	3	162
28	Southeast (Other)	433,295	0.06%	254	3	1	0	7	243
29	Twin Cities	4,644,423	0.11%	5,169	362	64	161	302	4,280
Total		7,959,549	0.11%	8,727	408	167	193	423	7,537