IMPROVING SURVEY ESTIMATES OF THE UNINSURED USING COMPUTER-ASSISTED INTERVIEWING LOGIC

Barbara Lepidus Carlson, Mathematica Policy Research, Inc.
P.O. Box 2393, Princeton, NJ 08543-2393, U.S.A.

Key Words: Uninsured, Health Insurance, CATI, Survey

INTRODUCTION

Estimating the number of people without health insurance is an important use of population-based surveys. As the health insurance arena becomes more complex, survey respondents often find it difficult to accurately report their insurance coverage. Several national surveys measuring health insurance go through a series of questions on the different categories of health insurance, and conclude that the lack of a positive response to all of these questions means the person does not have health insurance coverage. Back in 1993, we began to include a probe in our computer-assisted telephone interviewing (CATI) health insurance surveys that attempted to confirm that such people were in fact uninsured. Recently, we started to capture this information and examine the impact of this probe on estimates of the uninsured.

This paper reports findings based on three population-based health insurance CATI surveys that used the probe: one completed in Maine, one completed in North Dakota, and a national survey. The results include unweighted and weighted estimates of those who changed from uninsured to insured as a result of the probe, and what health insurance categories were chosen by those who changed insurance coverage status.

BACKGROUND

Several national surveys, including the Current Population Survey and the Medical Expenditure Panel Survey, obtain information about insurance status by posing a series of questions regarding the various types of insurance coverage. The series of questions generally ends with a catch-all, such as, "Other than the plans I have already talked about, was anyone in this household covered by any other type of health insurance plan?" A person is then classified as uninsured if he or she is in a residual category; that is, if the household or family informant has not responded affirmatively to any of the insurance questions for that person.

While such a person would in theory be uninsured, in practice this presumption is not always correct. For example, the informant may have neglected to mention one of the household or family members as being covered by a certain type of insurance, may not know the type of coverage of all members, may not understand or recognize some of the terms used to describe the various types of coverage, or may not consider certain types of coverage to be a health insurance plan. The interviewer can also make errors in recording the responses of the informant. The catch-all question about "other" types of insurance coverage can catch some, but not all, of the omissions. The only way additional omissions may be caught at a later point in the interview is if a question targeted at uninsured persons is clearly worded as such, and the respondent is forthright enough to challenge the question.

The use of computer-assisted technology in interviewing provides great opportunities for verifying information about insurance coverage. As soon as the series of insurance questions has been posed, including the catch-all question, the computer "knows" who is in the residual category. Mathematica has made use of this technology in two ways. After the series of insurance questions has been completed, the interviewer sees on his or her screen a summary of household insurance coverage as understood by the CATI program. At this point, the interviewer verifies the information with the respondent and has an opportunity to correct any errors or omissions. More importantly, immediately prior to this verification screen, the CATI program identifies anyone in the residual category and poses a question that specifically verifies whether the person is uninsured. This gives the informant the opportunity to confirm that the person is uninsured, to rectify an omission and provide information on the type of coverage, or to state that he or she does not know the insurance status or type of coverage for that household or family member.

Others have examined methodological issues relating to measuring the number of uninsured using surveys, in particular for comparing estimates of the uninsured across national surveys (Swartz 1986, Lewis et al. 1998, Rosenbach et al. 1998). In addition to the differences in question wording, sequence, and how the uninsured are defined, there are many survey characteristics that are believed to produce different uninsured estimates across surveys, including differences in: reference period; sampling and coverage; mode of interviewing; level of interviewer training on insurance questions; use of proxy respondents; and weighting, editing, and imputation. This paper focuses only on the impact of the uninsured probe on estimates *within* each of three surveys, and does not compare estimates across surveys; however, the results of this evaluation can be

used as a piece of the larger puzzle of trying to resolve differences in uninsured estimates *across* surveys.

METHODS

The results of this methodological investigation are based on the data from three surveys conducted by Mathematica Policy Research within the last couple of years:

- the Center for Studying Health System Change Community Tracking Study (CTS) Household Survey (1996-97), funded by the Robert Wood Johnson Foundation

- the Maine Health Insurance Coverage Survey (1997), under contract to the University of Southern Maine's Muskie School of Public Service, funded by a group of public and private organizations in Maine

- the North Dakota Health Insurance Survey (1998), under contract to the North Dakota Department of Health, funded by a grant from the Robert Wood Johnson Foundation.

The CTS Household Survey was a general population survey conducted primarily by telephone (using list-assisted random-digit dial [RDD] samples), with some additional interviews conducted in the field using cellular telephones to represent telephones with no or intermittent telephone service. The Maine survey was conducted entirely by telephone using RDD sampling, with the eligible population being households containing at least one child. The North Dakota survey was a general population survey conducted entirely by telephone using RDD sampling. More details on the methodologies of these three surveys can be found in final methodology reports available from the author.

All three surveys were administered using computer-assisted telephone interviewing (CATI), using CASES.[1] They all had a similar series of questions on health insurance coverage asked of a family or household informant. In all cases, each question in the series was worded as follows:

Are (either of / any of) you covered by .....?
IF YES: Who is covered?

Slight variations in wording among the three surveys are reflected below by words or phrases in square brackets. Alternate "fills" for pronouns or other words are reflected by words or phrases in parentheses. The various types of insurance coverage asked about sequentially[2] were:

- a health insurance plan from (either of /any of) your current or past employers or unions[3] [other than the military]

- a health insurance plan bought on your own [and not through an employer or union]

- a health insurance plan [provided by][held in the name of] someone who does not live in this household

- Medicare, the health insurance plan for people 65 years old and older or persons with certain disabilities (with a probe for persons 65+ not reporting yes)

- (Medicaid / Medical Assistance / MED-CAL / Welfare / MEDIKAN), [the government assistance program for people in need] [NYLCare Choice or PrimeCare {in Maine}]

- CHAMPUS, CHAMP-VA, TRICARE, VA, or some other military health care

- the Indian Health Service

- [(State-specific plan)]

- a [health insurance plan][state-sponsored or public {health insurance} program] that I have not mentioned. IF YES: What is the name of the [plan][program]?

This series of questions was followed up with the following question for each person without any "yes" response to the above series:

According to the information we have, (you/name) do(es) not have health care coverage of any kind. Do(es) (you/he/she) have health insurance or coverage through a plan I might have missed? INTERVIEWER: REVIEW PLANS IF INFORMANT IS UNSURE.
    <0> No/Not covered by any plan
    <1> Health insurance plan from a current or past employer/union/school
    <2> A health insurance plan bought on his/her own, professional association
    <3> A plan bought by someone who does not live in this household
    <4> Medicare

<5> Medicaid, [state-specific name]
[NYLCare Choice, or PrimeCare]
<6> CHAMPUS, CHAMP-VA, TRICARE,
VA, other military
<7> Indian Health Service
<8> [State-specific plan]
<9> Other plan
<98> Don't know
<99> Refused

Responses <1> through <9> would cause the CATI program to jump back to the appropriate prior question so that the answer could be changed to the correct response. At the end of the series of initial insurance questions, we had a verification screen [not sure about CTS] in which the interviewer verifies the insurance status of each household or family member. This is another point at which to confirm or disconfirm a status of uninsured. The effect of this verification screen is not the subject of this paper.

In the Maine and North Dakota surveys, we asked the CATI programmers to capture in a person-level variable whether this last question was asked and what the response was. At the time the CTS household survey was fielded, this information was lost. Other than for methodological research, there was no reason to keep the response, so it was treated as a standard verification question and wiped out when the prior response was corrected. Fortunately, we found a way to reconstruct whether the question was asked using a "trace" file that was set up for other quality control purposes. This trace file captured detailed information on the CATI sessions for most, but not all, of the data collection period. The results shown here for the CTS are only for those cases that were captured by the trace file. The CASES "history" file (similar to a keystroke file) was not usable as a resource of these purposes, because it would not have been possible to determine at what point in the interview a response was changed; that is, whether it was changed in response to the uninsured probe.

For each of the three surveys, I looked at the final person-level responses to the insurance questions and the response to the uninsured probe (or at least whether it was asked). From these, it was possible to determine the impact the probe had on people's responses to the insurance questions, and on the final weighted estimates pertaining to insurance coverage. This paper presents how many people changed their answer as a result of the probe, and what types of insurance were missed the first time through.

When making estimates from the data in each of the three surveys, it is necessary to incorporate analysis weights that account for variations in the probabilities of

selection, response rates, and coverage levels for various segments of the sampled population, in addition to inflating the sample so that it represents the population. Because of the stratification, clustering, and unequal weighting inherent in each of the surveys' samples, the variance of any survey estimate must account for these factors; therefore, the standard errors of the estimates of the uninsured rate were computed using SUDAAN software (SAS-callable version 7.5).

The weighted results presented here may not match other published results using these same data for a number of reasons. First, for the CTS household survey, the "other" insurance category presented here has not been backcoded to existing insurance categories, as the coding results from the verbatim responses were not available when this evaluation took place. Second, I have not disaggregated the "combined" insurance category into its various types--anyone reporting more than one type of insurance is placed into this category. Other resesarchers will undoubtedly set up certain hierarchies of types of insurance coverage, rather than include a "combined" category. (Note that the CTS household survey employed certain skip patterns so that some types of insurance coverage combinations were intentionally not ascertained.) Third, the way I dealt with "don't know," "refused," and other missing values may not match what other researchers have done. Finally, questions regarding whether the probe was in fact asked arose in several cases. The way I chose to classify them may not agree with the decisions of other researchers using these data. Furthermore, I have not excluded from these estimates those persons age 65 and older, whereas others may have.

RESULTS

Tables 1, 3, and 5 show the unweighted results for the three surveys, and Tables 2, 4, and 6 show the weighted results. The unweighted results give methodological results at the person level; that is, how many respondents were asked the probe and what their responses were. The weighted results show the impact of the probe on estimates of insurance status. In each table, the first column shows the distribution of responses to the series of insurance questions as they were ultimately answered. The second column shows the distribution of responses for those who were asked the uninsured probe. The third column shows what the distribution theoretically would have been had the probe not been asked; that is, all persons for whom there were no "yes" responses to the insurance question series would have been classified as uninsured.

It should be noted that, in practice, a family or household member misclassified as uninsured (in the absence of a probe) may have been later classified correctly if a subsequent question clearly stated the

assumption that the person was uninsured (or at the insurance verification screen). But given the way the CATI program was set up, the interviewers were told they could not go back and change the insurance status of household or family members after the insurance verification screen has been passed. It is at this point in the program that rosters and grids are set up according to the insurance status of each individual, which are then used for subsequent questions set up in macros. So it is essential that the insurance status of household or family members be correct before any follow-up questions are asked.

Maine Survey Results

Table 1 shows the unweighted results for the Maine survey. Responses were given by a household informant for 9,187 persons in households with at least one child. The results show that 12.3 percent of persons in the responding households were classified as uninsured. Note that this category includes persons with combinations of "no" and "don't know" responses to the series of insurance questions. More than two-thirds of persons were reported to have private insurance. Of the 9,187 persons in responding households, the probe was asked of 1,261; that is, 13.7 percent of persons went through the series of insurance questions without a "yes" response. While most of those asked the probe confirmed that they were uninsured (89.8 percent, including 50 who said "don't know" or refused to answer the probe), 129 persons were reported to be covered by some type of insurance as a result of the probe. (It is not clear whether the 50 persons with a response of "don't know" or "refused" to the probe were unable or unwilling to respond to *whether* the person was uninsured or to the *type* of insurance coverage.) Without the probe, the unweighted percent of persons in responding households that would have been classified as uninsured would have risen from 12.3 percent to 13.7 percent. Of the 129 persons for whom the probe prompted a change in insurance status, the vast majority (more than 97) said they were covered by private insurance.

The weighted results for the Maine survey in Table 2 show that 13.0 percent (standard error = 0.7 percent) of persons in households with children were uninsured, and more than two-thirds of such persons were covered by private insurance. Based on the results of the probe, we can postulate that the uninsured estimate would have been 14.3 percent (s.e.=0.7), a difference representing 8,655 people statewide, had the probe not been asked. The 95% confidence interval around this difference of 1.3 percentage points did not include zero.

North Dakota Survey Results

Table 3 shows preliminary unweighted results for the North Dakota survey. Responses were given by a

household informant for 13,156 persons in households. The results show that 8.6 percent of persons in the responding households were classified as uninsured. Again, this category includes persons with combinations of "no" and "don't know" responses to the series of insurance questions. Just under two-thirds of persons were reported to have private insurance only. Of the 13,156 persons in responding households, the probe was asked of 1,297; that is, 9.9 percent of persons went through the series of insurance questions without a "yes" response. While most of those asked the probe confirmed that they were uninsured (87.4 percent), 164 persons were reported to be covered by some type of insurance as a result of the probe. Without the probe, the unweighted percent of persons in responding households that would have been classified as uninsured would have risen from 8.6 percent to 9.9 percent. Of the 164 persons for whom the probe prompted a change in insurance status, the vast majority (more than 132) said they were covered by private insurance.

The preliminary weighted results for the North Dakota survey in Table 4 show that 9.2 percent (s.e.=0.4) of persons in households were uninsured. Based on the results of the probe, we can postulate that the uninsured estimate would have been 10.4 percent (s.e.=0.5), a difference representing 8,218 people statewide, had the probe not been asked. The 95% confidence interval around this difference of 1.2 percentage points did not include zero.

Community Tracking Survey Results

Table 5 shows the unweighted results for the CTS household survey. Responses were given by a family informant for 60,446 persons in households in the contiguous United States. The trace file mentioned earlier in the Methods section captured responses for 57,612 persons. The results show that 11.6 percent of persons in the responding households were classified as uninsured. Again, this category includes persons with combinations of "no" and "don't know" responses to the series of insurance questions. More than two-thirds of persons were reported to have private insurance, including some in the "combination" category. Of the 57,612 persons in responding households captured by the trace file, the probe was asked of 7,184; that is, 12.5 percent of persons went through the series of insurance questions without a "yes" response. While most of those asked the probe confirmed that they were uninsured (92.8 percent, including 85 who said "don't know" or refused to answer the probe), 515 persons were reported to be covered by some type of insurance as a result of the probe. Without the probe, the unweighted percent of persons in responding households that would have been classified as uninsured would have risen from 11.6

percent to 12.5 percent. Of the 515 persons for whom the probe prompted a change in insurance status, the vast majority (more than 334) said they were covered by private insurance.

Table 6 shows weighted results for the CTS household survey. We see that an estimated 13.0 percent (s.e.=0.6) of persons in the United States (excluding Alaska and Hawaii) were uninsured, and more than 60 percent were covered by private insurance. Without the probe, the uninsured estimate would have been 14.0 percent (s.e.=0.6), a difference representing 2.3 million people. The 95% confidence interval around this difference of 1 percentage point did not include zero.

CONCLUSIONS

While the impact of the uninsured probe is relatively small, it has been shown to be consistent in three separate surveys and to be significantly greater than zero. Using only a catch-all question at the end of the insurance questions does not sufficiently prompt household or family respondents to capture some previous omissions in reporting coverage. This is likely related primarily to two features of health insurance surveys. First is the fact that proxies are generally used for reporting insurance coverage. These respondents may forget to include a household or family member when listing who is covered by a particular type of health insurance. The uninsured probe verifies whether persons in the "residual" category were in fact uninsured, or were erroneously omitted from previous reports.

Second, the terminology for health insurance coverage, even terms like "insurance" and "plan," mean different things to different respondents. Furthermore, the line between public and private insurance plans is becoming blurred as Medicare, Medicaid, and state-specific health insurance programs contract with private insurance carriers to provide managed care. So some

respondents are waiting to hear a category that they think applies to their coverage, or that of their family members, and never hear it.

It is a fairly straightforward process to build into a CATI program a probe for persons in the residual category to verify whether they are in fact uninsured. Relying on respondents to respond correctly on the first pass through the questions, or to perhaps realize the error in a later question posed to the uninsured (and be forthright enough to bring it to the attention of the interviewer) is shown in this paper to be ill-advised. The findings in this paper also can be, and have been (Rosenbach et al. 1998), used to help explain differences in estimates of the uninsured among national surveys. We will continue to measure the impact of the probe on future CATI health insurance surveys.

REFERENCES

Lewis K, Ellwood M, and Czajka J (July 1998). "Counting the Uninsured: A Review of the Literature," Occasional Paper Number 8, *Assessing the New Federalism.* Washington, DC: The Urban Institute.

Rosenbach M and Lewis K (June 1998). "Reconciling Estimates of Health Insurance Coverage: Comparisons of the Community Tracking Study and the Current Population Survey." Report submitted to Center for Studying Health System Change. Cambridge, MA: Mathematica Policy Research.

Swartz K (1986). "Interpreting the Estimates from Four National Surveys of the Number of People Without Health Insurance," *Journal of Economic and Social Measurement*, 14:233-242.

END NOTES

1.The data collected and statistical results that are used in this paper were produced using computer programs made available through the Computer Assisted Survey Methods Program (CSM), University of California, Berkeley. Neither the CSM staff nor the University of California bear any responsibility for the results or conclusions presented here.

2. There were logical skips in the CTS so that certain types of coverage negated other questions on health insurance coverage. For example, if all persons in the family were covered by Medicare then the private insurance questions were not asked.

3. In the North Dakota survey, we first ask about employment for each adult, and whether the adult has coverage through his/her employer.

Table 1. Maine Health Insurance Survey (November 1997) - Unweighted Results

Insurance Classification	Distribution With Probe	Number Receiving Probe	Distribution Without Probe
Number of persons	9,187	1,261 (13.7%)	9,187
Private	6,171 (67.2%)	97 (7.7%)	6,074 (66.1%)
Medicare	69 (0.8%)	0	69 (0.8%)
Medicaid	875 (9.5%)	5 (0.4%)	870 (9.5%)
Military	167 (1.8%)	4 (0.3%)	163 (1.8%)
Indian Health Service	6 (0.1%)	0	6 (0.1%)
Other (coded private)	18 (0.2%)	8 (0.6%)	10 (0.1%)
Other (non-private)	19 (0.2%)	5 (0.4%)	14 (0.1%)
Combination	729 (7.9%)	10 (0.8%)	719 (7.8%)
Uninsured	1,133 (12.3%)	1,132 (89.8%)	1,262 (13.7%)

Table 2. Maine Health Insurance Survey (November 1997) - Weighted Results

Insurance Classification	Distribution With Probe	Number Receiving Probe	Distribution Without Probe
Number of persons	663,558	95,154 (14.3%)	663,558
Private	443,823 (66.9%)	6,654 (7.0%)	437,169 (65.9%)
Medicare	4,129 (0.6%)	0	4,129 (0.6%)
Medicaid	65,287 (9.8%)	249 (0.3%)	65,038 (9.8%)
Military	10,448 (1.6%)	270 (0.3%)	10,178 (1.5%)
Indian Health Service	188 (0.0%)	0	188 (0.0%)
Other (coded private)	1,155 (0.2%)	466 (0.5%)	689 (0.1%)
Other (non-private)	1,232 (0.2%)	228 (0.3%)	1,004 (0.1%)
Combination	50,741 (7.6%)	792 (0.8%)	49,949 (7.5%)
Uninsured	86,556 (13.0%)	86,498 (90.9%)	95,211 (14.3%)

Table 3. North Dakota Health Insurance Survey (Spring 1998) - Unweighted Results

Insurance Classification	Distribution With Probe	Number Receiving Probe	Distribution Without Probe
Number of persons	13,156	1,297 (9.9%)	13,156
Private	8,522 (64.8%)	132 (10.1%)	8,390 (63.8%)
Medicare	247 (1.9%)	3 (0.2%)	244 (1.9%)
Medicaid	273 (2.1%)	6 (0.5%)	267 (2.0%)
Military	366 (2.8%)	6 (0.5%)	360 (2.7%)
Indian Health Service	153 (1.2%)	3 (0.2%)	150 (1.1%)
Other (coded private)	32 (0.2%)	3 (0.2%)	29 (0.2%)
Other (non-private)	29 (0.2%)	2 (0.2%)	27 (0.2%)
Combination	2,400 (18.2%)	9 (0.7%)	2,391 (18.2%)
Uninsured	1,134 (8.6%)	1,133 (87.4%)	1,298 (9.9%)

Table 4. North Dakota Health Insurance Survey (Spring 1998) Weighted Results

Insurance Classificat'n	Distribution With Probe	Number Receiving Probe	Distribution Without Probe
Number of persons	666,408	69,234 (10.4%)	666,408
Private	405,148 (60.8%)	6,366 (9.2%)	398,782 (59.9%)
Medicare	14,913 (2.2%)	175 (0.3%)	14,738 (2.2%)
Medicaid	15,639 (2.3%)	436 (0.6%)	15,203 (2.3%)
Military	18,333 (2.8%)	276 (0.4%)	18,057 (2.7%)
Indian Hlth Service	10,562 (1.6%)	281 (0.4%)	10,281 (1.5%)
Other (coded priv.)	1,553 (0.2%)	207 (0.3%)	1,346 (0.2%)
Other (non-private)	2,307 (0.3%)	105 (0.2%)	2,202 (0.3%)
Combinat'n	136,874 (20.5%)	371 (0.5%)	136,503 (20.5%)
Uninsured	61,079 (9.2%)	61,017 (88.1%)	69,297 (10.4%)

Table 5. CSHSC Community Tracking Study Household Survey (1996-97) - Unweighted Results [captured by trace file]

Insurance Classification	Distribution With Probe	Number Receiving Probe	Distribution Without Probe
Number of persons	57,612	7184 (12.5%)	57612
Private	38,058 (66.0%)	334 (4.6%)	37724 (65.5%)
Medicare	7,123 (12.4%)	23 (0.3%)	7100 (12.3%)
Medicaid	2,327 (4.0%)	47 (0.7%)	2,280 (4.0%)
Military	686 (1.2%)	9 (0.1%)	677 (1.2%)
Indian Health Service	86 (0.1%)	2 (0.0%)	84 (0.1%)
State-Specific Program	88 (0.2%)	4 (0.1%)	84 (0.1%)
Other (uncoded)	844 (1.5%)	91 (1.3%)	753 (1.3%)
Combination	1,712 (3.0%)	5 (0.1%)	1,707 (3.0%)
Missing	19 (0.0%)	-	19 (0.0%)
Uninsured	6,669 (11.6%)	6,669 (92.8%)	7,184 (12.5%)

Table 6. CSHSC Community Tracking Study Household Survey (1996-97) - Weighted Results [captured by trace file]

Insurance Classificat'n	Distribution With Probe	Number Receiving Probe	Distribution Without Probe
Number of persons	251,501,843	35,128,170 (14.0%)	251,501,843
Private	155,150,122 (61.7%)	1,358,407 (3.9%)	153,791,715 (61.1%)
Medicare	33,668,003 (13.4%)	112,789 (0.3%)	33,555,214 (13.3%)
Medicaid	14,844,101 (5.9%)	339,342 (1.0%)	14,504,759 (5.8%)
Military	2,988,073 (1.2%)	32,322 (0.1%)	2,955,751 (1.2%)
Indian Health Service	479,521 (0.2%)	15,302 (0.0%)	464,219 (0.2%)
State-Specific Program	401,000 (0.2%)	13,342 (0.0%)	387,658 (0.2%)
Other (uncoded)	3,617,419 (1.4%)	437,632 (1.2%)	3,179,787 (1.3%)
Combination	7,482,045 (3.0%)	29,846 (0.1%)	7,452,199 (3.0%)
Missing	82,371 (0.0%)	-	82,371 (0.0%)
Uninsured	32,789,187 (13.0%)	32,789,187 (93.3%)	35,128,170 (14.0%)

N.B. These figures may not match published estimates from these surveys.