Memorandum on Data Guide: Analysis Results for Understanding Survey Estimates of California's Uninsured and Medi-Cal Populations

Authors:
Jack Feldman, Claudia Schur, Marc Berk, and Sreelata Kintala

January 2004

The purpose of this project was to produce a practical guide to using estimates of the number of uninsured California residents and of the Medi-Cal population for use by policymakers, legislators, media, and advocacy groups. The development of the guide was based on empirical analyses of three survey data sets—the Current Population Survey (CPS), the National Health Interview Survey (NHIS), and the California Health Interview Survey (CHIS). These analyses were conducted in order to identify sources of differences across the three surveys in estimates of the uninsured and Medi-Cal populations. This memo describes findings from these analyses, beginning with a description of the differences in the estimates. The sources of the differences and some general recommendations for using the estimates are also offered.

A great deal of methodological work analyzing the CHIS estimates of the uninsured and comparing them to CPS estimates has already been conducted, therefore an attempt was made to avoid conducting extensive analyses that would duplicate others' efforts.[i] Instead, the data guide primarily focuses on comparison of estimates for different population sub-groups in order to see if the survey differences varied in any systematic way. In addition to CPS and CHIS, an analyses of the California sample of the NHIS is also included.

Differences in the Estimates

We begin here by examining the percentage of the population without any health insurance coverage and the percentage of the population enrolled in Medi-Cal for the state as a whole and for population sub-groups within California. It should be noted that the accuracy of reporting for Medi-Cal bears a critical relationship to the accuracy of the estimates of the uninsured. If one of the surveys obtains a higher count of persons enrolled in Medi-Cal (or having private coverage), then this would likely result in a decrease in the estimate of the number of uninsured.

In order to examine differences across the surveys, we provide as an attachment to this memo estimates for each of the three surveys, focusing on point-in-time estimates.[ii] More detail on how the different time periods to which the insurance variables refer in each survey is provided in the next section of this memo. Tables 1 and 2 show the uninsured rates for selected groups for children and adults, respectively. Tables 3 and 4 provide data on Medi-Cal enrollment, again, separately for children and adults. The tables include both weighted percentages and the ratio of the percentage estimates—for CPS relative to CHIS and for NHIS relative to CHIS.[iii] These ratios allow us to see how the difference between surveys varies for different population sub-groups.

For the statewide estimate of the uninsured, CPS and NHIS give similar results with approximately 18 percent of the population uninsured in 2000 (not shown in tables). The CHIS estimate is lower at 14 percent. For children, NHIS and CPS estimates are also similar (15.5 and 15.4 percent uninsured, respectively), while the CHIS estimate is again much lower at 9.6 percent. The relationship among estimates from the three surveys is similar for adults, though the gap between the federal surveys and CHIS is much smaller for adults than for children. With respect to both children and adults, estimates from CPS and NHIS are fairly similar throughout.

The ratios in Tables 1 and 2 represent one indicator of the relationship between the survey estimates. For children the ratios are 1.6 and 1.61 for CPS and NHIS, respectively, relative to CHIS. For adults,

these ratios are much lower—1.22 and 1.25—meaning that the survey estimates are much closer together. One can see that this level of difference is fairly uniform across most of the sub-groups examined though there are a few exceptions. The survey estimates for recent immigrant children are similar for CPS and CHIS and closer together than for the overall population for NHIS and CHIS (ratio of 1.34). For adults, the ratios are also relatively constant across the population sub-groups examined.[iv]

In terms of Medi-Cal enrollment, the estimates across surveys are more uniform for children than for adults. With respect to children, the CPS and CHIS estimates align more closely than does NHIS with either survey. Both CPS and CHIS show about 23 percent of children on Medi-Cal while the NHIS estimate is slightly lower at 19 percent. The ratios show relatively little variation across population subgroups, particularly for the NHIS-to-CHIS ratio. The CPS/CHIS ratio for white children is somewhat larger at 1.39 and the gap between these two surveys is also somewhat wider for children in families with incomes above the poverty level. The estimates of the percent of adults enrolled in Medi-Cal are similar for CPS and NHIS (8.5 percent and 8.3 percent, respectively) and both surveys are lower than CHIS (11.6 percent). The magnitude of the gap across surveys is, again, fairly uniform.

Identifying the Sources of these Differences

There are a number of potential explanations for the differences in survey estimates, based on differences in the design and implementation of the surveys. ere we discuss some of these differences and the likely effect that we believe these differences have on the final results.

Response rates. It has generally been accepted that, all other things being equal, higher response rates result in higher quality data. This view is based on the assumption that a lower response rate is likely to be reflective of some systematic bias—in an employer survey perhaps it is small employers without benefits managers who disproportionately do not respond; in a survey about access-to-care problems it might be those *with* access problems who are more likely to participate. In either case, if characteristics of the non-responders are systematically related to the variables of interest, then a low response rate might impart considerable bias.

Response rates in both NHIS and CPS (usually upwards of 85 percent) are substantially higher than that in CHIS (approximately 37 percent). This is related to both cost (more resources can increase response rates by increasing the level of follow-up and supporting more experienced interviewers) and mode of interview (telephone surveys generally obtain lower response rates). Most survey methodologists would consider a response rate of 37 percent to be problematic, though some portion of the bias from low response rates can be corrected through non-response adjustments and adjustment to known 'control totals.' Estimates for all three surveys are adjusted to control totals for characteristics found on the decennial Census (such as age, sex, race, and urbanicity) that may be correlated with insurance status. There are no control totals for health insurance from the Census.

Mode of interview. Interviews for NHIS are conducted in person. The initial interview for CPS is conducted in-person as are all subsequent interviews for persons without telephones. CHIS is done over the telephone. This leads to two potential sources of error. First, there is the issue of coverage[v]—do all members of the target population have a non-zero probability of selection for the sample? Because some people don't have telephones—and the uninsured are more likely to be without

telephone coverage than the insured—there is some systematic bias in using telephone surveys to measure the rate of uninsurance. In calculating sampling weights, CHIS statisticians used information from the 2000 CPS on the characteristics of non-telephone households to adjust CHIS weighted population counts. This assumes that the characteristics of persons within given categories (e.g., AFDC recipients) are identical with regard to health insurance status, regardless of whether they are part of telephone or non-telephone households.[vi]

The other possible source of bias from telephone interviewing has to do with verification of insurance status. In-person interviews allow the interviewer to examine the respondent's health insurance card and confirm that the coverage is current and is true health insurance (rather than, for example, a long-term care insurance policy or dental-only plan). In addition, NHIS uses this opportunity to collect the name of the insurer and later cross-checks the name with a verified list of insurers. In each of the past several years (2000-2002), this editing process has led to a net increase of 1.42 to 1.49 percentage *points* in the percent uninsured.[vii] This could theoretically account for one-fifth of the gap between the estimates.

Unit of interview. In most household surveys, the interviewer first enumerates all the members of the household. One approach is to then collect pertinent information on each member sequentially. An alternative is to use an algorithm to select one member of the household for interviewing. Both CPS and NHIS enumerate the household and then collect information on all persons living in the unit. CHIS, on the other hand, enumerates the household but collects information concerning only one (selected) adult respondent, one child less than age 12, and one adolescent aged 12 to 17. Differences in the quality of reporting using these two approaches has been only minimally explored. To provide some sense of how these different approaches might affect estimates, we compared the percent uninsured and percent enrolled in Medi-Cal for one-person households vs. households with 2 or more persons. In this preliminary analysis, we see that estimates for one-person households are more similar across surveys than for 2-plus households. This suggests that reporting for another person is not as accurate as reporting for oneself and lends some credence to the CHIS estimates.[viii]

Reference period. Certainly much attention has been paid to the differences between estimates of the uninsured based on the different time periods to which they refer. With respect to estimates of the uninsured, each of the surveys has a slightly different approach. NHIS asks about insurance coverage at the time of the interview—referred to as a point-in-time estimate. The CPS question refers to the past calendar year in asking about coverage. If taken at face value, it should give an estimate of persons who were uninsured for the entire year and this is, in fact, how it is used by Census Bureau analysts; however, some other analysts interpret the estimates as being point-in-time. From CHIS, one can calculate three different estimates—a point-in-time estimate, an estimate of persons uninsured all of the previous year, and an estimate of persons uninsured at any time during the previous year.

For purposes of comparison, we generally use the point-in-time estimate from CHIS. While this may not be the estimate that is most comparable to CPS conceptually, it is a commonly used approach and provides a conservative estimate of the difference between the surveys.

Question Structure and Wording. The specific way in which questions about insurance coverage are asked—the directness and ordering of the questions—are likely to have a substantial impact on how respondents answer. Many surveys—including both CPS and CHIS—enumerate each type of insurance separately and in a direct fashion (i.e., Is respondent covered by Medicare?). NHIS does not directly name different types of coverage, but rather relies on respondents to identify the type of

coverage. In addition, some surveys include follow-up questions that attempt to confirm and/or correct a respondent—for example, for persons 65 and older, "You previously told us that you are not covered by Medicare. Is this true?" It is likely that the structure and wording of questions, as well as their placement in the overall interview and the total length of the interview have an effect on responses and on subsequent estimates of insurance status. It is difficult to say much about the size of that effect, but it could be substantial.

Recommendations for Using the Estimates

Our analysis indicates that there are substantial differences in the estimates from the three surveys, particularly with respect to measuring the number of uninsured. The biggest gaps in estimates are for children. Within children and adults, however, we have shown—with a few exceptions—that the differences are fairly constant across a number of different population sub-groups. From this, we infer that there is no magic bullet to explain the differences. Neither is there a gold standard against which to benchmark the estimates, particularly for the estimates of the uninsured. It should be noted that, with respect to Medi-Cal enrollment, there is a gold standard in the administrative data[ix] and so we focus most of our discussion on the estimates of the uninsured. Because we find no clear explanation, we are left concluding that none of the surveys is either right or wrong. While from a traditional survey methods perspective, CHIS contains more design elements that would lead to bias than does either NHIS or CPS, a recent and sophisticated assessment comparing a phone and in-person survey conducted by NCHS contributes to our uncertainty. This study compared estimates of uninsured children from NHIS and SLAITS (the State and Local Area Integrated Telephone Survey) and concluded that it was impossible to determine which survey was correct without external verification of insurance status.[x] Of particular interest, the SLAITS estimate of uninsured children in California is remarkably close to the CHIS estimate.

While this sort of conclusion may suffice for a methodological study, a practical approach to using conflicting estimates is needed. The recommendations we provide here are based on our analysis of 2001 survey data; while the general principles are applicable to other time periods, the specific multipliers used are not. Our recommendations on using the estimates begin with the assumption that the national surveys should be used for making state-to-state comparisons because it is generally better to make comparisons using data from the same source. Thus, even if the CHIS estimates were correct and the CPS/NHIS estimates were wrong, it would make practical sense to use the latter estimates when making state-to-state comparisons. Otherwise, California would artificially appear to have a relatively lower number of uninsured than other states.

To make estimates within California, however, one must primarily rely on CHIS since the federal surveys cannot support estimates of most local geographic areas. When comparing two geographic localities within California, using CHIS estimates without any adjustment would be appropriate.[xi] It makes sense to use CHIS to say that there is a higher percentage of uninsured in one county compared to another.

However, if the absolute number of persons is important when making a sub-state comparison or estimate (e.g., when calculating the cost of a change in policy), it may make sense to use a range estimate. The lower bound of the range would be the CHIS estimate and the upper bound would be an extrapolation to what the CPS or NHIS estimates would look like if sub-state estimates were available. To obtain this upper bound, we recommend using the statewide ratio of the two estimates.[xii] In other

words, one would multiply the CHIS estimate by 1.29 (for all ages), 1.6 (for children), and 1.22 (for adults) to get an upper bound for the range. In doing so, particular attention should be paid when making estimates concerning children since the gap between the survey estimates is so much larger than it is for adults. By blending the estimates from the different surveys, one takes advantage of the large sample size of CHIS as well as the validation procedures and in-person interviewing used in NHIS or CPS, resulting in a more complete and accurate picture of the health insurance experience of California residents.

Endnotes

[i] There are two analyses that deal directly with these issues—UCLA Center for Health Policy Research, The State of Health Insurance in California: Findings from the 2001 California Health Interview Survey, Appendix. Survey Methods and Effects on Results. Available at www.healthpolicy.ucla.edu/pubs/files/shic062002.pdf. Also see California Health Interview Survey. CHIS 2001 Methodology Series: Report I- Sample Design. Los Angeles, CA: UCLA Center for Health Policy Research, 2002. Available at www.chis.ucla.edu/pdf/CHIS2001_method1.pdf

[ii] For CHIS and NHIS we use a question asking about health insurance coverage at the time of the interview. For CPS, the question wording refers to the past year and is interpreted differently by different analysts—some interpret the question literally as being uninsured for the entire year while others view it as a point-in-time estimate.

[iii] We rely on percents because the weighted population estimates (the denominators of the percents) differ substantially across surveys and compound the differences in the percent estimates.

[iv] The differences for the elderly are due to small sample sizes and low percents.

[v] A number of telephone surveys have attempted to correct for potential bias through changes in survey procedures. Both the National Survey of America's Families and the Community Tracking Survey gave cell phones to respondents who did not have phones. The RWJF National Access to Care Family used a small in-person sample to supplement telephone interviews. These procedures may help to correct for the phone coverage problem, though not for verification issues.

[vi] A 1990 study of telephone vs. non-telephone households based on three years of NHIS data concluded that these were distinct populations and that it was not possible to predict health insurance status for non-telephone households based on data from telephone households. This would suggest that weighting could not fully correct for any bias. Corey CR and Freeman HE. "Use of Telephone Interviewing in Health Care Research," *Health Services Research* 25:1 (April 1990, Part I). An analysis of results from the National Survey of America's Families and the Community Tracking Study (both of which included an in-person sample) show that estimates of health insurance from the two samples are dramatically different. As an illustration, in CTS, the percent of adults without insurance from the RDD sample was 15.5% compared to 50.5% for the non-telephone sample. For more details, see Hall J, Kenney G, Shapiro G, and Flores-Cervantes I. "Bias From Excluding Households without Telephones in Random Digit Dialing Surveys: Results of Two Surveys." Proceedings of the Survey Research Methods Section, American Statistical Association (1999), pp. 382-387.

[vii] Information obtained from NCHS staff.

[viii] Previous research on proxy reporting (i.e., reporting for other persons) would lend tentative support to this finding.

[ix] Findings pertinent to this issue should be forthcoming from a RAND study being conducted for CHCF. From a conversation with the RAND analyst, it appears that under-reporting for Medi-Cal is substantial and, if corrected, may have a large impact on the count of the uninsured.

[x] Blumberg SJ, Osborn L, Luke JV, et al. "Estimating the Prevalence of Uninsured Children: An Evaluation of Data from the National Survey of Children with Special Health Care Needs, 2001." National Center for Health Statistics. Vital Health Stat. Forthcoming.

[xi] Due to small sample sizes in NHIS and CPS for sub-state areas, it is difficult to draw any conclusions as to whether specific geographic areas within California may be more or less subject to the statewide potential undercount of the uninsured.

[xii] While these specific ratios may vary from year to year, the approach of using the ratio of estimates for a given sub-group as the multiplier remains the same.