Methodology Explanation and Documentation
for
Nevada-Specific Estimates of the Uninsured

Presented to

Great Basin Primary Care Association

Carole L. Popoff
Betsy Fadali
Dean H. Judson
Decision Analytics, Inc.

6/3/00

(demogecon@aol.com)

Executive Summary

- In late 1999-early 2000, Decision Analytics, Inc., contracted with Great Basin Primary Care Association to produce a set of estimates of the number of children uninsured in Nevada, broken out by various characteristics. These estimates were to be made using a well-established and defensible method.
- Several options for data sources were considered: Using state-specific surveys, using state agency administrative data, and using the Current Population Survey March Insurance supplement by the U.S. Census Bureau. Existing state-specific surveys have significant flaws, and state administrative data did not adequately cover the population of interest. Furthermore, Great Basin Primary Health Care Association desired estimates at the county level.
- Given these considerations, the team chose to use Current Population Survey data sources ***tied*** to age, race, sex, and Hispanic origin estimates that were formerly used by the Nevada State Demographer's office. The method was developed in Oregon for use in the Oregon Health Plan administration, and similar methods have been proposed by the Urban Institute, National Institutes of Health, and others. The system provides estimates for each county in Nevada.
- The synthetic estimates system generates estimates at the state level that are broadly consistent with independent counts provided by the State Division of Health Care Financing and Policy, and that are consistent with similar estimates developed from the Behavioral Risk Factors Surveillance Survey, the Annie E. Casey foundation, and the U.S. Census Bureau.
- The synthetic estimates system generates the following estimates:

	July, 1998	July, 1999
All children (aged 0-18)		
Number uninsured	105,436	112,263
% of total population of children (0-18)	20.6%	20.7%
Children living in families whose income is less than 100% of the federal poverty level		

Number uninsured	31,306	33,388
% of total population of children (0-18)	6.1%	6.1%
Children living in families whose income is greater than 100% and less than or equal to 200% of the federal poverty level		
Number uninsured	31,458	33,660
% of total population of children (0-18)	6.1%	6.2%

Introduction

There is an acute need for additional detail in the areas of health care delivery, needs assessment, policy development, and others. With the rapidly changing nature of the health care system at both the state and national level, it is imperative that researchers and policy makers have access to useful and reliable information about the current environment. It is especially important that data be available to assess the need for publicly-funded programs given the size of the population they serve. Thus, the need for detailed estimates of population health care-related characteristics has increased dramatically in recent years, especially for smaller jurisdictions such as counties. This study details methodology that can be used by public officials and other interested parties to produce local-level estimates of health care-related characteristics or virtually any policy issue.

The characteristic of interest for this study is persons who have no health insurance[1], and are thus labeled uninsured. In particular, the focus of this study is to estimate the number of children, defined as 18 or younger, at or below 100%, 100%-to-200%, and below 200% of poverty level. Part of the study was to evaluate available methods of estimation and choose the method that satisfies certain criteria.

Method Selection -- Evaluation Criteria Used

In order to estimate the percentage and number uninsured for various groups, demographers have typically used one (or more) of three common techniques: 1) state- or program-sponsored regional surveys; 2) administrative data sets maintained at the state level; or 3) state level data from Census' Current Population Survey (CPS). We considered the following five goals in evaluating the feasibility of implementing each of the above-mentioned options for the purposes of this study. The estimation system must:

1. provide estimates useable (and reasonable) at the county level of geography;
2. serve intercensal estimation needs on a regular, periodic basis;
3. provide comparability over time;
4. be defensible -- statistically valid and reliable; and,
5. be within budgetary and time constraints.

Each of the three above-mentioned methods has features that make it useful for certain types of estimates, but we note that all these have weaknesses. Unfortunately, county-level special surveys are expensive, administrative records series often have very limited personal and household characteristics detail or are not designed for such work, and state-level data does not accommodate smaller aggregation. Therefore, we must use statistical

[1] For purposes of this study, our definition of "uninsured" is consistent with the Census Bureau's Current Population Survey definition – an uninsured person is someone has had no insurance coverage or access to coverage (private or government-sponsored) for the entire calendar year.

methods applied to the data that exist to arrive at reasonable estimates of the uninsured. Our responsibility is to pick the appropriate data source and the appropriate technique.

- A special survey designed to estimate the topic-specific characteristics

Some states and counties have developed topic-specific surveys to collect data. For example, the Nevada Center for Business and Economic Research at the University of Nevada, Las Vegas (UNLV) conducted a statewide demographic survey at a cost of approximately $93,000. Despite efforts to increase response rates, the initial 10.1% response rate was increased with a later telephone follow-up to only 18.2% (Pencek, Daneshvary, and Schwer, 1998:244). **However, this meager response rate raises a significant concern: The non-respondents are exactly the people more likely to be uninsured. Health insurance status is highly correlated with social and economic characteristics that also create survey non-response [2]. Thus, any attempt to estimate the number and proportion of uninsured is almost certainly badly biased downward (lower than the true number)**. (Dillman, 1978; Tanur, 1981; Judson, 1998; the response rate to this survey was described as "very very very very problematic."). Further, the cost of collecting local area data from these surveys can be exceptionally prohibitive, costing each county $12,000 (for 400 respondents at $30 per respondent) to $100,000 (for 2,000 respondents at $50 per respondent). Repeating this operation on a reasonable periodic basis to determine changes in eligibility can be financially prohibitive.

- Pure administrative record estimates of topic-specific characteristics

Several sources of administrative data are available to assist in creating estimates of health characteristics. For many diseases or conditions, state health officials maintain data registries such as cancer or immunization registries. Even with state or federal mandates to maintain them, these registries are often incomplete. Administrative data estimates are even more difficult to develop for a characteristic like health insurance coverage because, in most cases, one data source would not represent the total population. It would be necessary to combine and sort information from a variety of sources and make sure records are not duplicated. Also, Medicaid, Medicare and insurance department enrollment records help in estimating the population covered by insurance, but these records are not complete counts of the insured population. For example, if one person from a household accesses Medicaid, by program definition all family members are eligible for Medicaid. However, there is no reliable way to estimate the number of family members from these records, thus the estimates would be lower (biased downward) than the true number. Finally, in many cases these records are subject to privacy and confidentiality legislation that may prevent access. For a list of potential errors, biases, and assorted limitations of administrative records data, see Judson and Popoff, 1999.

[2] For example, those of Hispanic origin, or those who are primarily Spanish speaking. For example, using Current Population Survey data for Nevada, Hispanic white males under age 18 have a **38%** chance of being uninsured. Non-Hispanic white males under age 18 have a **16%** chance of being uninsured.

- CPS-based estimates of topic-specific characteristics

The US Census Bureau's Current Population Survey (CPS) is currently the only source for state-to-state comparative health insurance information for uninsured children.[3] However, using the CPS as the only source of data is not adequate to accurately assess health insurance coverage at the sub-state level. In fact, the CPS was not originally intended to act as a state level data source—while each states' sample is independent of other states, it was intended to represent national level information, (Bureau of Labor Statistics, 1999)[4]. Because the sample size of this survey, for the majority of states, is relatively small and the Census Bureau does not stratify the sample into sub-state areas, urban areas will tend to be heavily over-represented. Therefore, using the CPS survey data alone for county-level estimates will distort the number of people uninsured in counties whose population make-up is quite different from the state as a whole.

Given the goals set for this study we chose to use a mixed-method approach to estimate the uninsured portion of the population, which is based on an approach called "synthetic estimation". Using this approach, two (or more, in some cases) sources of data are brought together. One source contains information about the number and proportion of the population that are uninsured at the state level. The second source contains detailed information about the number of people in each county, grouped by certain demographic characteristics that have been shown to correlate with being uninsured.

A New Approach: A Synthetic Estimates System

This method serves the five goals above: county-level estimates are more likely to be representative; estimates can be made annually because the CPS and updated age/race/sex/hispanic origin estimates tied to current population estimates are done annually; is feasible to implement; and it is an economical system.

Basically the synthetic estimates system can be described as tying two independent sources of information together:

1) Age/race/sex/hispanic origin (ARSH) estimates at the county level, and
2) survey data at the state level of geography.

Basically, this method:

---

[3] The Behavioral Risk Factor Survey (BRFSS) gives state data on uninsured, but only for adults. Survey of Income and Program Participation (SIPP) has data on health insurance but isn't as timely (1996 is available now), and although the survey has state identifiers for most states it was not designed to be representative at the state level.

[4] The Source and Accuracy of the Data for the March 1998 Current Population Survey "Accuracy of State Estimates" states that the redesign of the CPS following the 1980 census provided an opportunity to increase efficiency and accuracy of state data. All strata are now defined within state boundaries. The sample is allocated among the states to produce state and national estimates with the required accuracy while keeping total sample size to a minimum. Since the CPS is designed to produce both state and national estimates, the proportion of the total population sampled and the sampling rates differ among the states. In general, the smaller the population of the state the larger the sampling proportion.

- Uses the ARSH estimates to take advantage of the high correlation between individuals' ARSH characteristics and other characteristics of interest;
- Constructs detailed estimates at a lower level of geography (county) because the ARSH estimates are done at this lower level of geography; and
- Links the county ARSH estimates to state-level survey data for proportion of uninsured.

The notion of synthetic estimation has been well-described by such authors as Gonzalez (1973), Gonzalez and Waksberg (1973) and Gonzalez and Hoza (1978). In the context of health-related characteristics, the National Center for Health Statistics has performed a variety of studies, both theoretical and empirical, on the usefulness of synthetic estimation techniques (Levy, 1971; National Center for Health Statistics, 1977; 1979; Schaible, Brock and Schnack, 1977). More recently, Reder (1997) developed methods for estimating literacy proficiencies for populations using the National Adult Literacy Survey, State Adult Literacy Surveys, Public Use Microdata Sample data, and 1990 Decennial Census tabulations for counties[5]. Finally, we note for the record that this approach has been used by the authors in previous estimates for use by the Oregon Health Plan Demonstration project (Sigmund, Judson, and Popoff, 1999).

The Two Sources of Data Explained– ARSH Estimates and Current Population Survey (CPS) Data

- ARSH Estimates

There are two powerful reasons to use ARSH estimates:

1) Because the changes in the proportions change slowly and systematically over time, updated ARSH estimates are highly likely to be representative of the current population; and
2) ARSH estimates capture characteristics related to age, race, sex or Hispanic origin that make estimates that are correlated or tied to age, race, sex or Hispanic origin characteristics more specific.

We note for the record that our method of making ARSH estimates, the cohort-component method, has been used in the state of Nevada for the past five years by the Nevada State Demographer's Office, including its use in executive branch mandated budgeting and strategic planning initiatives. Further, the cohort-component method is a standard technique used by the Census Bureau and the states of New York and Texas, among others, for making ARSH estimates. It is useful in this context because it is Nevada-specific: Components such as births and deaths and migration are calculated for Nevada, and the final totals are controlled to the official State Demographer estimates.

[5] The approach presented here builds directly on Reder's work. Though the methods differ in particulars, the notion of using survey data linked with local area ARSH characteristics is taken directly from Reder (1994; 1997).

Although the specifics of the ARSH model are complex, and it makes use of a very large amount of background data, the general structure is simple. It is:

Population (this year) = Population (last year) + [Births – Deaths] + Net Migration

This equation is applied to each age group. For years where the data are available, we use birth and death counts from Nevada's Vital Statistics. Where data are not yet available, we use known Nevada-specific fertility and mortality probabilities by ARSH category to produce estimates of the number who have died and the number who are born. See the appendix for a full explanation.

- Current Population Survey (CPS) – March Supplement -- Data

The Current Population Survey (CPS) is a joint project of the Census Bureau and the Bureau of Labor Statistics. The March Supplement of the CPS is an add-on to the regular monthly survey which focuses on labor force data. In March of each year the supplemental survey asks questions about health insurance. The universe for the survey is the civilian noninstitutional population as well as members of the armed forces living in civilian housing either on or off a base. Survey participants are asked whether they have had any health insurance coverage at any time during the previous year (thus the 1999 March supplement data refers back to the calendar year 1998). Estimates are then made using the survey answers. Those who do not report having had any source of health insurance during the previous year are counted as uninsured. Starting with the 1998 CPS files, people with only Indian Health Service coverage are also counted as uninsured – previously they were counted as covered by Medicaid.

The nationwide sample is around 50,000 households. There is only about a 5% non-response rate for the ongoing portion of the survey, and there is only about a 15% non-response rate for the supplement part of the survey. The census bureau weights the data with national age, race, sex and Hispanic estimates to compensate for non-response. We used three years of survey data combined to form estimates of the number and proportion of uninsured in each category. In Nevada, there were 611 households in the 1997 sample, 668 in the 1998 sample and 731 in the 1999 sample. When combined the three years of survey data total 2010 households. This results in a total of 5233 records on individuals. We needed a sufficient number of persons so that each ARSH cell would contain enough people to make the estimate of uninsured meaningful.

CPS data on uninsured are easily the most reliable, consistent, cheapest and timely data available for state-level estimates. The CPS has been in existence for most of the century, starting in 1940; it uses specially trained interviewers who specifically maintain over-time relationships with respondents to minimize non-response; extreme care is used in questionnaire design; and the design is standardized across the country, using standard definitions and methods.

For these reasons the CPS has been used extensively to estimate uninsured (U.S. Census Bureau, 1999; Annie E. Casey Foundation [Kids Count], 1999; Liska, Brennan and Bruen [Urban Institute], 1998; Employee Benefit Research Institute,1999; and Center for Budget Priorities, 1999, for example). However, there are some drawbacks to using these data, as there are to any sample. Because the data are so widely used their limitations (and corrections for these limitations) are well-known (Census Bureau, 1999; Liska, Brennan and Bruen, 1998). Three of the most important challenges are:

1) Medicaid underreporting -- the census bureau recodes as "covered by Medicaid" those who indicate they received some other benefit that would make them eligible for Medicaid, nevertheless, when compared to administrative records CPS still appears to underestimate Medicaid coverage.

2) Possible reporting of point-in-time insurance status rather than what is asked for: the previous year's coverage -- We assume (as the Census Bureau does when making their estimates of uninsured) that respondents are reporting their coverage for the entire year though there is some evidence that some respondents are reporting about a point-in-time or do not remember the situation for the entire past year (see Liska, Brennan, and Bruen, 1998). We note that a point-in-time estimate of uninsured would tend to be larger than an estimate of those who did not have any coverage throughout the entire year. Conversely, the point-in-time estimate would tend to be smaller than an estimate of people who had no insurance for a least a period of time throughout the year.

3) The design as a national survey means that small area data, having smaller sample sizes, is less reliable. A partial answer to this problem is combining more than one state or combining more than one year of data, as we have done**.** Combining years is also done by the Census Bureau, Urban Institute, Kids Count and others. It is a standard method for increasing sample size for the purposes of estimation. In fact, this method is part of the multi-million dollar American Community Survey project which will be replacing the Census long form in 2010.

- Basic Method Illustrated

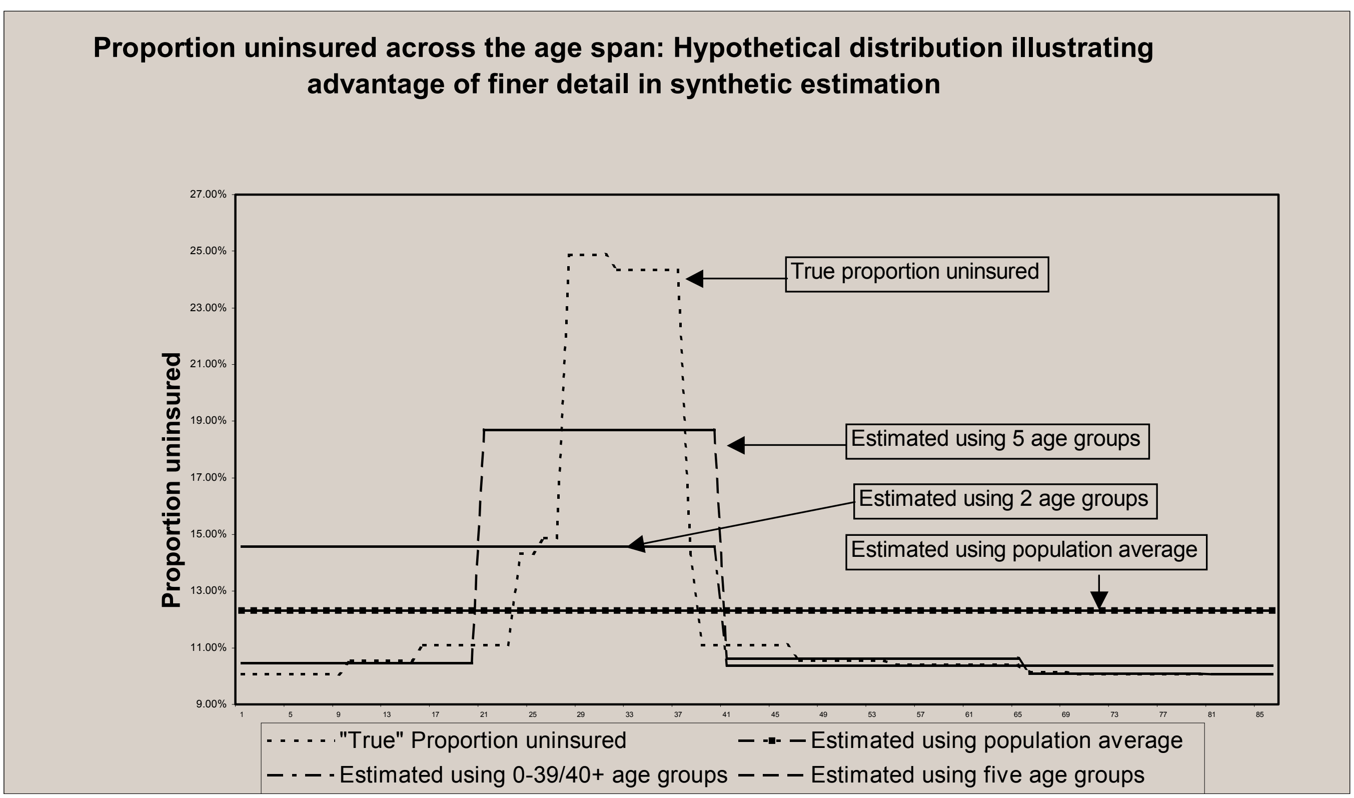
Step 1 – how ARSH Estimates are used: The basic synthetic estimation method, and our expansion of it, can be illustrated quite simply. For example, suppose that the average overall proportion of the population uninsured were 17%. The simplest kind of estimate would be to multiply the total population, times the overall population average proportion uninsured, 17%, to make an estimate of the number uninsured. For example, in 1999, the population of Nevada was estimated to be 1,965,659. Multiplying this number by 17% generates an estimated 334,162 persons who are uninsured.

However, this calculation is too simple. Because certain groups have uninsurance rates that are dramatically different than the average, this proportion is dramatically incorrect for those groups that deviate from the average.

We can illustrate this quite simply. For example, let us temporarily ignore race, sex, or Hispanic origin and focus only on age. If uninsurance rates vary by age (they do), then simply multiplying total population times overall proportion uninsured is not enough. Instead, we must make the estimate **group by group**, and only after making the estimate, we then add up the individual cell estimates to obtain the final number.

Figure one illustrates the advantage of making estimates by individual ARSH cells and is based on the following assumptions. First, that population can be broken into 5 age groups, 0-19, 20-39, 40-64, 64-79, 80+. Second, for this figure, we assume that **12%** of the total population is uninsured, but that the proportion is not distributed uniformly across the age groups. This figure demonstrates why breaking the population into finer groups must necessarily generate more correct estimates. The dotted line represents the "true" proportion uninsured by single year of age.

As can be seen in this figure, a method that estimates age-specific proportions provides a closer approximation to the truth than a method that only uses the population average. Thus, using five age groups is better than using only two age groups, which is in turn better than using only one, the entire population. In our case, we use single year of age, broken out by race groups, broken out by sex and broken out by Hispanic origin. (The technical appendix describes how we make this multiplication cell-by-cell to obtain a more finely disaggregated, and more precise, estimate.)



Step 2: How the CPS survey data are used: Next, we combine three years of CPS survey results and calculate the percent uninsured in each ARSH category (total number of uninsured in that category divided by the total estimated population in that category). This creates a matrix of "cells" that has a unique percent uninsured in each category.

For example, suppose that 3% of all 5-year-old Hispanic males are uninsured according to the CPS survey results. This forms our estimate of the percent uninsured for that specific category. Next we multiply each of the categories by the number of people estimated in each category from the ARSH tables. Now, if there are 100 5-year-old Hispanic males; then multiplying the percent uninsured according to the CPS, three of them are uninsured (.03 x 100 = 3)

It must be noted that some "cells" have very few people in them thus the estimate of proportion of uninsured has a very broad band, plus or minus, in which the "true" percent could fall. In order to make our estimates more reliable, we "collapse" cells (add them together) in reasonable categories to insure that the estimated percent uninsured is reliable – has a narrower "plus-or-minus" band. We used the criteria of approximately 50 persons per cell based on conversations with Census Bureau statisticians who work on these data (Moore and Seigel, personal communication, 2000). This required the "collapsing" of some cells into broader categories. For the record, with a 50 persons per cell size rule, the confidence intervals around the estimate of the proportion uninsured are substantial. The following table calculates the approximate confidence intervals for estimated proportions ranging from 10% to 35% and obtains the following 90% confidence intervals around the estimates.

Approximate confidence intervals associated with a sample size of 50 cases per cell	
P=	90% chance that true value is within the interval:
.10	.03 to .17
.15	.07 to .23
.20	.11 to .29
.25	.15 to .35
.30	.19 to .41
.35	.24 to .46

What these confidence intervals mean is that, as with any estimate, there is some uncertainty associated with our estimate of the proportion uninsured in each category. This fact is true for any estimation method known to man.

We repeat this process for the poverty categories, collapsing across categories when cell size is too small to provide adequate statistical reliability. Our collapsing scheme is based on information from other studies, from the total US CPS estimates, and from judgment. For example, it was found in the Urban Institute study of the US that males were more likely to be uninsured thus we did **not** collapse by sex; the same was true for

Hispanics versus non-Hispanics. However, we did collapse the races black and other into white due to the fact that they make up a relatively small proportion of the population and because we did not observe a substantial difference in uninsurance rates. We also collapsed age categories. For example, based on tabulations of CPS data, because of the large Medicare coverage, there is little variation in insurance rates for those over 65—almost all uninsurance rates are close to zero. Therefore, collapsing ages from 64 to 100 does not present any problems.

Note that it would be ideal to have a larger sample size for the state and stratification at a lower level of geography so that collapsing of cells would not be necessary. Collapsing creates “averages” of uninsured across the cells that are “summed” together. However, this is the more defensible method due to the large variance (plus or minus) to the disaggregated estimates, which would result from the uncollapsed cells.

To summarize, this method satisfies the goals of reliability, replicability, feasibility and defensibility. In addition, as we have noted above, while the sample size is adequate for the state as a whole, it is not adequate for individual counties. It is by using the ARSH estimates that we are able to “customize” state data for each county. We believe that this “customization” provides a finer-grained estimate of uninsured that takes into account the variation in composition in counties. When summed, it should make the state total more representative of the actual number.

Tests: Comparisons of These Estimates With Other Estimates

A natural test of these estimates is to determine how well they match up with other data series. A particularly useful test is to use this method to generate an estimate of the state population enrolled in Medicaid (using ARSH and CPS data), and compare that estimate to counts of Medicaid enrollees. A second test performed here is to determine whether the estimates generated by this method at the state level are consistent with Census Bureau estimates at the state level. However, we were not able to obtain unduplicated counts of state Medicaid enrollees in time for this comparison.

Given that unduplicated Medicaid enrollees were not available to us at the time of preparing this document, our test of these methods is to compare the estimated counts of uninsured against the numbers estimated by other methods. Unfortunately, not all methods use consistent definitions—for example some provide estimates only for adults[6] while others provide estimates by special categories. Where possible, we have tabulated our estimates to be consistent with the comparable definitions.

We begin with the Behavioral Risk Factors Surveillance Survey (BRFSS). This survey is an ongoing, monthly telephone survey, which is designed to assess a variety of health risks. Unfortunately, the BRFSS survey only provides information on adults in the household; thus we compare only our adult numbers with BRFSS adult numbers. BRFSS documentation states that because the survey is conducted by telephone and because not

[6] The Behavioral Risk Factors Surveillance Survey, for example.

having a telephone likely correlates with being uninsured, the proportion of uninsured people estimated by BRFSS may be underestimated. Differences in the survey questions used to ascertain health insurance coverage also are different for the CPS and the BRFSS. For the 15% uninsured rate reported below the question answered was "During the past 12 months, was there any time that you did not have any health insurance or coverage?"

Synthetic estimates of adults not covered by health insurance versus those reported by the Behavioral Risk Factors Surveillance Survey.		
	Number of adults who report not having health insurance	Percent of all adults In Nevada
Behavioral Risk Factors Surveillance Survey, 1998	204,996[7]	15%[8]
Number of persons 18 and over who do not have health insurance, estimated via synthetic methods	228,654	16.7%

We also compare our estimates with those generated at the state level by the Annie E. Casey Foundation in their "Kids Count" publication (1999). Note that the Annie E. Casey Foundation estimates are based on children aged 0-17, not 0-18 years of age. Further, the Annie E. Casey Foundation estimates use a five-year average of percentage uninsured, rather than a three-year average as we do. Because some evidence suggests that uninsurance rates have been increasing, this small methodological difference results in our estimate exceeding the Annie E. Casey estimate.

Synthetic estimates of children (0-17) without health insurance versus state estimates developed by the Annie E. Casey foundation		
	Number of children (0-17) without health insurance	Percent of all children (0-17) In Nevada
Annie E. Casey foundation estimates, 1998	77,039	19%
Estimated via synthetic methods, 1998	100,900	19.7%

Finally, we compare these estimates with those of the U.S. Census Bureau, itself. The Census Bureau also provides a total number of persons, so we can use Census Bureau numbers to estimate both the total number of persons without health insurance and the

[7] The estimate was formed by multiplying 1998 percentage reported on Centers for Disease Control (CDC) web pages (15%) by Nevada State Demographer's 1998 estimate of population 18 and older (1,366,643).
[8] The estimate of 15% reported on the CDC web pages differs from the one reported in a recent Nevada BRFSS report. The percentage uninsured reported there was 11.8%. At this writing we have not yet discovered why these numbers differ.

percentage of total population. Again, there is a difference that may be attributed in large part to the apparent upward trend in uninsured rates. The synthetic method is based on the three year average of uninsured from 1996-1998 and this smoothes the upward trend seen in the CPS uninsured rates for Nevada, giving a lower estimate for the synthetic method than for the state-wide Census Bureau estimate that uses only 1998 data.[9]

Synthetic estimates of persons without health insurance versus U.S. Census Bureau estimates.		
	Number of persons without health insurance	Percent of all persons In Nevada
U.S. Census Bureau estimates, 1998	394,000	21.2%
Estimated via synthetic methods, 1998	329,490	17.8%

As can be seen, in each case the estimates generated by our Synthetic Estimates System are comparable to their counterparts from other sources. One notable feature of our Synthetic Estimates System is that it is flexible enough to be conformed (approximately) to the definitions of the other data sources. This is a distinct advantage of such a system. Furthermore, at higher levels of geography it generates estimates comparable to other methods, and has the further advantage that it can be applied to lower levels of geography. This flexibility and accuracy are its two greatest strengths.

[9] Census Bureau estimates of uninsured in Nevada for 1996, 1997 and 1998 respectively were 15.6%, 17.5% and 21.2%.

References

Bureau of Labor Statistics (1999). "CPS: Annual Demographic Survey: March Supplement: Sampling and Weighting." http://www.bls.census.gov/cps/ads/1998/ssampwgt.htm, March 20, 1999.

Annie E. Casey Foundation. Kids Count (1999). Number of Children by Health Insurance Status (1994-1998 CPS, 5-year averages). http://www.aecf.org/kidscount/kc1999/data/kc99bins.xls

Cartland, Jenifer, and Beth Yudkowsky. (1993). "State Estimates of Uninsured Children", Health Affairs 12 (1): 144-151.

Centers for Disease Control and Prevention. 1999. "Behavioral Risk Factor Surveillance System: Prevalence Data, Nevada – 1998, Health Care Access" http://www2.cdc.gov/nccdphp/brfss/display.asp?cat=HC&yr=1998&qkey=7335&state=NV

Dillman, Don (1978). Mail and Telephone Surveys: The Total Design Method. New York, Wiley.

Employee Benefit Research Institute (1999). Sources of Health Insurance and Characteristics of the Uninsured: Mountain States, 1997. http://www.ebri.org/facts/0899fact2.htm

Gonzalez, Maria Elena (1973). Use and evaluation of synthetic estimates. In American Statistical Assocation, Proceedings of the Social Statistics Section. Washington D.C.: American Statistical Association.

Gonzalez, Maria Elena, and Waksberg, Joseph (1973). Estimation of the error of synthetic estimates. Paper presented at the first annual meeting of the International Association of Survey Statisticians, Vienna, Austria, August 18-25, 1973.

Gonzalez, Maria Elena, and Hoza, Christine (1978). Small-area estimation with application to unemployment and housing estimates. Journal of the American Statistical Association, 73:7-15.

Guyer, Jocelyn and Cindy Mann. (1999). "Employed But Not Insured: A State-by-State Analysis of the Number of Low-Income Working Parents Who Lack Health Insurance. Center on Budget and Policy Priorities. http://www.cbpp.org/2-9-99mcaid.htm

Judson, Dean H., and Ramirez, Brad (1996, 1997, 1998). Nevada County Population Estimates by Age, Race, Sex, and Hispanic Origin: 1995 (1996, 1997) Estimates, Projections, and Review of Methods. Reno, NV: Nevada State Demographer's Office.

Judson, Dean H. (1994, 1995, 1996). Nevada Population Estimation Method: Annual Estimates and Review of Methodology. Reno, NV: Nevada State Demographer's Office.

Judson, Dean H. (1995, 1996, 1997). Nevada Population Projections and Forecasts: 1995-2015 (1996-2016, 1997-2017) and Review of Methods. Reno, NV: Nevada State Demographer's Office.

Judson, D. H. (1999). Unpublished memorandum to Marla McDade of the Legislative Counsel Bureau, Carson City, NV.

Judson, D.H., and Popoff, C.L. (1998). Research Use of Administrative Records. Monograph manuscript in process.

Levy, Paul S. (1971). The use of mortality data in evaluating synthetic estimates. In American Statistical Assocation, Proceedings of the Social Statistics Section. Washington D.C.: American Statistical Association.

Liska, David, Niall Brennan and Brian Bruen. 1998. State Level Databook on Health Care Access and Financing. Washington, D.C., Urban Institute Press.

National Center for Health Statistics (1977). Synthetic Estimation of State Health Characteristics Based on the Health Interview Survey. Data Evaluation and Methods Research, Vital and Health Statistics Series 2, No. 75. Hyattsville, MD: U.S. Department of Health, Education and Welfare.

National Center for Chronic Disease Prevention and Health Promotion. 1999. "Behavioral Risk Factor Surveillance System: About the BRFSS" http://www.cdc.gov/nccdphp/brfss/about.htm

National Center for Health Statistics (1979). Small Area Estimation: An Empirical Comparison of Conventional and Synthetic Estimators for States. Data Evaluation and Methods Research, Vital and Health Statistics Series 2, No. 82. Hyattsville, MD: U.S. Department of Health, Education and Welfare.

Pencek, Bruce, Daneshvary, Rennae, and Schwer, R. Keith (1998). Health Insurance Coverage of Nevadans, 1997. Las Vegas, NV: CBER.

Reder, Stephen (1994). Synthetic Estimates of NALS Literacy Proficiencies from 1990 Census Microdata. Portland, OR: Northwest Regional Educational Laboratory.

Reder, Stephen (1997). Synthetic Estimates of Literacy Proficiencies for Small Census Areas. Available: http://www.casas.org/lit/litdata/reder.pdf, March 20, 1999.

Schaible, Wesley L., Brock, Dwight B., and Schnack, George A. (1977). An empirical comparison of the simple inflation, synthetic and composite estimators for small area statistics. In American Statistical Assocation, Proceedings of the Social Statistics Section. Washington D.C.: American Statistical Association.

Sigmund, C.L., Popoff, C.L., and Judson, D.H. (1999). A system for synthetic estimates of health-related characteristics: Linking a population survey with local data. Paper presented at the 1999 Population Association of America meetings, New York, NY.

Tanur, Judith (1981). Advances in methods for large scale surveys and experiments. In: National Science Foundation (Ed.), The Five-year Outlook on Science and Technology 1981. Washington, D.C.: Government Printing Office.

United States Census Bureau. (1999). “Health Insurance Coverage: 1998 – State Uninsured Rates” http://www.census.gov/hhes/hlthins/hlthin98/3yr98.html.

Technical Appendix: ARSH Estimates

The cohort-component method is based on the demographic balancing equation:

$$P_{t+1} = P_t + B_t - D_t + NM_t$$

In this equation, P_{t+1} is the population the target year, P_t is the population the previous year, B_t is the births to that group in the year, D_t is the deaths occurring to that group in the year, NM_t is the (net) number of migrants (in migrants minus out migrants) of that age.

This basic model is used to estimate the population in each age, sex, race and Hispanic origin (ARSH) group. A series of balancing equations are developed for each age-sex-race-Hispanic origin specific cohort. For example, population change for a specific sex and race cohort aged 45-49 in 1991 can be expressed as:

$$P^{45-49}{}_{t+1} = P^{45-49}{}_t - D^{45-49}{}_t + NM^{45-49}{}_t$$

There is no births term because no persons are born into this group between 1990 and 1991.

Some of the terms in the equations are known, others are assumed, and the remainder are solved. Among the known terms are population size for the base period, and deaths that have occurred during the period. The terms that are assumed are the in and outmigration rates for the period. Finally, the term that is solved for is the population size for the estimate year.

After an estimate for each age, sex, race, and Hispanic origin group is prepared, the totals are summed and compared to the independently-derived county total. (See Judson, 1995-1999 for details on how independently-derived county totals are developed.) If the ARSH total is higher than the official total, the ARSH numbers are reduced equally among all groups until the totals match. If the ARSH total is lower than the official total, the ARSH numbers are increased equally among all groups until the totals match. Implicitly, this method assumes that any differences between the ARSH totals and the official totals are caused by differentials in net migration; having no specific information on how those differentials are distributed, we assume that the differentials are distributed evenly among all groups.

The task list that follows represents the steps we went to determine the terms of the model.

1. Gather and Enter Primary Data:
    - Gather and enter data on 1980 and 1990 Census figures by age, race, sex, and Hispanic origin, by county.

- Gather and enter Nevada-specific historical data on vital events (births, deaths).

2. Perform Preliminary Data Analysis:
   - Estimate a Nevada-specific life table by age and sex, using 1990 Census data and 1990 death data.
   - Estimate Nevada-specific survival rates and survival probabilities by age and sex.

3. Develop Model:
   - Begin with 1980 ARSH data for Nevada.
   - Apply survival probabilities to each age, sex, race, and Hispanic origin group from the 1980 ARSH file.
   - Apply fertility probabilities to females of childbearing age to generate births by each ARSH group.
   - Each year, control fertility to match state and county totals from vital statistics.
   - Iterate this process forward until 1990 is reached.
   - Compare projected 1990 population by ARSH group with 1990 actual population from the 1990 Modified Age-Race-Sex-Hispanic Origin file.
   - Estimate ARSH group specific net migration rates for Nevada.

4. Apply Model to 1991-1995 Estimates.
   - Apply model to Nevada official estimates for 1980-1999.
   - Control ARSH group totals to 1999 official Nevada county estimates.
   - Obtain Nevada official 1999 estimates.
   - Apply model to Nevada official 1999 estimates by county.

In order to generate group forecasts using the cohort-component model, we must estimate survival probabilities, fertility probabilities, and net migration rates and apply those estimates to our 1990 base population. However, because many of the cell sizes are very small or zero (for example, the number of black Hispanic males aged 63 and living in Elko county), it is inappropriate and inadvisable to try to make estimates for each specific group. Therefore, we have collapsed our tables for estimation purposes only. We began with an overall collapse. For each of the probabilities and rates, we estimate it for the state level only: We do not attempt to estimate county specific mortality rates, for example.

Below we've listed the groups for which we calculated specific rates. In each case, in order to minimize distortion, we attempted to collapse on the categories that were least relevant to the rate in question. For example, it is known that mortality (survival) differs by sex. Thus, we did not collapse on sex. The list below illustrates:

Survival probabilities:
Collapsed across race and Hispanic origin; maintained distinctions for age (in five-year groups) and sex.

Fertility probabilities:
Collapsed ages into five-year groups; maintained distinctions for race and Hispanic origin.

Net migration:
Collapsed ages into five-year groups; collapsed across sex; maintained white/non-white and Hispanic/non-Hispanic distinctions.

After we estimate survival probabilities, fertility probabilities, and net migration, we apply the appropriate probability to the base population in 1990. The base population in 1990 is fully disaggregated by county of residence, single year of age, race, Hispanic origin, and sex. We maintain these distinctions through the estimation and projection period.

For presentation purposes we only show counts for age groups in 5 year intervals for ages 21+; however, detailed age estimates and projections are "underneath" the summaries. (More details on the estimation method can be found in Judson and Ramirez, 1996-1998.)

Technical Appendix: Synthetic Estimation Theory

The basic synthetic estimation equation we wish to use is:
$\hat{x}_{a,r,s,h} = P_{a,r,s,h} \cdot \hat{\mu}_{a,r,s,h}$.
Where,
$a \in \{0,...,85+\}$
$r \in \{W, B, API, AI\}$
$s \in \{M, F\}$
$h \in \{H, \sim H\}$
and
$P_{a,r,s,h}$ = the number of persons of age a, race r, sex s, and ethnicity h;
$\hat{\mu}_{a,r,s,h}$ = the proportion of persons of age a, race r, sex s, and ethnicity h that have the health-related characteristic of interest (in this case, who are uninsured), $\hat{\mu}_{a,r,s,h} \in [0,1]$;
$\hat{x}_{a,r,s,h}$ = the number of persons of age a, race r, sex s, and ethnicity h that have the health related characteristic of interest (in this case, who are uninsured).

Thus, we use our known age/race/sex/hispanic number, and an estimated group-specific proportion who are uninsured, to estimate the number of uninsured persons of the specific group.

The simplest form of a synthetic estimate is, of course, to simply multiply an overall population by an overall proportion to get an overall number. We introduce the "dot" notation to describe this method in this framework. For any variable $y_{a,r,s,h}$, we define:

$$y_{a,\bullet,s,h} = \sum_{r} y_{a,r,s,h} .$$

We define other sums similary if a, s, or h are "dotted," that is, when we "dot" a subscript, it merely indicates that we sum over all elements of that subscript, or, using more common language, we "collapse" that margin.

Using $P_{a,r,s,h}$ as an example, the total population of an area is equivalent to "collapsing" all margins, or:

$$\text{Total Population} = P_{\bullet,\bullet,\bullet,\bullet} = \sum_{a}\sum_{r}\sum_{s}\sum_{h} P_{a,r,s,h} .$$

If we wish to multiply the total population by some overall uninsurance proportion in the population, we would express this notion as:

$$\text{Total Pop. Uninsured} = \hat{x}_{\bullet,\bullet,\bullet,\bullet} = P_{\bullet,\bullet,\bullet,\bullet} \cdot \hat{\mu}_{\bullet,\bullet,\bullet,\bullet} .$$

However, instead of performing this collapsing, we maintain distinctions between groups. Thus, the equation that describes the fully-disaggregated synthetic approach is the following:

$$\text{Total Pop. Uninsured} = \hat{x}_{\bullet,\bullet,\bullet,\bullet} = \sum_{a}\sum_{r}\sum_{s}\sum_{h} P_{a,r,s,h} \cdot \hat{\mu}_{a,r,s,h} .$$

We make this calculation for each county in Nevada separately. This equation implies that we estimate the number of uninsured in a **particular** age, race, sex, and hispanic origin group, then "add up" by summing over all groups.