Economic Research Initiative on the Uninsured
Working Paper Series

COUNTING AND CHARACTERIZING THE UNINSURED

Pamela Farley Short, Ph.D.
The Pennsylvania State University

ERIU Working Paper 2
www.umich.edu/~eriu/pdf/wp2.pdf

Economic Research Initiative on the Uninsured
University of Michigan
555 South Forest Street, 3rd Floor
Ann Arbor, MI 49104-2531

12/4/01

This work was performed under the sponsorship of the Regents of the University of Michigan and the Robert Wood Johnson Foundation. The author is grateful to Kathy Swartz, Alan Monheit, Anne Polivka, and Catherine McLaughlin for their comments on earlier drafts. Peter Cunningham (Center for Studying Health Systems Change), Diane Makuc (National Center for Health Statistics), Chuck Nelson (Bureau of the Census), Steve Zuckerman and Jennifer Haley (The Urban Institute), and Jeff Rhoades (Agency for Healthcare Research and Quality) were helpful in providing information about and from the surveys conducted by their organizations. Paul Fronstin was also helpful in providing additional, unpublished details regarding his estimates from the CPS.

Introduction

Researchers have been making estimates of the number and characteristics of Americans without health insurance for nearly 50 years (Andersen and Anderson 1999). The Health Information Foundation, which later became the Center for Health Administration Studies (CHAS) at the University of Chicago, collaborated with the National Opinion Research Center (NORC) on the first nationally representative household survey of health care use, expenditures, and insurance in 1953. Over the next two decades, CHAS and NORC conducted three more national surveys on these topics.

In 1964, the National Center for Health Statistics (NCHS) published the first national estimates of health insurance (Hoffman 1964) from the National Health Interview Survey (NHIS). Ten years later, NCHS began planning for a panel household survey with short recall periods to capture reliable reports of health expenditures and utilization from household respondents over a calendar year (Andersen and Anderson 1999). These plans resulted in the fielding of the National Medical Care Expenditure Survey (NMCES) in 1977, a joint effort of NCHS and its sister agency in the U.S. Department of Health and Human Services, the National Center for Health Services Research (NCHSR).

NCHSR, now known as the Agency for Healthcare Research and Quality (AHRQ), supported the last of the CHAS-NORC surveys by contract in 1971 and shared an interest in insurance and expenditure data. With six interviews to collect data from each sampled household for the calendar year, NMCES allowed analysts at NCHSR to publish the first estimates of changes in insurance for individuals over a year. NMCES revealed that nearly as many people were uninsured for part of 1977 as for the entire year, and that more than two-fifths of Medicaid recipients moved on or off Medicaid in a year (Walden, Wilensky and Kasper 1985).

The National Center for Health Statistics still produces national estimates of the uninsured from the National Health Interview Survey (NHIS), now on an annual basis after fielding health insurance questions irregularly until 1989. With the Medical Expenditure Panel Survey (MEPS), the Agency for Healthcare Quality and Research continues the tradition of collecting longitudinal health insurance data begun as NCHSR. AHRQ has fielded a new, two-year MEPS panel in every year since 1996.

However, now there are at least six other ongoing national surveys that also produce information about the number and characteristics of Americans who are uninsured. In addition to the NHIS and MEPS, there are four national surveys that collect data on an ongoing basis about the uninsured of all ages:

- Current Population Survey (CPS) is a monthly survey conducted by the Census Bureau that has asked about health insurance in every March interview since 1980 (except 1981). Because of its long time series, timeliness, and a sample size that is large enough to make state-level estimates, the CPS is the most frequently cited source of statistics regarding the uninsured.
- Survey of Income and Program Participation (SIPP) is a multi-year panel survey that has been fielded somewhat irregularly by the Census Bureau since 1983. Data are now available from a four-year panel that was fielded in 1996.
- Community Tracking Study (CTS) began collecting data every two years in 1996-97, primarily by telephone, from nationally representative samples of households

that are concentrated in 60 communities across the country. The CTS is conducted by the Center for Studying Health System Change with funding from the Robert Wood Johnson Foundation.

- National Survey of America's Families (NSAF) also collects data on a two-year cycle, primarily by telephone. Fielded in 1997 and 1999, NSAF is part of the Urban Institute's "Assessing the New Federalism" project and is supported by funding from a number of foundations. The household sample is representative of the national population under age 65, but is concentrated in 13 states.

Two other surveys are limited to adults:

- Behavioral Risk Factor Surveillance Survey (BRFSS) interviews more than 150,000 adults each year by telephone. The BRFSS is conducted by the health departments of all the states and territories, with the support and guidance of the Behavioral Surveillance Branch of the Centers for Disease Control.
- Health and Retirement Survey (HRS) is a panel survey that follows the aging of cohorts of older Americans, starting at age 51. Interviews are conducted every two years and began in 1992 with a cohort born between 1931 and 1941. HRS is conducted by the Institute for Social Research at the University of Michigan, with funding primarily from the National Institute on Aging.

Detailed and up-to-date descriptions of all of these surveys, except for the HRS, are available from Fronstin (2000b) and from Lewis, Ellwood and Czajka (1998).[1] Both reports compare and contrast methodologies and estimates of the uninsured across surveys. As shown in Table 1, adapted from Fronstin, such comparisons show that there is considerable variation in estimates of the total number of Americans who lack health insurance from different surveys—even for the same year.

Where these earlier reports have focused on quantifying and reconciling the effects of methodological differences on the estimates from different surveys, this article attempts to provide readers with a more global and fundamental understanding of the conceptual and methodological issues involved in counting, characterizing, and studying the uninsured. In particular, it explains how most of the important issues can be traced back to two basic observations about the uninsured. The first observation is that the uninsured are a residual group by definition. They are the people who fall in the cracks left by public and private insurance programs (Figure 1). As a result, one cannot produce or make sense of statistics about the *uninsured* without first producing or making sense of statistics about the *insured.*

The second observation is a good illustration of the first. Namely, with relatively large numbers of Americans flowing in and out of Medicaid and employer insurance over short periods of time, people also move in and out of being uninsured at a fairly rapid rate. Because the insurance status of individuals changes noticeably over time, time is a very important consideration in counting and characterizing the uninsured.

The far-reaching implications of these two apparently simple observations are discussed in the next two sections of this article. Each section begins with conceptual issues related to one of the observations and then turns to its methodological and empirical implications. The next section draws on this framework to assess the relative strengths and weaknesses of different health insurance surveys. The last section offers suggestions regarding future directions for both data collection and analyses aimed at counting and characterizing the uninsured.

Observation #1: The uninsured are a residual.

Conceptual issues

What is health insurance? Because the uninsured are people who do not have health insurance, the first step in defining "uninsured" is to define "health insurance." Most surveys define health insurance as a list of different sources or types of coverage. Respondents are asked (separately) about each of the items in this list. Those without any of the listed types of coverage are considered not to have health insurance and are counted as uninsured. In order to achieve a reasonable level of reliability across respondents, surveys rarely begin by asking people if they are uninsured without implicitly defining health insurance with such a list.

For example, the list presented in the Current Population Survey defines health insurance as

- A health plan provided through [a household member's] current or former employer or union [as policyholder or dependent]
- A plan purchased directly from an insurance company [as policyholder or dependent]
- A health plan of someone who does not live in the household
- Medicare, the health insurance for person 65 years old and over or persons with disabilities

- Medicaid or [name(s) of Medicaid programs in respondent's state], the government assistance program that pays for health care
- TRICARE, CHAMPUS, CHAMPVA, VA, or military health care
- A health insurance plan such as [state-specific health insurance program for lo-income uninsured individuals] or any other type of plan

Historically, the Census Bureau also counted people covered only by the Indian Health Service (IHS) as insured, and included them with Medicaid recipients in published statistics from the CPS. Since 1998, the Bureau counts them as uninsured (Fronstin 2000b).

CPS interviewers are trained to exclude specialty plans that cover only dental, vision, or prescription expenses from the definition of health insurance. In its survey questions, MEPS specifically limits the definition of health insurance to plans that cover hospital or physician services (thereby excluding specialty plans). MEPS does not consider that either extra cash plans, which are sometimes linked to specific diseases, or the services provided directly by the Veterans Administration to selected veterans are health insurance.

No recent survey collects enough detailed information about health plans to quantify the effect of differences in the definition of health insurance on national estimates of the uninsured, but it is not likely that definitional differences have much empirical effect. Short and Vistnes (1992) found that about 1-2% of the elderly who reported private insurance in 1987 had only extra-cash or disease-specific coverage. Fronstin (2000b) estimates that the CPS change involving recipients of IHS services affected about 300,000 people and increased the estimated percentage of the population without health insurance from 18.1% to 18.3% in 1997.

Methodological implications

Because the uninsured are always identified as the residual group of survey respondents who do not answer positively to questions about different types of coverage, forgetfulness and underreporting are likely to inflate survey estimates of the uninsured. All of the coverage that respondents forget or do not recognize from the wording used in a survey shows up in higher counts of the uninsured. There are no national benchmarks for cross-checking survey estimates of the total count of people with private insurance, but administrative records for such public programs as Medicare or Medicaid can be used to evaluate the accuracy of survey-based counts of enrollment in these programs.

Administrative records also have limitations in terms of reliability and validity. Nevertheless, comparisons to Medicaid administrative data strongly suggest that Medicaid enrollment is underreported in surveys, implying that they overestimate the number of uninsured.[2] In the past, many data collection organizations edited survey responses to correct Medicaid underreporting, particularly by attributing Medicaid to welfare and SSI recipients (who automatically qualified for Medicaid before federal welfare reforms were instituted in 1996). Now, because Medicaid eligibility is no longer linked to welfare recipiency and has been expanded to many low-income individuals (especially children) who are not eligible for welfare, it is more difficult to identify likely Medicaid recipients from the responses to other survey questions. Because of the emphasis on health expenditures in MEPS, analysts using that survey can crosscheck and reconcile health insurance status with the sources of third-party payments that are reported. The Urban Institute customarily adjusts the CPS micro data, using the Institute's micro simulation model, to match administrative Medicaid enrollment by age and disability status (Lewis, Ellwood, and Czajka 1998).

To guard against underreporting, the majority of surveys (including CPS, NHIS, CTS, NSAF) explicitly ask respondents to confirm that they are uninsured after responding negatively to the list of questions about coverage types and sources. In addition, several surveys (including SIPP, MEPS, and CTS) ask a catchall question that allows respondents to report insurance that they did not associate with any of the types or sources mentioned by the interviewer. The use of a confirming question reduces the number of uninsured by 20% in the National Survey of America's Families, a telephone survey without a catch-all question, compared to 6% in the Community Tracking Study, a telephone survey with a catch-all question (Rajan, Zuckerman, and Brennan 2000). Reportedly, the confirming question in the NHIS has little effect in the context of the longer, more intensive, in-person interviews conducted for that survey (Rajan, Zuckerman, and Brennan 2000; note 6). Nelson and Mills (2001) reported recently that the addition of a confirming question to the CPS, to be incorporated in official estimates for the first time in 2000, will reduce CPS estimates of the uninsured by about 8%.

The fact that being uninsured is residually determined by the absence of public and private insurance has other important methodological implications. Because private insurance typically covers all members of the policyholder's immediate family, it is very important to adopt a household or family perspective in studying the uninsured.[3] For example, it is more meaningful to relate a person's insurance or lack of insurance to the presence or absence of a worker in the family than to whether or not the person (who may be the child or the nonworking spouse of a working adult) is employed.

To identify the set of people whose health insurance status “goes together,” health insurance analysts have invented and widely adopted an entirely new, insurance-related concept of a family unit. The Census Bureau’s official definition of “family” includes all of the people related by blood or marriage who are living at the same address; this is the definition that is used in aggregating and adjusting family income in official poverty estimates. The concept of a “health insurance unit” is limited to the people who would typically be covered by a family health insurance policy. At least one national survey, the Community Tracking Study, specifically samples and collects data for health insurance units instead of households or families. That sampling approach narrows the scope of data collection, by collecting data for fewer people, while it preserves a picture of important coverage connections within families.

In defining health insurance units, adults are grouped with their spouses and their dependent children, where dependent children are either high-school age and younger or college age and full-time students. A 15-year-old male and his mother constitute one family and one health insurance unit. A 50-year-old male who lives with his elderly mother are one family, but two different health insurance units. Unlike the teenager and his mother, the 50 year-old and his mother are unlikely to have insurance from the same source, and few programs or proposals for covering the uninsured would combine the income of the two adults in determining either’s eligibility for government assistance.

When income is aggregated over health insurance units, instead of families, significantly more of the uninsured are categorized as poor (Long and Marquis 1996).

The distinction is particularly important for young adults (Short 2000), the age group with the highest percentage uninsured. Using health insurance units to define poverty assigns young adults who live with their parents and are not in a school to a different health insurance unit than their parents, so that the parents' income is not considered in measuring the economic well being of the young adults.

Another important consideration in collecting and analyzing data about the uninsured, given the residual nature of that group, is each person's eligibility for public or private insurance. Because Medicare's nearly universal coverage of the population 65 and older means that virtually no one in that age group is uninsured, most statistics concerning the uninsured are limited to the population under 65. Now that Medicaid changes and the State Children's Health Insurance Program (SCHIP) have greatly expanded the coverage of children compared to adults, it is equally important to distinguish between adults and children in studying the uninsured.

Employment is importantly related to eligibility for insurance through employers and unions, the source of coverage for 80% of insured Americans under age 65 (Fronstin 2000a). Consequently, the most useful surveys for studying the uninsured ask lots of questions about employment status and job characteristics, in addition to asking about health insurance. The most useful surveys also provide employment data for all adults in a family or health insurance unit, because of the potential relevance in explaining the coverage (or lack of coverage) of other family members.

Given the number of public programs and legislative proposals that target eligibility for insurance by income, survey questions that are well designed to measure income are also important for health insurance analyses.

Still another key piece of information is whether any family members were eligible for health insurance through an employer or union, regardless of whether or not they actually enrolled. According to estimates from the Community Tracking Study (Cunningham, Schaefer, and Hogan 1999), all but 14% of people with access to insurance through their own or a family member's employment chose to enroll. Of that group, 9% had other coverage and only 5% were uninsured.[4] However, the uninsured who turned down insurance from a family member's employer accounted for about 20% of the total number of uninsured.

There are no comparable survey questions that can reliably establish eligibility for Medicaid or SCHIP. However, if a survey collects all of the necessary data elements (and does not suppress the respondent's state of residence because of confidentiality concerns), analysts can simulate Medicaid/SCHIP eligibility according to the rules in the respondent's state. This line of research has established that Medicaid eligibility is a good, but hardly perfect predictor of enrollment. Indeed, the discovery that a significant number of Medicaid-eligible children were uninsured in the mid-1990's (Selden, Banthin and Cohen 1998; Dubay and Kenney 1996; Sumner, Parrott and Mann 1997) touched off a concerted effort in Medicaid and SCHIP to increase participation rates among eligible children by greatly expanding outreach efforts and simplifying application procedures.

Finally, the residual nature of the uninsured also has implications for analyses designed to identify individual characteristics and other factors that contribute to the probability of being uninsured. It is hard to make much sense of the effect of these variables by estimating binomial multivariate models where "insured" and "uninsured" are the two outcomes. Consider, for example, the relationship between income and the probability of being uninsured in a binomial model. This relationship is surely not monotonic and is quite complicated, because low income is associated with Medicaid/SCHIP enrollment and high income is associated with enrollment in employment-related insurance. As a consequence, both high and low levels of income are associated with lower probabilities of being uninsured.

Multinomial models (for example, multinomial logit models with “private insurance,” “Medicaid/SCHIP”, and “uninsured” as outcomes) are likely to be more revealing and robust than binomial models in studying the uninsured.[5] By using multinomial models, analysts can specify separate, structural determinants of the probabilities of public and private insurance in the population under age 65—and then calculate the probability of being uninsured as the residual that it is. Age, income, and other eligibility criteria that often vary by state are the primary determinants of enrollment in Medicaid and SCHIP. The employment status and job characteristics of adult family members are the primary determinants of private enrollment.

Related empirical findings

While different surveys often disagree on the number of people who are uninsured, they all show that the risk of being uninsured is highest in the population subgroups that one would predict on the basis of eligibility (or lack of eligibility) for coverage from employers, Medicare, and Medicaid/SCHIP. For example, because the rate of employer insurance increases dramatically with earnings and income, while Medicaid and SCHIP cover fewer than half of the nonelderly poor, the uninsured rate decreases dramatically with increasing income (Figure 2).

The likelihood of being uninsured is also closely related to employment. Full-time workers and their families are much less likely to be uninsured than other people under age 65. Among working adults, those who work part-time or for small companies, are self-employed, or have jobs outside of manufacturing and the public sector are more likely to be uninsured.

Given the connections between age and eligibility for both employer insurance and Medicaid/SCHIP, the uninsured rate also varies noticeably by age. With the expansion of coverage to low-income children through Medicaid and SCHIP, children under 18 are now less likely to be uninsured than adults. Young adults in their late teens to mid-twenties are the age group with the highest percent uninsured. Young adults do not qualify as dependents under their parents' plans if they leave school and are likely to have entry-level or short-term jobs that do not offer insurance benefits if they work.

The uninsured rate also varies by socio-demographic characteristics that are correlated with income and family employment, such as race and ethnicity or marital status. Racial and ethnic minorities, as well as non-citizens, are more likely to be uninsured. Single individuals or people in families headed by single adults are more likely to be uninsured than people in families headed by married couples.

There are also noticeable geographic differences in the uninsured rate, which is generally higher in the South than elsewhere. A recent analysis using the Community Tracking Study (Cunningham and Ginsberg 2001) suggests that about a third of the difference in uninsured rates across communities is due to difference in racial/ethnic composition and socioeconomic status; about a quarter is due to differences in employment characteristics; and about an eighth is due to differences in state Medicaid eligibility.

Two kinds of statistics are frequently used in characterizing the uninsured by population subgroup. The uninsured rate within a subgroup (i.e., the percentage of the subgroup that is uninsured) measures the risk or likelihood of being uninsured for people in that subgroup. The preceding discussion focused on variation in uninsured rates across subgroups, indicating that some groups are at much higher or lower risk of being uninsured. By contrast, the percent distribution of the uninsured across subgroups (i.e., the percentage of the uninsured who are in each subgroup) describes the composition of the uninsured population and its concentration in certain subgroups.

The composition of the uninsured population depends not only on the uninsured rate in each subgroup, but also on the absolute size of each subgroup. A low uninsured rate in a very large subgroup can account for a surprisingly large share of the uninsured. For example, although the uninsured rate is relatively low for workers and their families, more than 4 out of 5 of the uninsured are in families with a working adult (Fronstin 2000a; Monheit and Vistnes 1996).

By the same token, about half of the uninsured under age 65 are in families with incomes *below* 200% of the poverty line (Fronstin 2000a), where the uninsured rate is relatively high. With nearly three quarter of the entire population under age 65 in families with incomes *above* 200% of the poverty line, the higher income group accounts for the other half of the uninsured.

Observation #2: Insurance status changes over time.

Conceptual issues

What is the best way to count and characterize the uninsured when people experience changes in insurance status over time, as illustrated in Figure 3? The simplest approach is to focus on the uninsured at a cross-section in time, such as the start of the calendar year represented by the dotted line in Figure 3. However, cross-sectional estimates understate the number of people who are uninsured over time. For example, a cross-sectional estimate at the start of the first full calendar year in Figure 3 will count A, B, and D as uninsured, but will miss C and E who are uninsured later in the year. As illustrated by the figure, longitudinal estimates that capture all of the uninsured spells for a sample of people over time should always produce higher counts of the uninsured than cross-sectional estimates. Furthermore, counting the uninsured over longer time periods (for example, two years instead of one year) should always produce a higher count of the uninsured, as illustrated by the addition of F to the count for one year.

Cross-sectional estimates of the uninsured are useful for projecting the average caseload that would be served by a new program to cover the uninsured. Cross-sectional estimates are also useful for projecting program costs. By counting the uninsured who would be covered by a new program on a given day, the cross-sectional approach yields an estimate of covered person-days for a one-day accounting period (and corresponds to a program's caseload). If there are no important seasonalities, multiplying the one-day estimate by 365 produces a reasonable estimate of the total number of covered person-days in a year (usually the accounting period of interest). Dividing by 365 then converts total person-days to person-years. Counting each uninsured person in the one-day cross section as an uninsured person-year is mathematically equivalent to these two operations, so that is how the calculation is actually performed in practice. Total annual program costs can then be approximated by applying an estimate of average costs per person-year (for example, the annual premium for covering each uninsured person in the cross section) to the total number of uninsured person-years.[6,7]

With longitudinal data that track a person's insurance status over days or months for a year or more, analysts can explicitly count uninsured person-days or person-months over a year and then convert to person-years. Consequently, longitudinal data of this type is also useful for projecting program costs or the average caseload that a new program will serve. Estimates from longitudinal surveys that do not break down insured and uninsured days (or months) for people who were uninsured for part of the year are not useful for these purposes. For example, the number or percentage of people who were *ever* uninsured in a year overstates the number of uninsured person-years. The number or percentage of people who were *always* uninsured in a year misses the part-year uninsured and understates the number of uninsured person-years.

The major advantage of longitudinal data, compared to cross-sectional data, is in assessing the welfare implications of being uninsured. First, recalling Figure 3, the risk of being uninsured is considerably greater than cross-sectional statistics describing Americans who are uninsured at a point in time would imply. Second, being uninsured for a long time is likely to have a much bigger effect on a person's health and finances than being uninsured for a short time. After all, the likelihood of getting through a short uninsured spell without a health crisis is much greater than the likelihood of getting through a long spell. Also, the health effects of foregoing appropriate care because of its uninsured cost are likely to compound and accumulate over time. By definition, cross-sectional data about the uninsured do not distinguish between the short-term and long-term uninsured.

Although cross-sectional data do not allow analysts to distinguish the long-term uninsured from the short-term uninsured, one can demonstrate mathematically that the long-term uninsured are implicitly represented more heavily in cross-sectional data than in longitudinal data. This difference between the two measurement approaches is formally attributable to length-based sampling, a feature of cross-sectional data that has been recognized in collecting statistics about other dynamic phenomena such as unemployment or poverty. The heavier representation of long-spells in the cross-section can be seen intuitively in Figure 3, where A's long spell is a third of the uninsured spells captured in the cross-section (A, B, C) but only a fifth of the uninsured spells for the year (A, B, C, D, E). Because the short-term uninsured generally come from higher in the income distribution than the long-term uninsured (Swartz and McBride 1990, Short and Klerman 1998), cross-sectional estimates can be expected to show that more of the uninsured are poor or come from groups with lower socioeconomic status than longitudinal estimates.

As the preceding discussion illustrates, important conceptual questions regarding the best way to count and characterize the uninsured for a particular analytic purpose are raised by health insurance dynamics (Swartz 1994). Acknowledging the importance of health insurance dynamics also means introducing new concepts into one's thinking about health insurance. Important concepts in a dynamic framework include health insurance spells (time periods with a specific type of insurance, including none at all), the duration of insurance spells, and transitions from one type of insurance to another.

Methodological implications

Health insurance dynamics have important methodological implications for the collection of health insurance data and for the comparability (or lack of comparability) of estimates from different health insurance surveys. First, because insurance changes over time, surveys that ask about insurance status over different lengths of time will produce different estimates of the uninsured. For example, when the NHIS or CTS asks if a survey respondent currently has any of a list of different types of coverage, that is not equivalent to asking in the CPS if he or she had any of the same types of coverage in the last year. Negative responses to the first set of questions imply that the person was uninsured for at least a day; negative responses to the second set of questions imply that the person was uninsured for at least a year. More people fall into the first category than the second. For that reason, surveys like NHIS or CTS that ask about current coverage should count more people as uninsured than surveys with longer reference periods like the CPS (the last calendar year), SIPP (the last 4 months), and MEPS (since the beginning of the calendar year or the last interview).

Extending the reference period back in time makes it more difficult for people who have experienced a recent change in insurance status to answer the questions about their health insurance correctly. Because it is undoubtedly easier for such people to answer correctly about their coverage "now," rather than about their coverage in the past, most surveys (NHIS, CTS, NSAF, BRFSS, and HRS) ask about health insurance in this cross-sectional fashion. When asked about the past, respondents have a tendency to forget coverage in the reference period that they no longer have.

In order to reduce the errors associated with imperfect recall, while capturing monthly insurance status over relatively long periods of time, SIPP and MEPS are designed as longitudinal surveys that re-interview respondents every few months. Despite the relatively short reference period in SIPP, most of the changes in coverage that are reported in SIPP occur in the months corresponding to the interviews. This pattern (the so-called "seam problem" in SIPP) suggests that respondents tend to focus on their current insurance status when they are being interviewed, even when asked to think only a few months into the past.

Although the CPS is the most frequently cited source of data regarding the uninsured, it poses the most difficult recall task for respondents. Respondents to the CPS are asked in March if they had any of the listed types of coverage in the preceding calendar year. Consequently, respondents must remember back 14 months and then ignore the two months immediately preceding the interview. The resulting, retrospective estimates of the number of people who were uninsured throughout a calendar year are much larger than comparable estimates from longitudinal surveys like SIPP and MEPS that interview people several times during a year. Even with the introduction of a confirming question in the CPS, which lowered the estimate for 1999 from 42 million to 39 million (Nelson and Mills 2001), the CPS annual estimate remains high in relation to annual estimates from MEPS and SIPP.

Methodological issues involving time also arise in measuring correlations and associations of other variables with health insurance status. For example, analysts using such surveys as NHIS and CTS should be cautious about relating cross-sectional information on insurance status from those surveys to annual measures of health services use (Long and Marquis 1994). The effects on utilization of being uninsured are disguised under those circumstances, because some of the people who were uninsured in the cross-section were not necessarily uninsured when the utilization occurred. At the same time, some people who were uninsured for part of the year (and may have used fewer services during that time) were classified in the cross-section as insured.

Finally, because insurance status changes over time, it is not really enough to know someone's current insurance status. It is also important to know how long each person was insured or uninsured before the time of the survey. Otherwise, one can neither estimate the duration of insurance spells nor distinguish between short and long uninsured spells in terms of their welfare implications. Nor can one model eligibility rules that are based on recent insurance history, such as the portability rules in the Health Insurance Portability and Accountability Act or state rules that discourage people from dropping private insurance and applying for Medicaid. All of the surveys, except for the CPS, currently ask something about coverage (or lack of coverage) prior to the start of the survey, but these questions seem to get surprisingly little use by analysts.[8]

Without retrospective information, cross-sectional surveys do not allow analysts to distinguish between such situations as A and B in Figure 3. Even in a longitudinal survey that covers the entire time window depicted in Figure 3, the length of B's uninsured spell is undetermined if retrospective questions are not asked at the start of the survey.

In technical terms, insurance spells are both left- and right-censored at the start and end of any survey, as illustrated in Figure 3. Special techniques, such as life tables or survival models, which focus on the conditional probability of ending a spell in period

$t+1$ (after the spell has lasted t time periods), can be used to accommodate right censoring.[9] However, these techniques require unbiased information about the prior duration of spells, without left censoring.

Related empirical findings

Going as far back as the 1977 NMCES, longitudinal surveys consistently show that turnover in the uninsured population is empirically important. As Table 2 illustrates with longitudinal monthly data from the 1996 MEPS panel, the hard core of people who were uninsured throughout the two years of the survey comprised a little more than a quarter of the people who were ever uninsured over two years (23.5 million out of 80.2 million). Out of the average of 45 million who were uninsured in each month in 1996, 32 million were uninsured throughout the year. As demonstrated by the difference between the 32 million who were uninsured throughout 1996 and the 64 million who were ever uninsured in 1996, the number of people of who began or ended an uninsured spell during the year was about the same as the number who remained uninsured.

The relatively high turnover in the uninsured population is attributable to a large number of fairly short spells. Half of uninsured spells end within 5 or 6 months (Swartz, Marcotte, McBride 1993b; Bennefield 1996b). Nevertheless, the long-term uninsured are also important in terms of their numbers and policy significance (Swartz 1994). Because people with long spells accumulate in the uninsured population, while people with short spells replace each other, nearly three-quarters of the people in an uninsured cross-section have been uninsured for 6 months or more (Short and Klerman 1998).) More than 40%

have been uninsured for 18 months or more. Recall that such a cross-section approximates the caseload of a universal program that would insure all of the uninsured.

Studies that have examined the factors associated with the likelihood of ending an uninsured spell and transitioning into coverage (Swartz and McBride 1990; Swartz, Marcotte, and McBride 1993a; Short and Friedman 1998; Bennefield 1996b) suggest that people with higher socioeconomic status have shorter uninsured spells. Also, although young adults are more often uninsured than older adults, they have shorter uninsured spells (Swartz, Marcotte, and McBride 1993a; Short and Friedman 1998; Bennefield 1996b). Nonworkers, Hispanics, and high school dropouts have longer uninsured spells (Bennefield 1996b; Short and Friedman 1998). There is also strong evidence that the likelihood of regaining coverage declines as people remain uninsured for longer and longer periods of time (Swartz, Marcotte, and McBride 1993a; Short and Friedman 1998).

Alternative ways of counting the uninsured in MEPS are arranged in Table 2 so that the estimates increase from left to right across the table. As they increase, the estimates also give increasing weight to the short-term uninsured. The leftmost column counts only those people who were uninsured for 24 months and ignores anyone else who was ever uninsured. The rightmost column gives the same weight to people who were uninsured for as little as 1 month as those uninsured for 24 months. (Each person is counted once, regardless of the length of time that they were uninsured.) The center column weights each person according to the length of time that he or she was uninsured in 1996.

By looking across the table at the relative risk for each population subgroup, one can see how that group's risk of being uninsured varies in comparison to the lowest risk group when different time frames are used to count and characterize the uninsured. As expected, the relative risk of being uninsured generally declines from left to right across the table for lower socioeconomic groups, as the different time frames give less weight to the long-term uninsured. There is a noticeable exception to this pattern in the leftmost column, where the relative risk of being uninsured for all of 1996 and 1997 is lower for low income groups and African Americans than the relative risk (shown in the next column) of being uninsured for all of 1996. This is attributable to the underlying trends from 1996 to 1997 in the risk of being uninsured all year by subgroup (data not shown).

Comparative Strengths and Weaknesses of Different Surveys

The Census Bureau's dependability in releasing CPS data on a regular, timely basis over many years is one of the major reasons for the popularity of that survey. As shown in Table 3, micro data are released from the CPS more quickly than from any other survey. In addition, although the CPS health insurance questions have been redesigned on several occasions since 1980 (Swartz 1997; Fronstin 2000b), no other survey offers as long a time series for studying trends in health insurance. The sample size, which is large enough to make state-specific estimates, is another important advantage of the CPS. The rich employment and economic data, collected for all adults in the household, has been useful in many analyses. As the source of official estimates of poverty in the United States, the CPS also has good income measures.

Unfortunately, there is overwhelming evidence that many respondents answer the CPS insurance questions incorrectly. Because the CPS estimates are more similar to the cross-sectional estimates from other surveys, and because respondents apparently tend to report their current coverage when asked about the past, many analysts ignore the question wording and treat the estimates from the CPS as a one-month cross section (Swartz 1986). Analysts are happy to treat the CPS counts as cross-sectional, because cross-sectional estimates are more useful in costing out reform proposals than the all-year estimates of the uninsured that the CPS is designed to produce. Nevertheless, if some respondents interpret and answer the CPS questions correctly, then the CPS estimates are really an ill-defined amalgam of annual and cross-sectional insurance status (with a good measure of recall error thrown in).[10] Ironically, if the Census Bureau continues to improve the validity of the all-year CPS estimates by adding the new confirmation question and experimenting with other changes, analysts will have to stop interpreting the CPS data as a cross-section that can be used for cost estimates.[11]

Table 3 reveals that there are two other government surveys, NHIS and SIPP, that begin to approach the CPS in terms of sample size, but ask questions with shorter recall periods that respondents can answer more accurately. The NHIS questions about current insurance status are undoubtedly the most straightforward for respondents with recent changes in insurance status to answer. Recent improvements in the NHIS, which shifted the insurance questions to the core where they are asked each year and added more economic and employment questions, have made the NHIS more suitable for studying health insurance issues that have historically been investigated with the CPS. However, some important employment questions are not asked for all of the adults in the family;

estimates and public use files have not been released as quickly for the NHIS as the CPS; the sample is too small to make reliable estimates for all states; and the income questions are probably not as good as in the CPS.

The 1996 SIPP panel also approaches the sample size of the CPS, but is not quite large enough to make estimates for all states. The Census Bureau plans to maintain the larger panel size instituted in 1996, by fielding one large panel every three years instead of overlapping panels every year. With a fixed, four-month reference period, each wave of SIPP is a nationally representative cross-sectional survey that could be used to make quick estimates of health insurance status. Like the CPS, SIPP has the advantage of collecting a wealth of economic and employment data. Furthermore, the shorter reference period for these economic data (4 months in SIPP compared to a year in CPS) is probably better for measuring or simulating the relationship between insurance status and economic variables (such as employment or eligibility for Medicaid) that also change over time.

Because they ask retrospectively about insurance over a four-month reference period, the SIPP questions are known to have their own flaws in terms of recall. However, the apparent tendency of respondents to report changes in insurance (and many other variables) in the interview month should not detract from the validity of SIPP as a cross-sectional "snapshot," because most changes in insurance should still be captured from interview to interview.

Probably the biggest factor discouraging use of SIPP is the hit-or-miss fielding of the survey. Not only have there been several lapses when SIPP was not fielded at all, but the panels have varied noticeably in length and by sample size. Despite the potential advantages of SIPP compared to the CPS, analysts have been understandably reluctant to invest in learning how to use a fairly complicated survey that has not been fielded on a consistent basis.

All three of these surveys (CPS, NHIS, and SIPP) are limited in the quality and amount of health care utilization data that analysts can connect to health insurance differences and changes. With its shorter recall periods for collecting utilization over an extended period of time, MEPS is the survey that best captures the effects of insurance on access and utilization. Although the individual panels are relatively small in comparison to CPS, NHIS, or SIPP, extensive oversampling and the possibility of combining overlapping MEPS panels enables analyses involving relatively small, particularly vulnerable population subgroups.

Timeliness is inevitably an issue with MEPS, because of its design. The quick estimates that AHRQ publishes from the round 1 interview with each panel are not as analytically useful as the monthly data that are released much later. The round 1 questions count people who are uninsured for varying lengths of time (3 to 6 months, depending on the time of interview). There are also two potential problems with a reported lack of insurance from interviews conducted toward the end of the round 1 reference period. First, recall error is likely to be significant. Second, because 6 months is a fairly long time in relation to the turnover in the uninsured population, and correct answers to the questions imply that uninsured respondents lacked coverage for the entire 5-6 months, the round 1 estimates undercount uninsured person-years. Furthermore, because the MEPS income questions are not asked until the calendar year has ended, the round 1 estimates cannot be tabulated by income, an essential variable in most policy analyses.

NSAF and CTS are privately funded efforts that are not really aimed at producing regular national estimates of the uninsured. While the designers of both surveys maintained the capability to make reliable national estimates, the NSAF sample is optimized for state-specific estimates and the CTS is optimized for community-specific estimates. The central questions in both studies have to do with comparing estimates across geographic units and across time, rather than making precise estimates of national totals that are important in evaluating current and proposed national policies. To limit the cost of conducting so many interviews in so many different locations, both surveys rely heavily on telephone (instead of personal) interviews, and their samples are primarily based on random digit dialing instead of area probability sampling.

New Directions for Data Collection and Analysis

The estimates from the six different surveys that were shown at the outset of this article in Table 1 suggest that counts of the uninsured cluster into two different ranges. Estimates from NHIS for 1997 (41 million) and MEPS for 1999 (43 million) suggest that the number of uninsured at a point in time in 1999 was in the neighborhood of 43 million. Given the ambiguous time frame for estimates from the CPS, the CPS estimate for 1999 (42 million, revised downward to 39 million with the new confirmation question) seems consistent with point-in-time estimates in this range. However, Fronstin's monthly SIPP estimate (35 million for October 1994 to September 1995) is at odds with the NHIS and

MEPS and is more consistent with the lower national estimates from the two telephone surveys, the NSAF and the CTS (around 35 or 36 million). Given these inconsistencies, the first research priority is to try to corroborate estimates in one of these ranges by producing more current monthly estimates from the 1996 SIPP panel for the years from 1996 to 1999, as well as more recent monthly estimates from MEPS and cross-sectional estimates from NHIS.

Despite their limitations in capturing important insurance dynamics, accurate cross-sectional estimates of the uninsured can be produced more quickly and cheaply than longitudinal estimates. Consequently, cross-sectional estimates and surveys will continue to be important in monitoring and analyzing trends in health insurance, especially when there are policy changes (such as SCHIP) to be evaluated quickly.

Although there are plans to expand the CPS sample to facilitate the evaluation of SCHIP on a state-by-state basis, and analysts often interpret CPS estimates as a cross-section, it is risky to translate annual questions incorrectly answered by respondents into program caseloads and costs. Furthermore, to the extent that respondents answer the CPS questions correctly, the focus of the questions on coverage held at anytime during a year gives relatively little weight to transitions out of coverage, a potentially important consideration in evaluating programs that experience as much turnover as SCHIP and Medicaid. For these reasons, there should be a concerted effort to improve the usefulness and timeliness of truly cross-sectional surveys for policy purposes.

Some of the most fruitful new analyses of cross-sectional data are likely to involve pooling multiple years of data from the same survey. Pooled cross-sections will provide more variation for exploring the determinants of health insurance status than can be found in a single cross-section. Pooled micro data will also be instrumental in evaluating the effects of policy changes and investigating trends over time. It is important for data collection organizations that conduct cross-sectional surveys to maintain the consistency of survey questions and procedures over time, in order to facilitate pooled and trend analyses.

Because of the importance of state policies in determining Medicaid and SCHIP enrollment, the probability of being uninsured is necessarily correlated among observations within the same state. Consequently, analysts need access to state identifiers and need to use statistical techniques that recognize intra-state correlations in cross-sectional studies of insurance status and the uninsured.

A lot of important territory remains to be explored with longitudinal data. As argued above, one cannot accurately assess the personal and social consequences of being uninsured without taking account of the distribution and duration of uninsured spells across people over time. Also, by analyzing longitudinal data for individuals over time, analysts may better succeed in distinguishing causality from correlation. They can observe the ordering of events, model changes in individual behavior, and study outcomes more effectively with longitudinal data.

Historically, the lack of regular and timely longitudinal data has been one of the biggest impediments to studying the causes and consequences of gaps in health insurance over time. That situation may now be improving. AHRQ recently released the second year of data for the 1996 MEPS Panel and is committed to keeping MEPS in the field continuously. Because the median length of an uninsured spell is about half a year, longitudinal panels do not need to be longer than the two years covered by MEPS to be useful in studying spell durations. SIPP data from the 12 core interviews (3 each year) and the first 6 topical modules from the 1996 panel are now available for downloading from the Census Bureau's website. However, the Bureau has not yet released a longitudinal file that compiles selected variables from each of the 12 interviews and includes a longitudinal weight to adjust for sample attrition. A 2001 panel went into the field in February, and the Bureau plans are to field abutting (instead of overlapping) SIPP panels every 3 years.

Future research that takes a dynamic view of health insurance will focus on describing and modeling transitions from one insurance status to another. Instead of static analyses that count and characterize the uninsured in cross-sectional data, these dynamic analyses will count and characterize *changes* in coverage (such as the likelihood and predictors of leaving Medicaid or employer insurance and becoming uninsured). Ideally, these dynamic models will imply flows from one type of coverage to another that are consistent with the stocks of each type in cross-sectional estimates.

Future research will also be able to produce a more definitive picture of the consequences of being uninsured by characterizing the uninsured in terms of the timing and duration of their lack of insurance. Using longitudinal data to distinguish between the long-term and short-term uninsured is an obvious next step in this line of outcomes research. More generally, research that relates individual or family changes in insurance status to changes in utilization, changes in employment, and changes in health status (both in a causal sense and in time) will improve our current understanding of the social and personal costs of being uninsured.

REFERENCES

Andersen, R., Anderson, O.W. (1999). "National Medical Expenditure Surveys: Genesis and Rationale." In A.C. Monheit, R. Wilson, R.H. Arnett III (eds.), Informing American Health Care Policy: The Dynamics of Medical Expenditure and Insurance Surveys, 1977-1996. San Francisco: Jossey-Bass.

Bennefield, R.L. (1996a). A Comparative Analysis of Health Insurance Coverage Estimates: Data From CPS and SIPP. Joint Statistical Meetings, Chicago, IL, Bureau of the Census.

Bennefield, R.L. (1996b). "Dynamics of Economic Well-Being: Health Insurance 1992-1993 Who Loses Coverage and for How Long?" Current Population Reports P70-54 (May): 1-5.

Chollet, D. (2000). "A Survey of Surveys: What Does It Take to Obtain Accurate Estimates of the Uninsured?" State Coverage Initiatives 1(5-8). Washington DC: The Alpha Center.

Cunningham, P.J., Schaefer, E., Hogan, C. (1999). "Who Declines Employer-Sponsored Health Insurance and is Uninsured?" Issue Brief Findings from Health System Change (22).

Cunningham, P.J., Ginsburg, P.B. (2001). "What Accounts for Differences in Uninsurance Rates Across Communities?" Inquiry 38(1): 6-21.

Dubay, L., Kenney, G. (1996). "The Effects of Medicaid Expansions on Insurance Coverage of Children, " The Future of Children 6(1):152-161.

Fronstin, P. (2000a). "Sources of Health Insurance and Characteristics of the Uninsured: Analysis of the March 2000 Current Population Survey." Issue Brief Employee Benefit Research Institute 228.

Fronstin, P. (2000b). "Counting the Uninsured: A Comparison of National Surveys." Issue Brief Employee Benefit Research Institute 225.

Hoffman, C.W. (1964). "Health Insurance Coverage: United States, July 1962-June 1963." Series 10, Number 11. Rockville, MD: National Center for Health Statistics.

Johnson, R.W., Crystal, S. (1997). "Health Insurance Coverage at Midlife: Characteristics, Costs, and Dynamics." Health Care Financing Review 18(3): 123-148.

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

Long, S.H., Marquis, M.S. (1994). Universal Health Insurance and the Uninsured: Effect on Use and Cost. Washington DC: Office of Technology Assessment and Congressional Research Service Report for Congress, 94-689EPW.

Long, S.H., Marquis, M.S. (1996). "Some Pitfalls in Making Cost Estimates of State Health Insurance Coverage Expansions." Inquiry 33(1): 85-95.

Monheit, A.C., Vistnes, J.P. (1997). "Health Insurance status of workers and their families: 1996." MEPS Research Findings No.2 97-0065.

Nelson, C. T, Mills, R.J. (2001). "The March CPS Health Insurance Verification Questions and Its Effect on Estimates of the Uninsured. Paper presented at the annual meetings of the American Statistical Association in Atlanta, August.

Rajan, S., Zuckerman, S., Brennan, N. (2000). "Confirming Insurance Coverage in a Telephone Survey: Evidence from the National Survey of America's Families." Inquiry 37(3): 317-327.

Rhine, S., Ng, Y. (1998). "The Effect of Employment Status on Private Health Insurance Coverage: 1977 and 1987," Health Economics 7(1): 63-79.

Rhoades, J., Chu, M. (2000). Health Insurance Status of the Civilian Noninstitutionalized Population: 1999. MEPS Research Findings No. 14, AHRQ Pub. No. 01-0011. Rockville, MD: Agency for Healthcare Research and Quality.

Selden, T.M., Banthin, J.S., Cohen, J.W. (1998). "Medicaid's Problem Children: Eligible But Not Enrolled." Health Affairs 17(3): 192-200.

Short, P.F., Freedman, V.A. (1998). "Single Women and the Dynamics of Medicaid." HSR: Health Services Research 33(5): 1309-1336.

Short, P.F., Klerman, J. (1998). Targeting Long-and Short-Term Gaps in Health Insurance. New York: The Commonwealth Fund, Report #259.

Short, P.F., Shea, D., Powell, M.P. (2001). Health Insurance on the Way to Medicare: Is Special Government Assistance Warranted? New York: The Commonwealth Fund, Report #472.

Short, P.F., Vistnes, J.P. (1992). "Multiple Sources of Medicare Supplementary Insurance." Inquiry 29: 33-43.

Short, P.F. (2000). "Hitting a Moving Target: Income Related Health Insurance Subsidies for the Uninsured." Journal of Policy Analysis and Management 19(3): 383-405.

Sloan, F.A., Conover, C.J. (1998). "Life Transitions and Health Insurance Coverage of the Near Elderly." Medical Care 36(2): 110-125.

Summer, L., Parrott, S., Mann, C. (1997). Millions of Uninsured and Underinsured Children Are Eligible for Medicaid. Washington DC: Center of Budget and Policy Priorities.

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.

Swartz, K. (1994). "Dynamics of People Without Health Insurance: Don't Let the Numbers Fool You," *JAMA* 271(1): 64-66.

Swartz, K., McBride, T.D. (1990). "Spells without Health Insurance: Distributions of Durations and Their Link to Point-in-Time Estimates of the Uninsured, *Inquiry* 27(Fall): 281-288.

Swartz, K., Marcotte, J., McBride, T.D. (1993a). "Personal Characteristics and Spells Without Health Insurance." Inquiry 30(1): 64-76.

Swartz, K., Marcotte, J., McBride, T.D. (1993b). "Spells Without Health Insurance: The Distribution of Durations When Left-Censored Spells are Included." Inquiry 30(1): 77-83.

Swartz, K. (1997). "Changes in the 1995 Current Population Survey and Estimates of Health Insurance Coverage." Inquiry 34(1): 70-79.

Walden, D., Wilensky, G, Kasper, J. (1985). Changes in health Insurance Status: Full-year and Part-Year Coverage. NHCES Data Preview 21; DHHS Pub. No. PHS 85-3377. Rockville MD: National Center for Health Services Research and Health Care Technology Assessment.

NOTES

[1] Health insurance estimates from the HRS, together with information about the HRS survey design, are available from Short, Shea, and Powell (2001), Sloan and Conover (1998), and Johnson and Cystal (1997).
[2] The undercount of Medicaid in the Current Population Survey been discussed fairly widely (Fronstin 2000b; Chollet 2000; Lewis, Ellwood, and Czajka 1998). Lewis, Ellwood and Czajka report that there is also evidence of this problem in other surveys.
[3] When Medicaid was automatically linked to the welfare payments for a family, family eligibility for Medicaid was another reason for studying the insured and uninsured within families.
[4] Similarly, estimates from MEPS for 1996 indicate that only 5% of workers ages 16-64 who were offered insurance were uninsured (Monheit and Vistnes 1997).
[5] There are a few examples of such multinomial models in the literature, including Rhine and Ng (1998), Johnson and Crystal (1997), and Short and Friedman (1998).
[6] In reality, these calculations are only the starting point in costing proposed programs. Cost estimates incorporate additional adjustments to account for less than full participation in public coverage programs among the uninsured, for example, and for incentives to substitute enrollment in the new program for existing coverage among the insured. Adjustments to cross-sectional estimates may also be necessary to account for the frequency and timing of eligibility determinations. If a new program grants coverage for a specified time period, such as 6 months, its coverage will extend beyond the currently observed ending of some uninsured spells and will cover currently insured person-years as a result (Kathy Swartz, personal communication).
[7] In reality, so-called cross-sectional surveys do not collect information about the health insurance status of each sampled person on exactly the same day. Rather, they typically ask about the status of each sampled person on the day of the interview (or "now," in the wording of the survey questions). Although everyone in the sample is not literally interviewed on the same day, the survey still characterizes the status of each person on one day. Because one day is the same as another in the absence of seasonalities, the logic described in the text still applies.
[8] For example, MEPS asks when each person uninsured at the start of the calendar year was most recently covered by health insurance and the source of the person's prior insurance. NHIS asks how long it has been since each uninsured person last had health care coverage. NHIS also asks if there was ever a time in the past 12 months when each insured person did not have any health insurance or coverage and, if so, for how many months.
[9] For example, see Swartz, Marcotte and McBride (1993a), Bennefield (1996b), and Short and Friedman (1998).
[10] When the Census Bureau experimented with questions about current insurance status in the March 1995 CPS, there was a discrepancy of about 10 percentage points between the

cross-sectional estimates and those based on questions about the preceding year (Bennefield 1996a).

[11] Even if the absolute numbers from the CPS do not have a lot of meaning, because of ambiguities about the timing of the coverage reported in the CPS, it is possible that the CPS is still reliable in measuring relative differences across the population or over time. That is to say, the estimates can still be reliable, even if they are not valid. However, because the reporting errors in the CPS are associated with changes in insurance status, the CPS is unlikely to be reliable in comparing subgroups or time periods where the mix of short-term and long-term uninsured differs or changes over time.

Table 1. Estimates of the Uninsured Under Age 65 from Various Surveys

Source	Year	Uninsured	Sample size	Timeframe of uninsured estimate	Location identifiers	Other limitations
CPS[a]	1999	42 million	131,000	Uninsured throughout calendar year	All states	Some states have small sample sizes
	1998	44 million	116,000	Uninsured throughout calendar year	All states	Some states have small sample sizes
	1995	40 million	118,000	Uninsured throughout calendar year	All states	Some states have small sample sizes
SIPP[b]	1994	19 million	47,000	Uninsured throughout calendar year	45 states & DC	Many states have small sample sizes
	10/94-9/95	35 million	47,000	Average monthly number of uninsured	45 states & DC	Many states have small sample sizes
MEPS[c]	1999	43 million	22,000	First 3-6 months of year	None	
	1998	42 million	13,000	First 3-6 months of year	None	
	1996	45 million	24,000	First 3-6 months of year	None	
	1996	32 million	24,000	Uninsured throughout calendar year	None	
NHIS[d]	1997	41 million	103,000	Uninsured at time of interview	All states	Many states have small sample sizes
CTS[e]	7/98-7/99	36 million	60,000	Uninsured at time of interview	60 communities	
	7/96-7/97	35 million	60,000	Uninsured at time of interview	60 communities	
NSAF[f]	1997	36 million	110,000	Uninsured at time of interview	13 states	
	1999	36 million	110,000	Uninsured at time of interview	13 states	

[a]Current Population Survey; [b]Survey of Income and Program Participation; [c]Medical Expenditure Panel Survey; [d]National Health Interview Survey; [e]Community Tracking Survey; [f]National Survey of America's Families

Adapted and updated from P. Fronstin, *Counting the Uninsured: A Comparison of National Surveys* (Table 1, p.4). EBRI Issue Brief Number 225. Washington, DC: Employee Benefits Research Institute, 2000.

Table 2. Number and characteristics of the uninsured under age 65 according to different time frames (United States, 1996-1997)

	Uninsured throughout 1996-1997			Uninsured throughout 1996			Average monthly uninsured 1996			Ever uninsured in 1996			Ever uninsured in 1996-1997		
	Millions	Rate[1]	Rel. risk[2]	Millions	Rate[3]	Rel. risk	Millions	Rate[4]	Rel. risk	Millions	Rate[3]	Rel. risk	Millions	Rate[1]	Rel. risk
Total	23.5	9.9%		31.6	13.5%		45.0	19.4%		64.0	27.3%		80.2	33.7%	
Age															
Under 18	4.5	6.2	1.00	7.0	9.8	1.00	10.9	15.7	1.00	17.4	24.4	1.00	22.1	30.7	1.00
18-24	3.5	13.4	2.16	5.6	22.3	2.28	8.3	33.0	2.10	11.6	45.8	1.88	13.6	52.4	1.71
25-54	13.6	11.5	1.85	16.8	14.3	1.46	22.9	19.6	1.25	31.2	26.7	1.09	39.3	33.3	1.08
55-64	2.0	9.1	1.47	2.2	10.7	1.09	2.9	13.9	0.89	3.8	18.3	0.75	5.2	23.4	0.76
Race/ ethnicity															
White	12.4	7.5	1.00	16.9	10.3	1.00	24.9	15.2	1.00	36.5	22.2	1.00	46.9	28.3	1.00
African American	3.5	11.1	1.48	5.2	16.8	1.63	7.6	24.8	1.63	11.1	36.0	1.62	13.3	42.0	1.48
Hispanic	6.5	22.0	2.93	7.8	27.5	2.67	10.2	36.6	2.41	13.3	46.8	2.11	15.8	53.3	1.88
Other	1.1	10.1	1.35	1.7	15.9	1.54	2.4	22.9	1.51	3.1	29.2	1.32	4.2	37.1	1.31
Percent of poverty															
<100	5.5	16.5	3.67	8.0	23.4	4.42	11.1	33.2	4.00	15.4	45.1	3.44	17.2	51.5	2.72
100-125	1.8	18.4	4.09	2.6	25.5	4.81	4.1	40.4	4.87	5.5	53.9	4.11	6.1	61.3	3.24
125-199	6.1	18.8	4.18	8.0	24.2	4.57	10.7	32.7	3.94	14.5	43.8	3.34	16.3	50.6	2.68
200-399	6.4	8.0	1.78	8.8	11.4	2.15	12.5	16.4	1.98	18.1	23.6	1.80	24.8	31.3	1.66
400+	3.8	4.5	1.00	4.3	5.3	1.00	6.7	8.3	1.00	10.5	13.1	1.00	15.8	18.9	1.00

[1]Denominator is all persons in the civilian noninstitutionalized population in 1996 or 1997 who were under age 65 in 1997. [2]Relative risk is the ratio of uninsured rates. [3]Denominator is the civilian noninstitutionalized population under age 65 in 1996. [4]Denominator is total person-years in the civilian non-institutionalized population under age 65 in 1996.

Source: Author's tabulations of the 1996 MEPS Panel.

Table 3. Current features of six national surveys with health insurance data

	CPS	NHIS	MEPS	SIPP	CTS	NSAF
Organization managing the survey	Census Bureau	National Center for Health Statistics	Agency for Healthcare Research & Quality	Census Bureau	Center for Studying Health System Change	The Urban Institute
Survey design	March supplement is a cross-section[1]	Cross-section	Panel, 5 interviews over 2 years	Panel, every 4 months over 3-4 years	Cross-section	Cross-section
Mode	In person and telephone	In person	In person and telephone	In person and telephone	Telephone, supplemented in person	Telephone, supplemented in person
Response rate	86% (2000)	88% (1999) 90% (1998)	75% Round 1 (1998) 68% Annual (1998)	64.5% (1996), all interviews	65%	65% (1997) 60% (1997)
			Sample design			
Universe	Civilian non-institutionalized pop.	Civilian non-institutionalized pop.	Civilian non-institutionalized pop.	Civilian non-institutionalized pop.	Civilian non-institutionalized pop., excluding Alaska and Hawaii	Civilian non-institutionalized pop. under age 65
Sample frame	Area probability	Area probability	NHIS	Area probability	Random digit dialing (supplemented by area probability)	Random digit dialing (supplemented by area probability)
Sample size	50,000 households 130,000+ people	Varies around 40,000 households 100,000+ people	7-13,000 households 15-35,000 people[2]	37,000 households	60,000 people	44,000 households 100,000 people
Oversampling	Hispanics	Blacks, Hispanics	Blacks, Hispanics, disabled, low income, high expense, elderly	Low income	For community-specific estimates	For state-specific estimates Low income
Individual data	All members of household	Health insurance - all members of household Some topics - 1 adult and 1 child	All members of household	All members of household	All adults members of family insurance unit 1 child per unit	1 child under 6 1 child 6-17 Parents of sampled children Sample of

	CPS	NHIS	MEPS	SIPP	CTS	NSAF
						childless adults
Location-specific estimates	All states	27 states	No state estimates	Some states	60 communities	13 states
			Health insurance questions			
Timeframe	Preceding calendar year	Time of interview, throughout year, ever in year	Since start of calendar year or last interview; monthly	Last 4 months; monthly	Time of interview	Time of interview
Respondent	Household informant	One person familiar with family's health coverage	Family informant	Self reporting for adults (15+)	Family informant	One spouse reports for parents and children. Self-reporting or proxy for other adults.
Logical imputation	Medicaid of adults is attributed to their children	Medicaid attributed to AFDC/SSI recipients until 1996	Minimal	Similar to CPS	Minimal	Medicaid attributed to TANF recipients
Catch all question	Asked	Not asked	Asked	Asked	Asked	Not asked
Confirm uninsured	Added in 2000	Added in 1997	Reasons why uninsured[3]	Reasons why uninsured[3]	Asked	Asked
Eligibility for employer insurance	Not asked in March Supplement	Asked	Asked	Asked	Asked	Asked
Status before survey	Not asked	Asked	Asked	Asked	Asked	Asked
			Other data			
Health care use	Not asked	2-week and 12-month recall	4-5 month recall	Selected interviews 12-month recall	12-month recall	12-month recall
Employment data	Extensive	Limited for all adults.[4] Additional detail for one adult.	Extensive	Extensive	Extensive	Extensive

	CPS	NHIS	MEPS	SIPP	CTS	NSAF
Data availability						
Health insurance time series	Annual since 1980	Annual since 1989, selected years 1960-1986	Annual since 1996 1977, 1980, 1987	Panels cover most years since 1983	1996-97 1998-99	1997, 1999
Latest published tables (as of August 2001)	P-60 report with data for 1999, September 2000	NCHS's *Health United States* with 1999 data, planned Aug. 2001	Annual AHRQ report with R1 estimates for 2000, July 2001	P-70 report with data for 1992-1993, May 1996	*Issue Brief 29,* 1998/1999 for children, April 2000	*1999 Snapshots of America's Families II*, October 2000
Latest micro data (as of August 2001)	2000 survey with 1999 data, released in 2000	1999 release planned Aug. 2001	1998 annual file 2000 R1 file planned Aug. 2001	1996 core files, 6 of 12 topical modules	1996-1997	1997 1999 release began Summer 2001

[1] Each CPS panel is interviewed for 4 months and then interviewed in the same 4 calendar months a year later. [2] Panels alternate between larger and smaller sample sizes. Two panels can be combined for 1997 and all subsequent years. [3]Lack of insurance is not explicitly confirmed, but the uninsured are asked for "reasons why uninsured." [4]Employment status, hours, earnings, insurance eligibility for all adults

Figure 1. Measuring and describing the hole left by public and private insurance

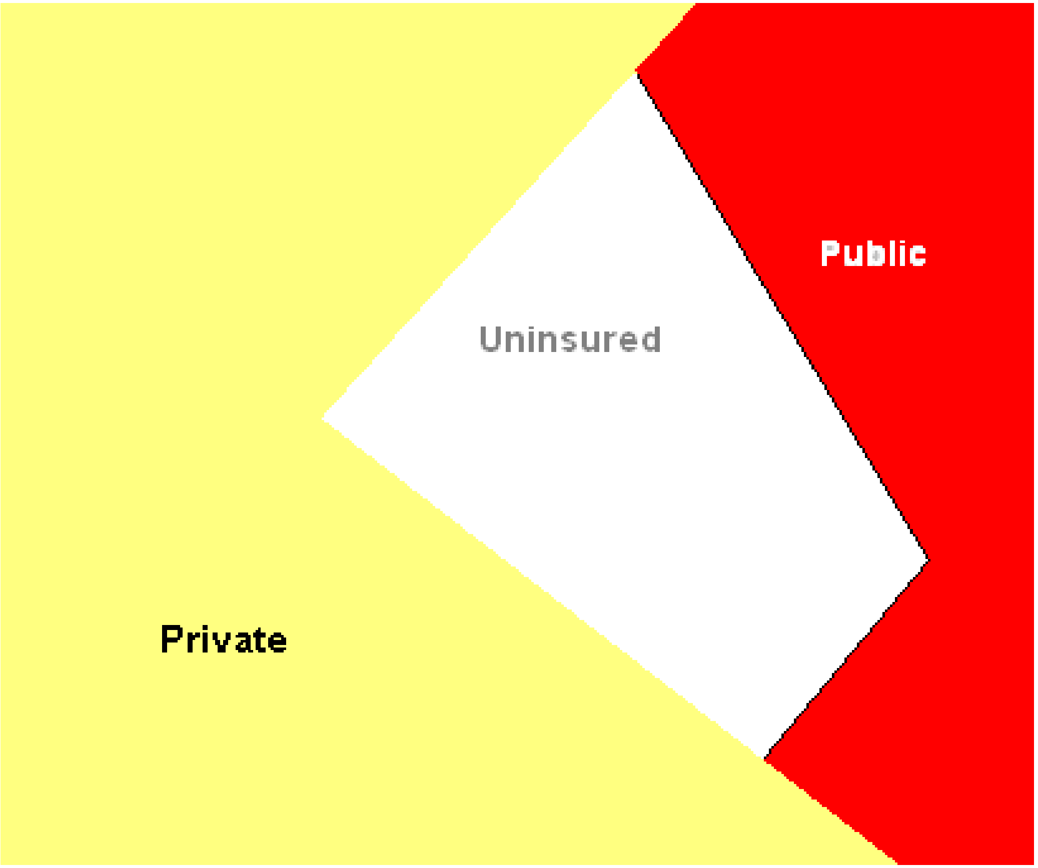

Figure 2. Percent distribution of insurance status by family income

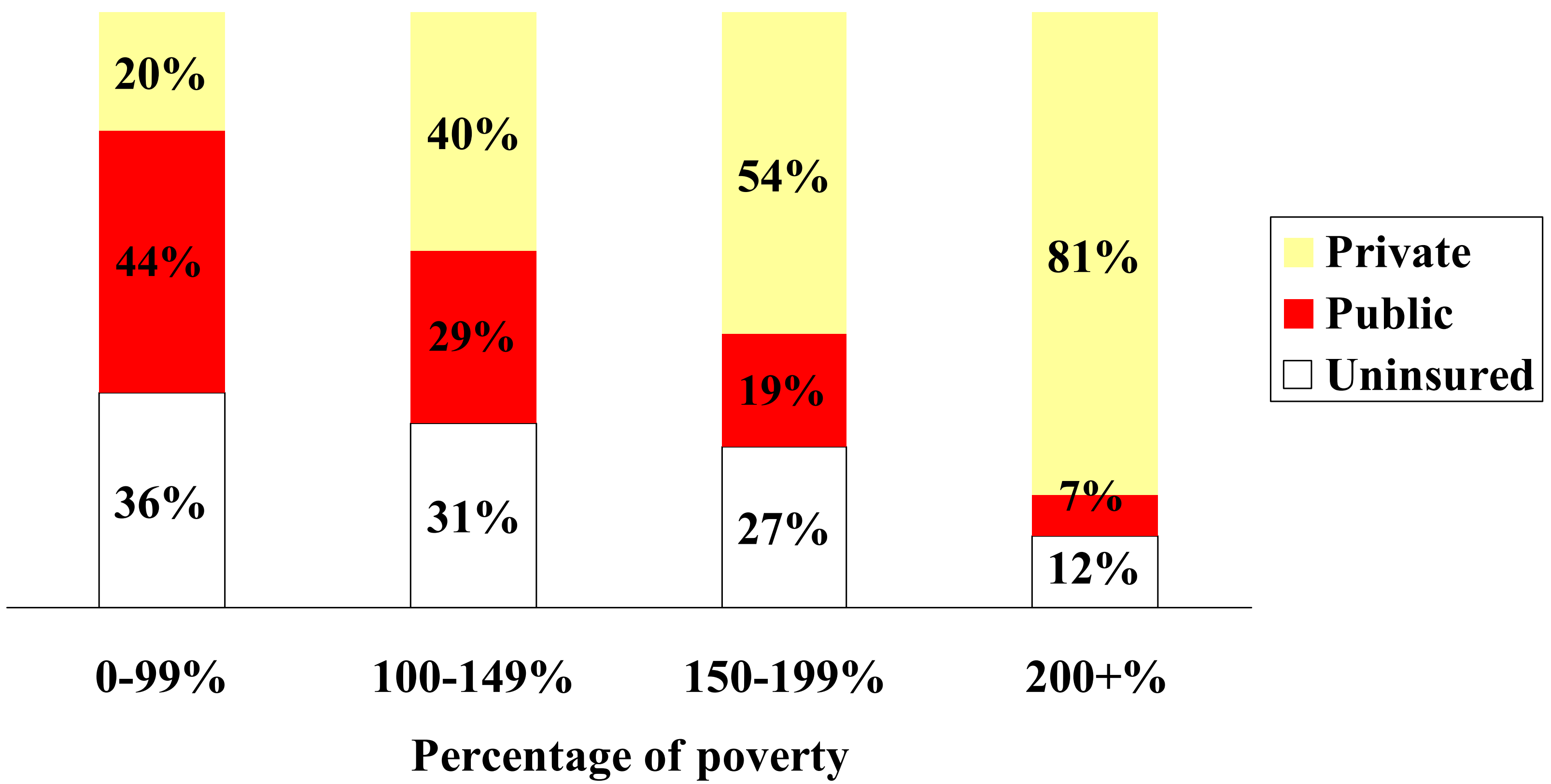

Figure 3. Changes in Insurance Status Over Time